

Stress  
in  
modern Welsh

Dissertation presented for the degree of Doctor of Philosophy of the  
University of Cambridge, September 1983.

UNIVERSITY  
LIBRARY  
CAMBRIDGE

Briony J. Williams

Selwyn College, and  
Department of Linguistics,  
University of Cambridge.

September 1983

Briony Jane WILLIAMS

### Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration.

This dissertation does not exceed 80,000 words in length.

*Briony Williams*

*21st September 1983*

### Acknowledgements

I should like first and foremost to thank my supervisor, Dr. Francis Nolan, for his invaluable advice and for assistance in practicalities even when inconvenient. Thanks are also due to Nigel Vincent for the first germs of an idea, to Richard Baron for clarification of a statistical point, and to Mike the technician for instruction in the operation of some equipment. Any errors or omissions in the thesis are, of course, entirely my own responsibility.

Stress in modern Welsh: abstract of Ph.D. dissertation

Briony J. Williams

September 1983

Selwyn College and Department  
of Linguistics, University of Cambridge.

The first chapter reviews twentieth-century work on stress, considering stress from several points of view: as intensity, pitch and rhythm. The weaknesses of some approaches are pointed out. Some past work on isochrony is summarised and the metrical theory of stress introduced. The problem presented by stress in Welsh is introduced informally.

The second chapter describes a series of measurements and experiments carried out to investigate stress in Welsh. Acoustic measurements are made of  $F_0$ , intensity, duration, and lengths of rhythmic feet. It is concluded that stress in Welsh is determined by rhythm rather than inherent acoustic cues. Statistical analysis of further measurements of feet reveals the effects of syllabification and also supports the previous findings. Perceptual experiments using native speaker judgements of resynthesised speech are presented and analysed; it is concluded that Welsh stress is linked with rhythm and with consonantal length (the consonant after a stressed vowel being longer than one after an unstressed vowel), while it has no direct link with pitch-prominence as such.

The third chapter outlines a theoretical framework based on the 'metrical' theory of stress, and proceeds to apply it to Welsh. This modified framework makes use of Selkirk's 'prosodic units', but differs from the usual forms of the theory by also taking account of the semantic weight of words in the utterance context. It is agreed that this makes for a more explanatory theory, and some puzzles of English intonation are shown to be described more simply using such a theory. The patterns of intonation in Welsh are also described in this framework.

The fourth chapter discusses the implications posed by the experimental results for the general theory of stress. Reference is made to some research in the psychology of stress perception, and evidence is also presented from the historical development of Welsh from ancient British times, and also from the Welsh bardic rules of versification. It is argued that Welsh has a unique contribution to make to our understanding of the nature of and possible forms taken by stress, and the conclusion is that stress in any language, and more particularly in English, may not be as straightforward a matter as has been hitherto assumed.

## CONTENTS

1. Introduction	1
1.1. Approaches to the definition of stress	1
1.1.1. Measurable acoustic cues	1
1.1.2. Perceived loudness	2
1.1.3. Lexical stress versus sentence stress	3
1.1.4. Philological/phonological view	3
1.2. Past work on stress	4
1.2.1. Stress as primarily psychological reality	4
1.2.2. Stress as intensity	6
1.2.3. Stress as pitch	8
1.2.4. Stress as accent	13
1.2.5. Stress as a component of rhythm	16
1.2.6. Stress as rhythm: the "metrical" theory of stress	20
1.3. Informal observations of Welsh stress	23
1.3.1. Welsh stress rule	23
1.3.2. Reaction of English listeners	23
1.3.3. Robustness of the ultima	24
1.3.4. Weakness of the penult	25
1.4. Plan of the thesis	26
2. Experimental work	28
2.1. Words in citation form	28
2.1.1. Aim	28
2.1.2. Method	28
2.1.3. Results	30
2.1.4. Discussion	31
2.2. Connected speech	32
2.2.1. Unrehearsed speech: segmental measurements	32
2.2.2. Unrehearsed speech: foot measurements	36
2.2.3. Rehearsed speech	39
2.2.4. Segmental measurements	44
2.3. Perceptual experiments	47
2.3.1. Post-stress consonant	47
2.3.2. Phonological segmental length	55
2.3.3. Effect of pitch-prominence	58
2.3.4. Comparison of perceptual experiments	67

2.4. Experimental findings	80
2.4.1. Summary of experimental results	80
2.4.2. Implications of the results	84
3. Theoretical considerations	88
3.1. Syllable level	88
3.2. Word level	92
3.2.1. Syllable, or ' $\sigma$ '	92
3.2.2. Stress foot, or ' $\Sigma$ '	94
3.2.3. Prosodic word, or ' $\omega$ '	98
3.3. Semantic level	104
3.3.1. Thompson's features	104
3.3.2. Matrix versus grid	106
3.3.3. A semantic grid	107
3.3.4. Semantic grid versus n-ary features alone	109
3.3.5. A possible solution	116
3.3.6. Semantic and metrical grids	119
3.3.7. Some advantages of this type of analysis	121
3.3.8. [content] and the prosodic representation	124
3.3.9. The need to take account of semantics	125
3.4. Metrical structure (phrasal level)	127
3.4.1. Selkirk's units	127
3.4.2. Discussion	130
3.4.3. Intonation	135
3.4.4. Building up the tree	141
3.4.5. Selkirk's Monosyllable Rule	142
3.4.6. Examples	148
3.5. The metrical grid	151
3.5.1. Introduction	151
3.5.2. Description	152
3.5.3. Use of the metrical grid	154
3.5.4. The need for the metrical grid	159
3.5.5. Stress shifting	163
3.6. Some intonational considerations	172
3.6.1. Congruence	172
3.6.2. 'Neutral' pattern	173
3.6.3. Compound nucleus	174

3.6.4. Complex nucleus (i)	178
3.6.5. Complex nucleus (ii)	180
3.6.6. Stressed nonlexical word	182
3.7. Location of [3 content]	183
3.7.1. Neutral pattern	183
3.7.2. Contrastive stress	183
3.7.3. Emphatic stress	184
3.7.4. Discussion	186
3.8. Application of the system to Welsh	190
3.8.1. Vowel level	190
3.8.2. Word level	190
3.8.3. Semantic level	202
3.8.4. Phrasal level	203
3.8.5. Metrical grid	209
3.8.6. Intonational considerations	210
4. Discussion	221
4.1. Penultimate stress in Welsh	221
4.1.1. Phonetic mechanisms	221
4.3.2. Penultimate stress; non-linguistic evidence	237
4.3.3. Final stress; non-linguistic evidence	238
4.4. Implications for the theory of stress	241
4.4.1. Two trends	241
4.4.2. Stress accent and pitch accent	243
4.4.3. The notion of stress	244
4.4.4. Uncertainty in the description of stress	246

## 1. INTRODUCTION

### 1.1 Approaches to the definition of stress

The word 'stress' has been used in so many different senses in phonetic and phonological literature that it is a far from simple task to unravel the knot of meanings and pick out the particular strand that is being used at any one time. The strands are of widely varying types, some blending into neighbouring shades more readily than others, and the only point on which there is unanimity is that stress exists. Individual views of stress have been researched in depth and advocated with conviction, leading to a true entanglement of the various strands; it may be that our understanding of stress will only be complete when we can unpick and make use of each one of these views in an interwoven whole.

Different types of approach to stress are described briefly below, and there follows a discussion of past work on the subject. Nearly all work in this area has been done using English, and this should be constantly borne in mind; the conclusions drawn need not be relevant to other languages.

#### **1.1.1 Measurable acoustic cues**

Much useful work on the nature of stress has concentrated on its physical realisation in ways that can be quantified by machines such as the sound-spectrograph and, nowadays, the computer with wave-form display. The relevant acoustic parameters are: fundamental frequency (henceforth  $F_0$ ), intensity and duration. These approximately correspond to perceived pitch, loudness and length - approximately, since there are interaction effects between the cues such that what the brain registers may not necessarily be what the machine records.

These measurements are carried out on a syllable-by-syllable basis, the researcher concentrating on the inherent properties of a particular spoken syllable or vowel that cause it to be heard as stressed or unstressed, as the case may be. The overall structure of the utterance is not considered, and stress is seen as a binary feature, while the search is for an acoustic cue or cues that can give as nearly as possible 100% accuracy in identifying stressed syllables. This 'absolutist' approach was popular when the sound-spectrograph first became widely used, and much work was done on inherent acoustic cues to stress; however, it is less influential nowadays, since the focus has shifted to the formulation of more satisfactory theories to account for the acoustic facts discovered.

### 1.1.2 Perceived loudness

On the whole, loudness is the perceptual correlate of the acoustic cue of intensity (see section 1.2.2.2. for details of work in this area). The advantage of the view of stress as perceived loudness is that this leaves untouched the cue of  $F_0$ /pitch variation, so that the latter can be reserved to figure in the description of intonation patterns. Both the 'British' and the 'American' schools of thought on suprasegmentals separate stress from intonation, and the identifying of stress with perceived loudness is a popular means of doing this (see Ladd 1980, Chapter 1, for a clear summary of these two schools of thought). While loudness is the auditory correlate of the acoustic cue of intensity, the latter is in its turn the acoustic correlate of the articulatory/physiological cue of greater muscular effort - in this case, in the lungs. This greater effort leads to heightened subglottal pressure, which may cause a rise in the frequency of vibration of the vocal folds quite apart from any muscular adjustments in the larynx itself. The so-called 'larynx versus lungs' controversy (see section 1.2.3.1.) centred on the question of whether laryngeal adjustments or increased subglottal pressure was the cause of the  $F_0$  rise observed on stressed syllables in English.

### 1.1.3 Lexical stress versus sentence stress

Every polysyllabic word in English has a fixed place for stress (ignoring stress shifting for the present - see section 3.5.5.). This one stressable syllable is stressed when the word is spoken in citation form; this is known as 'lexical stress'. Similarly, every phrase or short sentence has one syllable of especial prominence, usually near the end of the string; this is known as 'sentence stress'. In addition, before the most prominent syllable in a sentence, some syllables with a lesser degree of stress may be heard; however, not all the lexically-stressed syllables in the sentence need actually be made prominent in this way in a particular utterance.

The different grades of stress may lead to confusion as to which type of stress is being discussed on any given occasion, and may tend to class together dissimilar phenomena. The term 'sentence stress' is best abandoned in favour of the British-school term 'nucleus', which allows the influence of intonation as well as of stress. Section 1.2.4. presents Bolinger's rather different solution to the problem of distinguishing between lexical and sentence stress.

### 1.1.4 Philological/phonological view

In the historical development of any language, some sound changes occur in stressed syllables and others in unstressed syllables. The former sound changes are usually of the 'strengthening' type, and involve e.g. for vowels: diphthongisation, movement towards a more peripheral vowel quality, and lengthening; and for consonants: gemination, mutation, and insertion of a new consonant. The changes occurring in unstressed syllables include e.g. for vowels: monophthongisation, movement towards schwa in quality, shortening or deletion; and for consonants: simplification of clusters, lenition, and deletion of intervocalic consonants. The overall effect is that stressed syllables are more auditorily distinct and easier to perceive.

In the synchronic description of any language, some phonological processes occur in stressed syllables and others in unstressed syllables. The processes are of the same type as those described above. Thus it is

possible to define stress, for both diachronic and synchronic purposes, purely on distributional grounds, without appealing to phonetics to provide a physical basis to the distinction stressed/unstressed.

The philological/phonological view described above will not be further developed in this study. It may be suitable for the abstract phonological analysis of a language, but in the end it is not an explanation, merely a restating of the problem; it offers no reasons why particular qualities of segments should be better exponents of stress than others, nor why particular syllables should be picked out to receive stress. Neither does it explain the effect of the string of stressed and unstressed syllables on the listener, who is able to extract from the signal the perceptual cues to stressed syllables. Though the listener may also draw on his or her own knowledge of the language, this point will not be further considered. The mere fact that there exist certain cues to stress in the speech signal makes it unlikely that knowledge of the language is all that is required to apportion stress correctly; therefore this study contains a lengthy consideration of the physical aspect of stress. The philological/phonological view treats stress as 'given'; in this study, stress is treated as the unknown quantity and the problem is approached from first principles.

## 1.2 Past work on stress

### **1.2.1 Stress as primarily psychological reality**

Of the 'physical' definitions of stress, the one that verges most nearly on the 'abstract' is the view of stress as existing in the brain as a psychological reality. The earliest modern proponent of this view is André Classe, who states:

'Stress may be defined as an impulse (primarily of a psychological nature) which expresses itself in the first place by an increase of pressure in the speech-canal and approximately coincides with the point of greatest pressure.' (Classe 1939:37)

While Classe does not deny the relevance of physiological cues to stress, he asserts the primacy of the physiological dimension in the definition of stress. A similar view is taken by Gimson, who writes:

'...stress certainly exists in English in terms of a mental pulse or beat, measurable perhaps in the nervous activity of the brain.'  
(Gimson 1956:149)

Gimson likewise states that this 'mental beat' may be realised by extra articulatory effort or greater loudness; however, he sees the psychological dimension as primary. More recently, Chomsky and Halle have attempted to place stress within the native speaker's linguistic 'competence', writing:

'...representations of stress contours and similar predictable phenomena correspond, up to a point, to some perceptual reality that can be brought to consciousness with training and care.' (Chomsky and Halle 1968:26)

Since 'competence' for Chomsky is a psychological reality, it follows that he also regards stress as primarily psychological. Unlike Classe and Gimson, however, Chomsky and Halle have not the slightest interest in the physical manifestation of stress, preferring to concentrate on its functioning in the abstract linguistic system as a whole. In fact, they have serious doubts as to whether phonetic representations of stress describe any kind of physical or acoustic reality at all. Their favoured view is that the native speaker (and native speaker phonetician) depends purely on his or her knowledge of the language when picking out stressed syllables. This is an extreme view. Most workers on stress would allow the physical facts at least some space in the description, while a few would reject entirely the notion that listeners could impose their own structuring on the speech signal, and would attempt to devise stress-locating algorithms that a computer might use. For English and many other languages, some degree of correlation of stress with acoustic and physiological phenomena has been discovered. Accordingly, the research using this line of approach will be outlined below, and the 'abstract' approach will not be further considered.

## 1.2.2 Stress as intensity

### 1.2.2.1. Production aspect

One very old tradition treats stress as an increase in intensity on a particular syllable, caused by extra effort in the lung musculature and leading to greater perceived loudness. This approach first finds a clear formulation in the work of Stetson (1928) whose research, though somewhat unreliable, sparked off much useful work in the area. He states that each syllable is associated with a 'chest pulse', or contraction of the expiratory muscles, while stressed syllables receive an especially 'heavy stroke', which leads to a raising of the  $F_0$  of vibration of the vocal folds. He writes:

'...the heavy stroke of the accent involves the chest pressure and is apt to change the pitch because the laryngeal musculature is often affected by tensions in the other musculatures of speech.'  
(Stetson 1928:141)

This particular assertion is based on no supporting evidence and has subsequently been found to be most improbable. Since the adoption of more sophisticated experimental techniques, phoneticians have found that it is not the case that every syllable has a concomitant 'chest pulse'. Ladefoged's experiments led to his concluding that '...every stress is accompanied by an extra increase of subglottal pressure' (Ladefoged 1967:46); in other words, Stetson's 'chest pulse' was occurring on stressed syllables alone.

This would have been an ideal candidate for the physiological definition of stress, were it not for later experiments by Ohala and by Van Katwijk, who both concluded that the only significantly greater-than-normal decreases in lung volume occurred during those stressed syllables that were emphatically stressed (Van Katwijk 1974, cited by Ohala 1977). Van Katwijk, working with Dutch speakers, and Ohala, working with English speakers and a Swedish speaker and using a whole-body pressure plethysmograph, found no appreciably greater-than-normal decrease in lung volume on non-emphatically stressed syllables. So it seems that the action of the respiratory muscles cannot give a reliable cue to stress.

This conclusion would disappoint the large number of earlier workers who identified stress with greater muscular effort, though without performing physiological experiments to investigate the assumption. The earliest statement of this position within the modern tradition of phonetics is probably that of Sweet, who defined stress as the 'comparative force with which the separate syllables of a sound-group are pronounced' (Sweet 1878: paragraph 263). The same view was taken by subsequent scholars such as Abercrombie (1923:19), Palmer and Blandford (1924:5), Bloomfield (1933:110-111), Jones (1950 3rd. edn.:137), Trager and Smith (1951), Kingdon (1958:1), O'Connor and Arnold (1961, 1973 2nd. edn.:287), Lieberman (1967:144), Catford (1968:317), and Allen (1973:77). A recent, and more sophisticated statement of the same position is that of Sommerstein, who writes:

'...the hearer is using all available information to determine the degree of muscular effort with which each syllable was uttered, and what we call stress corresponds not to the data he uses or any part of it, but to the conclusions he reaches.' (Sommerstein 1977:36)

Such a view may reflect the results of modern research which rejects a simplistic view of increased muscular effort leading directly to increased intensity as the sole cue to stress. However, the primary factor in the definition of stress is still seen as muscular effort, the 'production aspect' of increased intensity. Some workers in this field, on the other hand, consider the perceptual correlate - increased loudness - to be the primary cue to stress, and are less concerned about the physiological or articulatory origins of this increased loudness.

#### 1.2.2.2. Perception aspect

The view of stress as mainly increased muscular effort implies a consequent increase in loudness, but this is seen merely as a by-product. Some, however, view the increase in loudness as the primary criterion of stress, for example Pike (1945: section 4.4.7.), Malmberg (1963:80), and Crystal, whose definition of stress is: 'Perceivable increase in loudness accompanied by unmarked pitch movement' (Crystal 1969:158) - i.e. an increase in intensity that is not also intonationally prominent. This reference to pitch movement reflects the general consensus that, of the acoustic correlates of stress discovered for English, Fo/pitch variation

is by far the most reliable, though by no means the only, cue. The next section describes some of the experimental results that have led to the adoption of this view.

### 1.2.3 Stress as pitch

#### 1.2.3.1. Production aspect: the "larynx versus lungs" controversy

The question of whether  $F_0$  changes were caused by the laryngeal muscles ('larynx' view) or by the pulmonary muscles leading to an increase in subglottal pressure ('lung' view), caught the attention of phoneticians after the publication of a version of Lieberman's doctoral thesis in 1967 as the book 'Intonation, Perception, and Language'. In this book Lieberman posits the 'archetypal normal breath-group' as being characteristic of those patterns used to delimit the boundaries of unemphatic, declarative sentences in normal speech. This type of breath-group, in American English at least, shows a fall in both  $F_0$  and amplitude at the end of a sentence. In yes-no questions, the pitch rises at the end of a sentence, and Lieberman concedes that this pitch variation is caused by changes in the tensioning of the laryngeal muscles. For all other pitch changes, however, he firmly upholds his theory that their sole source is the pulmonary musculature, writing: 'The speaker simply maintains about the same laryngeal tension throughout the entire expiration' (Lieberman 1967:26). The effect of this constant laryngeal tension (in American English), states Lieberman, is that there is no compensation for the fall in subglottal pressure (henceforth s.g.p.) at the end of an utterance, and thus the  $F_0$  also falls; while pitch rises are caused by the peak in s.g.p. near the end of the breath-group. [1]

This view has been challenged since, especially by experimental phoneticians who question Lieberman's assumption that the laryngeal

---

[1] In British English, he writes, the tension of the laryngeal muscles is gradually relaxed in a controlled way through the breath-group, while the s.g.p. rises and then falls as in the American pattern. The net effect is of initial high pitch (tense laryngeal muscles, medium s.g.p.), remaining high through the main part of the breath group (less tense laryngeal muscles, high s.g.p.), then showing a marked drop in pitch towards the end of the breath group (relaxed laryngeal muscles, low s.g.p.).

muscles remain at a constant tension in American English. However, even before publication of Lieberman's claims, there had been much evidence to support the 'larynx' view. Such evidence took various forms, e.g.-

- a) The observation that the larynx moved up and down in the neck during  $F_0$  changes suggested that it was somehow involved in these changes;
- b) A common clinical observation was that paralysis of some laryngeal muscles often led to defects in pitch control, while respiratory paralysis had no such effect;
- c) Direct electromyographic recordings of the activity of the laryngeal muscles during steady-state phonation or during singing showed the laryngeal muscles to be very active during  $F_0$  variations - it was assumed that the same would hold for speech conditions;
- d) In experiments where subjects produced a steady-state vowel at a constant pitch while receiving slight pushes on the chest at unexpected moments, the pushes produced brief involuntary increases in s.g.p. which produced brief  $F_0$  increases - the s.g.p. effect on  $F_0$  could then be calibrated. After using these calibrations to factor out that (small) part of the  $F_0$  contour of samples of connected speech which was probably due to s.g.p. variations, it was concluded that most  $F_0$  changes in speech were due to the action of the laryngeal muscles.

After publication of Lieberman's claims, more specific work was done to examine the relative contribution of s.g.p. and the laryngeal musculature to  $F_0$  changes. The general conclusion was that, while an increase in s.g.p. indeed accompanied a rise in  $F_0$  (which often coincided with a stressed syllable), this increase was far too small to be a significant cause of the  $F_0$  rise, and that therefore the laryngeal muscles were the main cause of  $F_0$  variations. Ohman and Lindqvist (1966, cited in Ohala 1977) concluded that the change in trans-glottal pressure drop that was due to stress was always much smaller than that due to stop consonants; while the  $F_0$  changes in stressed syllables did not correlate well with the stress-induced changes in trans-glottal pressure drop either in phase or in amplitude. Vanderslice (1967) found that the recording both of vertical movements of the larynx and of s.g.p. during connected speech showed the laryngeal movements to be in better synchronisation with the  $F_0$  changes than was s.g.p. Ohala (1977) observes that not only is it improbable that

the s.g.p. variations could cause much of the observed Fo changes, but that it is also probable that, to some extent, the s.g.p. fluctuations were themselves caused by the laryngeal muscles' adjusting of the vocal folds, since an increase in the Fo of the vocal fold vibration would reduce the mean glottal area, which would reduce the airflow and thus lead to increased s.g.p. Thus the increase in s.g.p. on a syllable with higher Fo, far from being the sole cause of the Fo rise, is reduced to a negligible contributory factor or even to nothing but a subsequent effect of the Fo rise. The evidence is weighted firmly against Lieberman's claims, and the consensus now is that 'larynx', not 'lungs', is responsible for Fo variations in speech.

#### 1.2.3.2. Perception aspect

While the production aspect of the view of stress as significant pitch variation led to a controversy, there has been no such disagreement in the case of the perception aspect. For some time, acoustic phonetic experiments have been conducted with the aim of determining the acoustic correlates of stress in English. All have concluded that Fo change is the most important and consistent cue. Fry (1958) used a 'Pattern playback' machine to synthesise speech at a constant 120 Hz, with a 'Voback' vocoder to vary the Fo, and used noun/verb pairs distinguished solely by stress (e.g. *object*/*object*) to elicit stress judgements from listeners. He found that higher Fo produced an impression of greater stress, whichever syllable it was on, while the magnitude of this pitch change had little effect on stress judgements beyond the fact that it had to be perceptible. He termed this the 'all-or-none' effect. What was far more important than the amount of pitch change was its direction, in that a very large step-up was in more danger than an equally large step-down of being perceived as unnatural and therefore of confusing stress judgements. Fry also found that duration and intensity had an influence on stress judgements, the former more than the latter. Thus the view of stress as intensity was shown to have no basis in fact. The cue of Fo change could outweigh both the other cues in nearly all cases; he termed this cue 'sentence intonation', as it involved a glide in pitch on a single syllable.

Similar results were discovered by Lieberman (1960), who also used noun/verb pairs, but whose experiment took the form of measurement of the relevant acoustic parameters of syllables unanimously judged stressed; he did not attempt synthesis of these parameters. Lieberman found that higher  $F_0$  was the most important single acoustic correlate of stressed syllables. The next most important cue was the integral of the amplitude with respect to time (performing almost as well as the  $F_0$  cue), while duration, though certainly a cue to stress, was less reliable - thus Lieberman differs from Fry in his assessment of the importance of duration as a cue to stress. These results also cast doubt on the view of stress as intensity, since  $F_0$  here is a marginally better cue. He also discovered 'trading effects' between  $F_0$  and amplitude cues, such that changes in  $F_0$  coherent with perceived stress could compensate for non-coherent changes in amplitude, and vice-versa. Despite such success in isolating reliable unidimensional cues to stress, however, Lieberman saw that stress judgements seemed to be made using several cues, and accordingly devised a program for mechanically recognising stressed syllables in noun/verb pairs which he claimed to be in agreement with perceptual stress judgements 99.2% of the time. This program took account not only of  $F_0$  change, but also of peak amplitude, the integral of the amplitude with respect to time, duration, and the pitch to amplitude relation (i.e. the 'trading effects', if any). In its main conclusions, Lieberman's work agrees with Fry's.

There are two observations to make on this approach to the definition of stress. First, the  $F_0$  cue is different in kind from the duration and intensity cues, being binary ('all-or-none') rather than scalar, as in the latter two cues. Fry found that duration and intensity, in the absence of  $F_0$  variation, had a marked effect on stress judgements, and so they must be considered as valid cues to stress; particularly since Lieberman found amplitude to be practically as important as  $F_0$ . However, neither Fry nor Lieberman attempted any explanation of the all-or-none effect, and no satisfactory explanation has been formulated since. Thus there remains a situation in which one acoustic cue to stress is of a radically different kind from the other cues, with no apparent basis for this distinction. The second observation is that, while Fry is commendably cautious about the implications of his results, emphasising the complicated interaction of a

number of cues in actual language use, and the limitations of the radical simplifications necessary in experimental work, Lieberman, on the other hand, is bolder in his claims and more confident about the ability of his program to locate stressed syllables in noun/verb pairs. A later experiment by Lieberman, however, casts doubt on the possibility of absolute certainty when using only inherent acoustic cues to locate stress.

Lieberman (1965) conducted an experiment to test whether the linguist using Trager-Smith notation makes use of any kind of 'objective' procedure in which he or she concentrates on the physically present acoustic signal. He prepared stimuli using a POVO synthesiser to give four types of stimulus from the original recording:

- a) Fo and amplitude information preserved and superimposed on a synthesised [a] vowel;
- b) Only Fo information preserved and superimposed on the synthesised [a];
- c) Only smoothed Fo information preserved and superimposed;
- d) Only amplitude information preserved and superimposed.

Two linguists then transcribed the resultant stimuli using the standard Trager-Smith notation, and the results were compared with the mechanically-obtained Fo contours. The overall finding was that the linguists often used their 'subjective' judgement, and used the notation appropriate to the structure of the sentence (which could be in one of eight 'modes'), even where there was a different objective Fo pattern. The same tendency was found in the stress transcription, where the transcribing linguist seemed to be inferring the presence of secondary and tertiary stresses from his knowledge of the grammatical attributes of the words, rather than hearing these stress levels in the physical attributes of the speech signal. Lieberman suggests that perhaps only two degrees of stress may have acoustic correlates independent of vowel quality. This stance is more conservative than his earlier one, and may be a closer approximation to the truth, since it does not make the impossible demand that all stress-locating algorithms based purely on acoustic cues should be completely infallible. However, the fact remains that some degree of stress at least may be located using acoustic cues, and thus the extreme

'abstract' view is probably untenable.

#### 1.2.4 Stress as accent

##### 1.2.4.1. Accent equals stress

Following on from the findings of acoustic studies that pointed to Fo as the most important cue to stress in English, Bolinger formulated his theory of 'pitch accent' which defines stress solely in terms of pitch variation. Whereas the mainstream 'British' and 'American' schools of thought on suprasegmental phenomena had separated stress phenomena (seen as intensity-based) from intonational phenomena (seen as pitch-based), Bolinger, on the other hand, conflated stress with intonation in his notion of 'accent'. In common with all other scholars at that time, he was working under the assumption that there existed inherent cues to stress contained within each stressed syllable and within no unstressed syllable. His particular innovation was the identification of 'accent' as a candidate for this inherent cue.

Bolinger's primary distinction was between 'stress' and 'accent'. His 'stress' is the potential for accent, and is a lexical abstraction that is part of the unchanging properties of a word. His 'accent' is the realisation of this potential for prominence on a particular syllable in a particular utterance. Thus Bolinger's 'accent' is what most previous scholars had been investigating under the name of 'stress'. 'Accent' takes the form of pitch obtrusion or pitch change, which can be of three kinds; A, B, or C.

a) Accent A shows a marked pitch drop during or immediately after the accented syllable.

b) Accent B, on the other hand, shows a marked pitch rise, either during/immediately after the accented syllable, or from the preceding syllable to the accented one.

c) Accent C shows a marked pitch drop from the preceding syllable to the accented syllable. (It is not clear why accents A and C are held to separate accents, while the analogous (rising) accents are subsumed into the one accent B.)

Bolinger characterises the pitch changes of the accents as being 'all-or-none', while pitch changes due to intonation are 'gradient'. While intonational pitch changes may involve large pitch differences, these occur on a scale (e.g. a scale of 'degree of finality'), and at no point on this scale does the accent change its type. This distinction is similar to Crystal's distinction between the 'simple' and 'complex' pitch-range systems (Crystal 1969:141).

Although Bolinger seems to define the three accents purely in terms of pitch patterns, in practice other factors may be considered, e.g. 'contextual redundancy', phonetic reduction of a vowel, duration, and listener's knowledge of the language. He sees duration as being subordinate to and co-variable with accent, in that a certain amount of duration is required for pitch obtrusion to be executed at all (Bolinger 1958:138). In the case of potentially ambiguous pitch patterns, where there are two possible 'corners', either of which may be the significant one, Bolinger writes:

'The fact that the linguistic signal is mixed does not refute it, but simply means that under the circumstances the hearer looks to another cue' (op. cit.:130-131).

This cue is 'contextual redundancy', the accent being deduced to lie on the stressed syllable of the word which contains new information. He adds that in place of or along with the contextual cue, the unaccented syllable may be 'phonetically degraded', and in this case the accent is thrown onto the other word. He writes:

'In this sense, the allegro and lento forms of words perform an accentual function. Normally an accented syllable must not be phonetically degraded' (loc. cit.).

This sounds very much as though duration were challenging the supremacy of pitch, while traditional stress is creeping in by the back door and pragmatics is arriving on the scene quite unheralded. But since Bolinger does not explicitly justify this unlooked-for deviation from his pitch-based definition of the accents, it is not a straightforward matter to work out what he is actually doing with the accents. Ladd, however, argues (1980: chapter 2) that Bolinger does indeed define prominence independently of pitch movement, in that he makes implicit use of other

criteria besides pitch movement. If this is true (and the evidence suggests that it may well be), then the distinction between the 'traditional' and 'pitch-accent' analyses of suprasegmentals breaks down, since all pitch movement will correspond only to intonation, and stress/prominence will be defined on a separate basis. In such a situation, Bolinger's 'pitch accents' become entities of the same order as the British school 'nuclear tones' (e.g. fall, fall-rise, low rise); that is to say, they become pitch movements that are merely associated with an already prominent/stressed syllable, rather than pitch movements that define this prominence in the first place.

#### 1.2.4.2. Accent equals stress-plus-more

A possible source of confusion is provided by Crystal's (1969) use of the terms 'stress' and 'accent', where the distinction is not at all of the same kind as Bolinger's. For Crystal, stress and intonation are separate, and stress and accent both make use of a 'bundle' of the phonetic features of pitch, intensity, and duration; the difference lies in 'which of the attributes of sound is the perceptually most dominant feature of utterance' (Crystal 1969:120). For stress, the dominant perceptual component is loudness, an increase in which can allow a syllable to achieve a certain (minor) degree of prominence in the sentence. For accent, the dominant perceptual component is pitch, in the form of a 'linguistically marked' pitch movement which lends the syllable a major degree of prominence suited to its greater 'emphasis' or information value; concomitant features of accent are the presence of a degree of stress and any other prosodic or segmental factors, such as relative sonority of sounds or durational variation. Thus, for Crystal, 'accent' refers to the prominent syllables of words with greatest information value and includes stress, while 'stress' has nothing to do with lexical abstractions (unlike Bolinger's view) but instead refers to the prominent syllables of words with any degree of information value; words with no information value at all are unstressed.

The difference in usage here may perhaps reflect the different emphases of the two traditions; the British school (to which Crystal belongs) has tended to concentrate on developing a system to represent the prosodic

features of an utterance that will prove useful in the teaching of English to foreigners. To this end, much work has been done on developing a workable intonation transcription system that is reliable in most contexts, easy to master, quick to write, and not too encumbered with detail (see O'Connor and Arnold 1961 for a classic example of this approach). The American school (with which Bolinger may be associated due to its probable influence on his early years as a linguist) has tended to concentrate, on the other hand, on more sophisticated theoretical models of prosodic phenomena. To this end, much work has been done on the status of intonation with regard to segmental phenomena (e.g. whether there is such a thing as a pitch-level phoneme) and, since ease of use was not the main concern, articles in this tradition fairly bristle with superscript numbers and juncture and boundary symbols, as well as four different kinds of stress mark.

Crystal's use of the terms 'stress' and 'accent' is paralleled by others within the British tradition, who equate 'accent' with 'stress-plus-pitch-prominence', and 'stress' merely with increased loudness; e.g. O'Connor and Arnold's statement that 'Pitch as well as stress is involved in the recognition of accent' (O'Connor and Arnold, 2nd. edn. 1973:31). However, Bolinger's distinction remains a useful one, and if it had become more widely used, it would have obviated the need for such phrases as 'stressable syllable' or 'lexically stressed syllable', which are not as theoretically well-defined.

### **1.2.5 Stress as a component of rhythm**

#### **1.2.5.1. Isochrony**

A phenomenon closely related to stress that has been studied for centuries is rhythm or, to be precise, isochrony. Isochrony is the condition whereby a given type of unit (for English, a stressed syllable) occurs at equal intervals of time however many other units of a different type (unstressed syllables) lie in between. This means that, for English, interstress intervals containing few syllables may show marked lengthening of those syllables, while interstress intervals containing very many syllables may show marked shortening and reduction of those

syllables, if isochrony could be shown to exist.

It is generally accepted that there is no perfect isochrony in English, not even in the most rhythmic verse. The range of possible durations of interstress intervals (or 'feet') is far too large for absolute, objective isochrony to be a tenable view. In actual utterances, interstress intervals can vary in length by as much as 6:1 (Buxton 1983:112). The most that can be said on the objective level is that there may be a tendency towards isochrony, and therefore a valid basis for the characterisation of English as a 'stress-timed' language (see Pike 1945 for the first use of this term); while, on the subjective level, hearers may subconsciously normalise quite large discrepancies in duration to arrive at an impression of isochrony only loosely based on fact. This latter view is the one currently gaining favour, as will be seen from a brief summary of experimental work on isochrony.

#### 1.2.5.2. Experimental findings

A lucid and exhaustive review of past work on the subject may be found in Lehiste (1977); only the main lines of argument will be presented here.

Isochrony in production has been found to be unreliable at best; there is certainly no absolute isochrony, and even a tendency towards isochrony may at times be so faint as to be almost undetectable. Some workers (e.g. Shen and Peterson 1962) have rejected altogether the validity of isochrony in production for English; while others give only a qualified assent, such as Classe, who states that isochrony in English exists only in very favourable circumstances, viz.:

'the groups [i.e. feet] concerned must not contain very different numbers of syllables; the phonetic structure of the component syllables must not differ too widely; the grammatical connexion between the groups and the grammatical structure of these groups must be similar' (Classe 1939:85).

The problem with isochrony in production is the difficulty of measurement; the techniques employed until now have been somewhat crude statistically speaking, and thus there is no way of telling what counts as a significant tendency and what is a chance artefact. More refined statistical techniques are only beginning to be employed, and are able to

show up phenomena that before would have gone unnoticed; hence Buxton's justified complaint that descriptions of temporal structure in speech 'have been limited to approximate claims about a "tendency to isochrony"' (Buxton 1983:119). In section 2 of the present study, statistical techniques are employed that make clear the underlying patterns and tendencies that cannot be seen on the surface. While these techniques are by no means advanced from a statistician's point of view, they are nonetheless an improvement on most of what has gone before.

Isochrony in perception has been found to be much easier to demonstrate than isochrony in production. First, Lehiste found (op.cit.) that many of the durational differences between feet found in the production of some 'four-measure sentences' were actually below the threshold of perception and therefore quite irrelevant for linguistic purposes. Since the threshold of perception of differences in foot lengths for speech stimuli was found by her to be markedly higher than for non-speech stimuli (Lehiste 1975b), the odds are weighted in favour of perceptual isochrony even before the utterance is produced. If the utterance, once produced, is not exaggeratedly rhythmic, some of the foot length differences will be above this perceptual threshold for speech stimuli and thus potentially perceivable as non-isochronous. However, there is a further safeguard built in; the listener quite often subjectively imposes a rhythmic structure on such feet, when there is a momentum of rhythm sufficient to override the perceivable duration differences. Thus the listener perceives the rhythm as isochronous when in fact it may be very far from being so.

This is the phenomenon known to psychologists as 'subjective rhythm', which has been well documented (see G.D. Allen 1975 for a review of work in this field). There are two aspects of subjective rhythm:

- a) Imposing a rhythmic structure on a sequence of identical pulses, so that some pulses sound subjectively more prominent;
- b) Under- and overestimating the duration of long and short time intervals respectively, to 'normalise' subjectively towards isochrony.

G.D. Allen (op. cit.) observes that listeners tend to adjust their perception of the duration of intervals towards an average duration. Taken together with the tendency to impose a subjective rhythm on any sequence of intervals, whether equal or not, this tendency to even out the durations contributes to the perception of regular rhythm in languages, such as English, which have a 'stress accent' rather than a 'pitch accent' - i.e. where pitch change is not the only indication of prominence.

Even after subjective rhythm has been imposed, however, there will always be at least a few isolated feet which upset the pattern radically and break the minimum of rhythmic momentum necessary for subjective rhythm to be imposed successfully. In such a case, it could be that there remains an overall impression of isochrony in the utterance as a whole, due less to any isochrony actually found than to the movement in the direction of isochrony that may lie below the surface. Since the factors constraining the duration of segments are fairly well known, for many segments (see Lehiste 1970 for a summary), it is possible to identify modifications in these constraints in a particular direction, e.g. that of isochrony. Such adjustments are actually made in production by speakers, who partially modify the duration of some speech sounds and their sequences in the direction towards isochrony, leaving the listener's subjective sense of rhythm to take its cue from these modifications and create an impression of isochrony (Rapp 1971, Lehiste 1975a). The situation is summed up neatly by Barnwell, who writes:

'The implication is that speech attempts to be very rhythmic, but fails because of durational constraints. Hence it may be that what is heard as rhythmic may be really the interpretation of changes in duration in the direction of true rhythm. Hence, just as the perception of pitch is not necessarily directly related to  $F_0$ , so the perception of rhythm may not be directly related to true time intervals in speech.' (Barnwell 1971:88)

Thus isochrony may be seen as a real component of English suprasegmental structure, but only with the qualification that the regularity resides mainly in the mind of the hearer (and possibly of the speaker also). The objective correlate of isochrony is never perfectly realised, and can only be said to be present in a virtual form as the target towards which move certain articulatory durational adjustments.

This insight of an undercurrent of rhythm in English forms the core of the 'metrical' theory of stress, first formulated in Liberman 1975 and developed in many subsequent articles by various scholars, notably in Liberman and Prince 1977. Metrical phonology attempts to state the suprasegmental structure of English in terms of a hierarchy of strong/weak relations, defined on units of at least one syllable. The binary-branching nature of this structure reflects Liberman's claim that the fundamental distinction in metrical structure is a binary one between 'strong' and 'weak', where these labels refer to a complex phonetic and phonological entity that has only indirect repercussions on the physical level. The main feature of this structure is that it is relational; that is, units are not inherently and unchangeably either strong or weak, but take their labels from their relationship with a comparable unit. Liberman separates stress from intonation, thus agreeing with both the British and the American traditional schools of thought. Where he parts company with traditional views is that, while he sees intonation as based on pitch change, his view of stress is based on rhythm. Previously it had been held that rhythm was based on stress; but Liberman turns the relationship on its head by saying, effectively, that rhythm is stress. The following section (1.2.6.) gives a very brief outline of the main features of metrical phonology, while chapter 3 contains both an exposition and a refinement of it based on subsequent work.

### **1.2.6 Stress as rhythm: the "metrical" theory of stress**

#### **1.2.6.1. Metrical theory and the cycle**

Liberman discusses Chomsky and Halle's 'Nuclear Stress Rule' which applies cyclically, reassigning [1 stress] to the stressed syllable of each word and subordinating all [1 stress] vowels to the left within the domain of application of the rule (Chomsky and Halle 1968). He then shows how the metrical theory can account for the same data, but in a much simpler way, in that the latter theory has no need for the cyclic principle, for the principle of stress subordination, nor for the variable in the formulation of the rule, and also needs no ordering of the constituent rules since it applies at all levels simultaneously. This is seen in a comparison of the two types of rule:

### Chomsky and Halle Nuclear Stress Rule

- a) word stress rules: assumed
- b) Nuclear Stress Rule: 
$$\left[ \begin{array}{c} V \\ 1 \text{ stress} \end{array} \right] \text{---> } 1 \text{ stress /\_Q}$$

(applies only to phrasal constituents)  
condition: Q contains no 1 stress
- c) Principle of cyclic application of NSR
- d) Principle of stress subordination: when primary stress is placed in a certain position, then all other stresses in the string under consideration at that point are automatically weakened by one degree.

### Liberman Nuclear Stress Rule

- a) word stress rules: assumed
- b) Nuclear Stress Rule (applies only to phrasal nodes):

In every phrasal constituent, put the strong position on the right.

Formalization:  $N \text{ ---> } S/[M\_]$  where N and M are nodes.

A general convention of metrical phonology ensures that the sister node is automatically labelled 'weak' after the assignment of 'strong' to one node.

(adapted from Liberman 1975:130-1)

The structure to which the latter rule applies is based on the syntactically given constituent structure. The difference between the two types of rule is that, whereas the Chomsky and Halle rule assigns stress vowel-by-vowel and cycle-by-cycle, the Liberman rule assigns a pattern simultaneously to a whole structure.

More telling still, the phonological cycle in the conventional theory is an optional additional assumption, not logically related to the nature of the particular rule involved. Thus it would be as reasonable to postulate it for segmental rules as it is for prosodic rules. As it is, prosodic phenomena are the only ones for which the cycle can be plausibly argued. On the other hand, the metrical theory lacks any apparatus for cyclic rule application, since it accounts for the 'cyclic' properties of prosodic phenomena on the basis that stress is a hierarchically defined relation. Thus the metrical theory accounts for these properties on the basis of the intrinsic nature of the phenomenon itself. This characteristic is not shared by segmental rules, since rather than

defining hierarchical relations among phonological entities, they define intrinsic properties of the phonological entities themselves. The metrical theory predicts that prosodic phenomena will, by their very nature, have 'cyclic' properties (i.e. the preservation of relative stress under embedding) but makes no predictions for segmental rules (see Liberman 1975:129-138 for a fuller discussion of the 'cyclic' nature of the metrical theory, and section 3.4. of the present study for full details of metrical or 'prosodic' structure).

#### 1.2.6.2. The metrical grid

Once the metrical structure has been assigned, on the basis of constituent structure and rules for nuclear stress and compound stress, then the 'metrical grid' is characterised. This is an abstract entity reflecting the metrical structure but independent of it, and thus only loosely tied to syntactic surface structure. It reflects the temporal patterning and the relative prominence of the syllables in the utterance - in short, the rhythm. A syllable marked as prominent in the metrical grid is highly likely to begin a rhythmic foot and to be stressed. The intonation of the utterance, however, is more closely linked to the metrical structure, and so is not necessarily synchronised with the rhythmic patterning (see section 3.5. for a full discussion of the metrical grid).

The fact that it is the rhythmically-prominent syllables as defined by the metrical grid which are characterised as stressed, gives a clue as to the metrical theory's attitude to stress. This is seen as defined not on a syllable-by-syllable basis in terms of inherent acoustic properties, but rather in terms of the rhythmic structure of the whole utterance, apart from which it would make no sense to call any syllable 'stressed' - a viewpoint quite in accord with the relational nature of this theory, which defines a node as 'strong' only in relation to a 'weak' sister node, never on its own merits. Ladd (1980:26) sums up this attitude concisely by stating that measurements of fundamental frequency, duration, etc.,

'...should not be interpreted as demonstrating the acoustic correlates of a particular stress peak, but rather as indicating the sorts of acoustic features which induce us to hear an overall rhythmic organization which includes the particular stress peak.'

In other words, stress is rhythm, and without rhythm one cannot talk about stress, except possibly in terms of a potentiality on the stressable syllable of each word. This is the insight gained from the experiments with Welsh stress described in the second chapter of this study, and it informs the theoretical treatment of stress in the third chapter. The general plan of this study is described in section 1.4. below, but first, in section 1.3., some observations made in the past concerning Welsh stress will help to provide the background to the problem under investigation.

### 1.3 Informal observations of Welsh stress

The following points are of an informal and pre-theoretical nature, but it is hoped that they will help to set the scene for those with no knowledge of Welsh.

#### **1.3.1 Welsh stress rule**

Stress in Welsh is fixed on the penultimate syllable, for polysyllables. In the few cases where the stress does not follow this rule, nearly every example has stress on the final syllable. Only a very small minority of words have stress on a syllable before the penult; such words are mainly recent loans from English, and thus can perhaps be discounted as not being part of the native word-stock.

#### **1.3.2 Reaction of English listeners**

While non-Welsh-speaking English people would probably experience no difficulty in identifying stressed syllables in their own language, the situation is different when they are presented with Welsh. Rees (1977) remarks in a footnote that he has observed, in an informal experiment, that English people consistently judge the stress in Welsh to be on the final syllable rather than the penult, and thus are consistently wrong. My own experience bears out this observation: as a phonetically-trained English speaker having learned a little Welsh, I carried out an undergraduate study of intonation in Welsh in the framework used by O'Connor and Arnold (1961) for English. I quickly found that the syllables I had learned were

supposed to be stressed did not in fact sound stressed, while the supposedly unstressed final syllables sounded clearly stressed in very many cases; identifying the stressed syllables was a most unexpectedly difficult task. Similarly, in a study of the Cwm Tawe (Swansea Valley) dialect of Welsh, Watkins writes that:

'I have also been told that to the foreign ear, unaccustomed to dialect Welsh, it would seem that there is a stronger stress on the final than the penultimate syllable' (Watkins 1953:8).

It was this kind of observation that sparked off the present study; while it is not claimed that stress in Welsh is now completely accounted for, nevertheless much light has been thrown on the question from several angles.

### 1.3.3 Robustness of the ultima

In the available literature on the acoustic or auditory nature of Welsh stress (the references being both minimal and largely imprecise), there seems to be some degree of agreement that the final syllable of polysyllables is more prominent acoustically and auditorily, and that it has higher pitch than the penult (Jones, D.B. 1949, Watkins 1953, 1972a, Thomas, C.H. 1967, Rees 1977, Griffen 1979, Thomas, A. 1979). Other rather subjective comments are also made; for example, Watkins (1972a) states that the Welsh accent is not as 'strong' or 'loud' as that of English, and that the stress is more evenly distributed, 'particularly as between penult and final'. D.B. Jones (1949) states that the stress difference between the penult and the ultima is small, and is much less than the difference between the penult and the preceding syllable(s); this, he writes, can be heard by the ear, and is also seen in phonetic changes due to the accent. Griffen (1979) sets up three phonological levels of stress in the word; in polysyllables, the ultima has level 2, the penult level 1 (primary stress), the antepenult (if any) level 3, and the preantepenult (if any) level 2. Thus this analysis also postulates only a small difference of stress between the penult and ultima. There seems to be only one past study of Welsh stress in acoustic terms; Watkins 1953. This rudimentary investigation confirms experimentally that the pitch of the ultima is higher than that of the penult, in some words at least, and

also finds that the 'stress' (by which is meant greatest intensity and longest duration) may fall either on the penult or the ultima. Watkins concludes:

'...it would seem that the stress and pitch in modern Welsh especially in the Southern dialects are not on different syllables. The pitch has a tendency to follow the stress, but one thing is sure, and that is there is no great stress in modern Welsh. There is very little difference between the penultimate and final syllables, pitch and stress accents resting upon one or the other depending upon the position of a word in a sentence and the importance which is attached to the word. It is true to say, however, that the final syllable is longer (comparatively) than the penult and that this is a contributing factor in lending it greater clarity in dialect words.' (Watkins 1953:9)

Thus the ultima is very often more prominent in phonetic terms than the stressed penult. This is also true from a phonological point of view: while the central vowel schwa is never found in the ultima, it is often found in non-final syllables, even in stressed penults. Similarly, the phenomenon of 'vowel mutation', whereby some non-final high vowels lower to schwa in certain contexts, never causes the vowel of the ultima to reduce to schwa (see Williams (in press) for an attempted explanation of this phenomenon). Thus the vowel of the ultima is more 'prominent' or 'robust' in both phonetic and phonological terms; as A. Thomas puts it,

'...the ultimate syllable is often (perceptually) the stronger of the two [i.e. penult and ultima], and always structurally the stronger' (Thomas, A. 1979:170).

This is aided by the fact that the penult is very 'weak' in some sense, as will be seen below.

#### 1.3.4 Weakness of the penult

The stressed penult has been observed to be somewhat weak, both acoustically and auditorily (Jones, D.M. 1949, Watkins 1953). It is also weak structurally; in rapid speech the penult of some weakly-stressed disyllables can be lost altogether, while their final syllables are never lost, in cases where the vowel of the ultima is 'of greater natural sonority' than that of the penult (Jones, op. cit.) - e.g.:

eto --->      'to ('yet')                      /ɛ tɔ/ --->      /tɔ/

<u>yma</u> --->	'ma ('here')	/ə ma/ --->	/ma/
<u>yna</u> --->	'na ('there')	/ə ma/ --->	/na/
<u>dyna</u> --->	'na ('there is/are')	/də na/ --->	/na/
<u>acw</u> --->	'cw ('yonder')	/akə/ --->	/kə/

This observation is confirmed by Watkins (1953), who adds his own examples, and also gives cases where the stressed penult, though not lost, at least has its vowel reduced to schwa before a long consonant or consonant cluster. In dialect Welsh, the vowel of the penult can sometimes change its quality, shifting to a rounded vowel under the influence of a following labial stop; this does not occur in the final syllable (Watkins, op. cit.). Jones sees this weakness of the penult as a relic of the Old Welsh Accent Shift which, he writes, left a high pitch accent on the ultima while the 'stress element' shifted to the penult. He writes; 'No doubt the frequently higher pitch of the final syllable plays a part in this process' (Jones, D.M. 1949). Watkins sees the Accent Shift in different terms; he writes that, in Old Welsh, the stress was evenly distributed on all the syllables of polysyllables, with higher pitch on the ultima. The Accent Shift was not so much a shift as a gradually-developing stress on the penult, with higher pitch remaining on the ultima (Watkins, 1972a). However, the general agreement is that nowadays the final syllable is very often both phonetically (in acoustic and auditory terms) and phonologically (in structural and performance terms) in some way 'stronger' or more prominent than the stressed penult.

#### 1.4 Plan of the thesis

The second chapter of this thesis concerns a series of measurements and experiments undertaken to ascertain the acoustic nature of stress in Welsh. Both rehearsed and unrehearsed data are used, as well as perceptual experiments, and the same general results obtained for each, suggesting that the findings are indeed related to what happens in actual language use. An attempt has been made to provide a rather more rigorous statistical treatment than has been usual in earlier studies of stress, and it is this statistical treatment that shows up tendencies which would have been almost impossible to discover in any other fashion.

The third chapter of the thesis concerns the development of a theoretical model of stress that will be appropriate to the experimental results. The theory is based on Liberman's 'metrical' theory of stress, but incorporates refinements suggested by Selkirk, and also modifies it further to allow the theory to take account of semantics.

The initial exposition of this modified theory is carried out in terms of English, in order to demonstrate the workings of the theory in terms of a language known by all readers, and also to show that the theory is at least descriptively adequate for English. After the initial exposition, some puzzles in English intonation are tackled. These puzzles, set in an O'Connor and Arnold type of framework, are shown to be more easily explained using this modified 'metrical' approach. The theory is then applied to modern Welsh, which turns out to be a little simpler than English owing to the fixed nature of stress in Welsh. Thus the theory may be more immediately suited to Welsh than to English, particularly since experimental results support the application of this theory to Welsh (in chapter 2), while no such concrete evidence can be cited for applying the theory to English. Be that as it may, it is suggested that the theory provides a simpler and neater account of the facts, for both Welsh and English.

The fourth chapter of this study discusses the implications of the experimental findings for the theory of stress in general. It is discovered that the ramifications spread far beyond the Welsh language, and may influence our assumptions of what constitutes 'stress' in a language. In this respect, the present study is not relevant merely to Welsh (though it throws much light on old puzzles in that language), but has a far wider application that concerns the theory of stress in any language of the world.

## 2. EXPERIMENTAL WORK

### 2.1 Words in citation form

#### 2.1.1 Aim

The aim of this first set of measurements was to discover what, if any, were the acoustic correlates of stress in Welsh. This was envisaged as being of the same order as stress in other languages that have been investigated in the past, such as English, French or Polish, where increases in certain intrinsic acoustic properties can be shown to coincide reliably with a stressed syllable.

#### 2.1.2 Method

A recording was made of a male native (South) Welsh speaker reading from a prepared list. The subject was placed in an acoustically-treated recording studio, and read into an AKG 451 microphone leading to an Ampex AG440 open-reel tape recorder.

The list contained both isolated Welsh polysyllables, and a few Welsh polysyllables embedded in a frame sentence that placed the word in question in nuclear position after a momentary pause. The frame sentence in question was: 'Dydi hi ddim yn ddigon i YSGRIFENNU "..." (/dɛdɪ hɪ ɔ̃ɪm ən ɔ̃ɪgən ɪ sgrɪvənɪ .../, 'It's not enough to WRITE "..."). Thus the words used were in citation form and carried the intonational nucleus, in most cases. Twenty words were used in all; naturally, this number was not enough to give a detailed account of stress in Welsh, but was sufficient to make plain the main trends.

Two non-Welsh-speaking English speakers and a Welsh speaker (not the recorded subject) then listened separately to the tape and made stress

judgements for each word. In the few cases where one or both of the English speakers found it impossible to make a judgement, the word in question was not counted in the subsequent analysis. Sound-spectrograms were made of each word, using a Kay Electronics 7029A 'Sonagraph', and making both wide-band and narrow-band spectrograms. Measurements were made by hand from these spectrograms, and various calculations made.

The division of the speech signal into segments was made on the basis of the wide-band spectrogram; for the purposes of measurement, it was decided that a phone would be said to have begun only when it showed a clear pattern of voicing, formant structure, or silence, as appropriate. There was necessarily some idealisation at this stage, but the precision of measurement was not of the highest degree, and so it was felt that such considerations would make little difference to the results. The mean amplitude was estimated by eye from the narrow-band amplitude display, a horizontal line being drawn as an attempted best mean of the variations in amplitude (which were surprisingly small in most cases, since sudden sharp peaks were rare). The area bounded by this line of mean amplitude, plus the baseline and the start and end of the segment as defined on the basis of the wide-band spectrogram, was designated the 'overall' or 'envelope' amplitude, informally speaking; it was hoped that this was a fair approximation to Lieberman's 'integral of the amplitude with respect to time' (Lieberman 1960).

In considering the stress of the words, only the vowel stretches were measured, firstly for reasons of ease of measurement, and secondly since Fry (1958) had found that, in English words, the stress was carried almost entirely by the vowel stretches of the signal. He had accordingly varied only the vowel portions in his subsequent experiments with synthesised speech, and this produced no results which might have called the procedure into question.

When measuring changes in fundamental frequency ( $F_0$ ) within a single syllable, such a change was only acknowledged if the  $F_0$  change (in either direction) was greater than or equal to 15 Hz. This was mainly due to the poor precision of the measuring technique (use of a roamer to estimate  $F_0$  from the frequencies of the higher harmonics); however, this restriction

has the advantage that a change in  $F_0$  of 15 Hz or more would most probably be perceptible, while a smaller change might well go unnoticed in a rapid conversational situation, and thus would be irrelevant for linguistic purposes.

### 2.1.3 Results

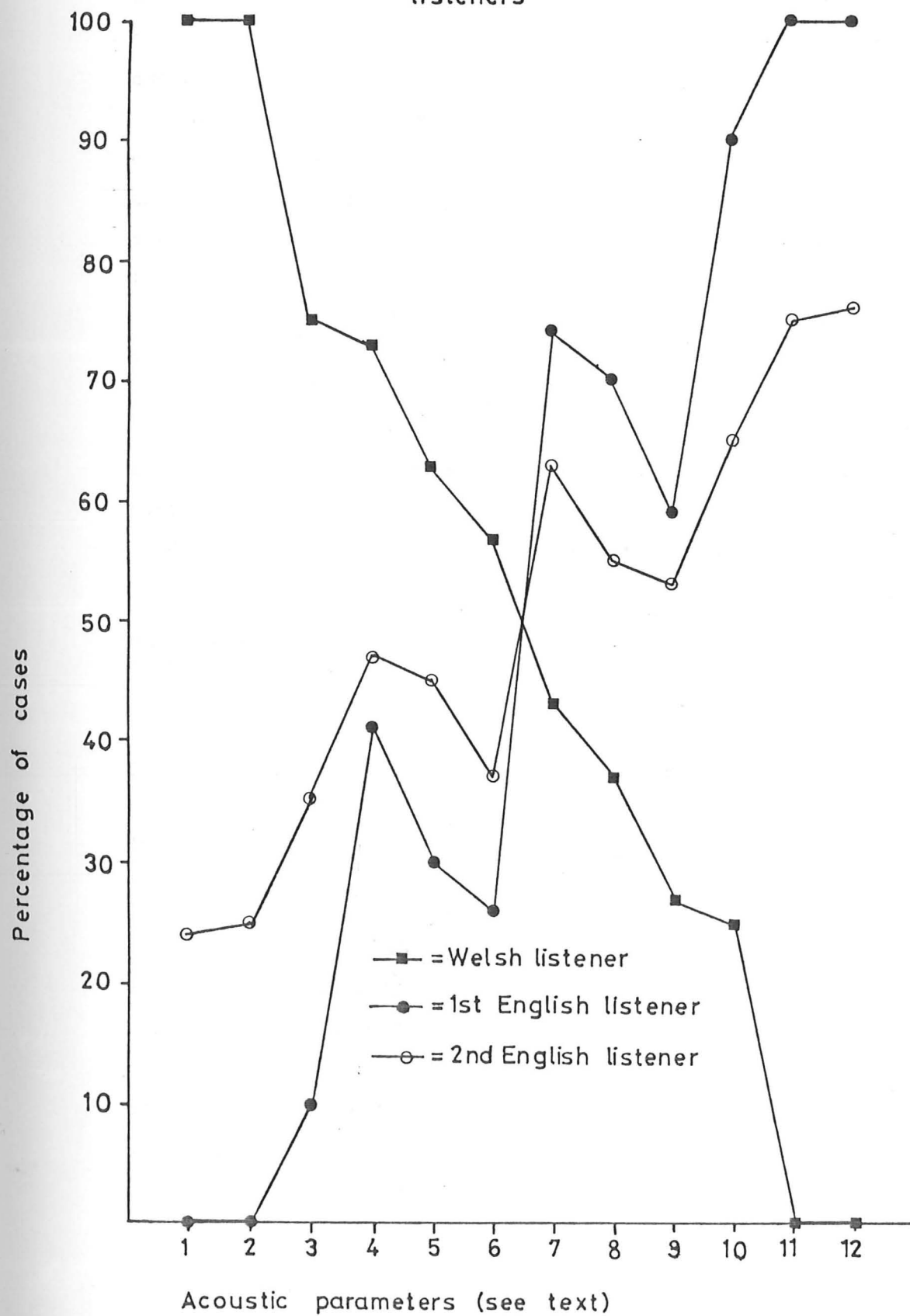
The results of these measurements are presented graphically in Fig. 1, which shows the properties of syllables judged stressed by each listener. The twelve 'acoustic parameters' are as follows:

- 1: shorter duration of vowel
- 2: lower estimated amplitude integral of vowel
- 3:  $F_0$  change (within vowel) of less than 15 Hz
- 4: higher  $F_0$  at start of vowel
- 5: greater mean amplitude of vowel
- 6: greater peak amplitude of vowel
- 7: lower peak amplitude of vowel
- 8: lower mean amplitude of vowel
- 9: lower  $F_0$  at start of vowel
- 10:  $F_0$  change (within vowel) greater than 15 Hz
- 11: greater estimated amplitude integral of vowel
- 12: longer duration of vowel

In each case, the comparison is made between the given vowel and the other syllable(s) of the same word. It will be noted that half of the above acoustic parameters are the converse of the other half; this is because, although in nearly all cases the value for a cue is easily deducible from that of its complement, the few cases of equality between two vowels in one or other of the acoustic dimensions made it necessary to include converse cues as well. The words in question are all regularly stressed (i.e. on the penult); this restriction is introduced since an informal, similar, test suggested that stressed final syllables in Welsh show quite a different pattern of acoustic cues from stressed penultimate syllables, and thus it would be wiser to concentrate on the more common type of stress pattern.

Fig. 1

Properties of syllables judged stressed by three  
listeners



#### 2.1.4 Discussion

The most striking feature of Fig. 1 is the sharp contrast in general tendency between the two English listeners and the Welsh listener. The English listeners, following the cues for stress in English, interpreted as stressed those Welsh syllables which had one or more of the 'standard' cues: Fo change (glide or step), greater amplitude, or longer duration. On the other hand, the syllables judged stressed by a Welsh speaker had the opposite attributes. To make matters more confusing, it was found that the final, supposedly unstressed, syllable of these polysyllables almost always had longer duration, a pitch glide, and greater envelope amplitude. This was the syllable typically chosen as stressed by the two English listeners, thus supporting the informal observations described in section 1.3.2. earlier.

The implications of this result are that the commonly-accepted view of stress (as intrinsic acoustic cues) will not do for Welsh; the basis of stress in Welsh is of a different kind. However, it should be borne in mind that this was merely a limited and preliminary experiment.

### 2.2 Connected speech

#### 2.2.1 Unrehearsed speech: segmental measurements

##### 2.2.1.1. Aim

Since the preceding experiment was merely of a preliminary nature, it was decided to check its findings in the case of words in connected speech, i.e. not in citation form and not necessarily in nuclear position. The first type of connected speech to be investigated was (relatively) spontaneous speech, in order to find out whether the above findings held for Welsh outside experimental conditions.

#### 2.2.1.2. Method

Recordings were made of Welsh speakers at the 1981 National Eisteddfod at Machynlleth, using a Sony TCM 260 portable cassette tape-recorder with separate microphone. Because of the recording conditions (crowded field background noise to a) poor-quality loudspeaker b) single Welsh speaker c) tape/slide presentation), the quality of the recording was poor. However, the main acoustic features were recoverable from spectrograms. English listeners' judgements could not be tested, since a non-Welsh speaker could not have divided each utterance into words.

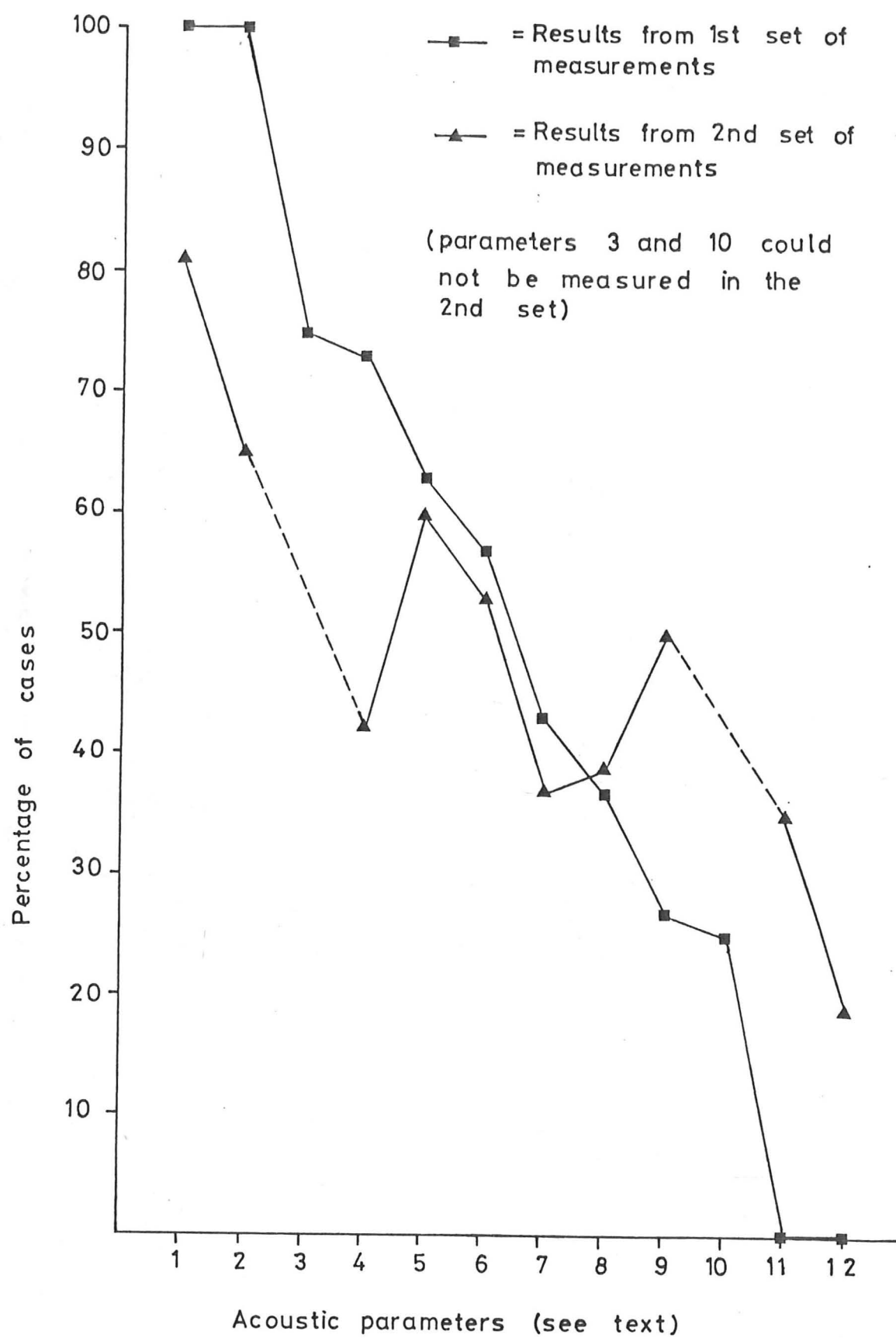
As before, sound-spectrograms were made of twenty-four of the more intelligible sentences and long phrases. The same measurements were made as before, with the exception of the  $F_0$  change within the vowel, which could not be measured in these utterances. Of these twenty-four samples, twelve were spoken by the master of ceremonies on the platform in the main pavilion (these are mainly words of thanks and scene-setting, in a slightly formulaic style but not read from a script); four of the samples were taken from a tape/slide presentation of the abbeys of Wales (though not colloquial in style, these utterances were spoken fluently); and eight of the samples were spoken by a middle-aged female native speaker who held a short conversation (this lady turned out to have strong views on the subject of the Welsh language, and thus her speech was highly likely to be without influence from English). The vocabulary and grammatical structures used in these samples were mainly of an everyday kind.

#### 2.2.1.3. Results

The results are presented in graphical form in Fig. 2, in the same format as before (minus parameters 3 and 10). The number of words measured was forty-six. Only non-prepausal and regularly-stressed words are represented, since prepausal and irregularly-stressed words (i.e. with stress on the ultima) show marked deviations from this pattern. Superimposed on these results is the line representing the properties of syllables judged stressed by the Welsh listener in the first set of measurements. Fig. 3 shows a wide-band and a narrow-band spectrogram, with amplitude display, of one of the unrehearsed utterances, as an example of the type of recording quality obtained.

Fig. 2

Properties of stressed syllables in two sets of measurements



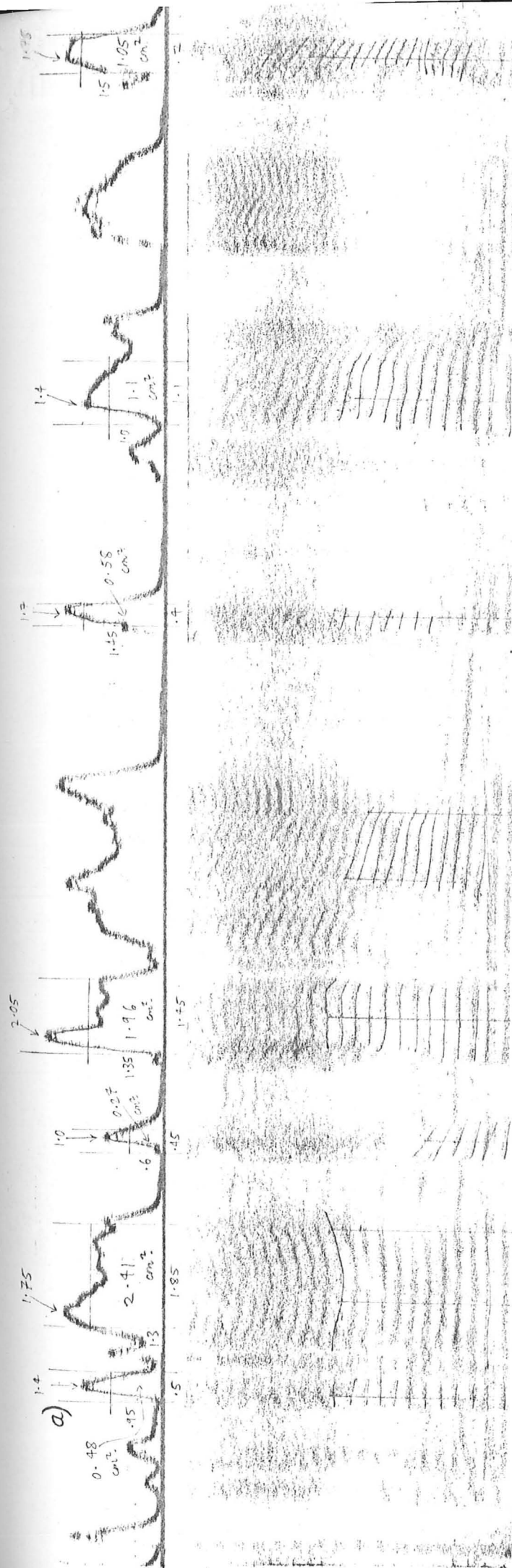


Fig. 3: Wide-band and narrow-band spectrograms, with amplitude display, of one of the unrehearsed utterances.

kəməɹə ɪ g ə d h i ɹ j ə i θ g h ə n t h a i v i g h ə r

The pattern for this second set of measurements is not as clear-cut as in the first set, but the same trend is present; the stressed vowel is the one with shorter duration and greater peak and mean amplitude. The only major difference between the two sets of measurements was that the results for the second set indicated that the envelope amplitude of the vowel (parameter 2), and a step-up in  $F_0$  (parameter 4), were most unreliable cues to stress or the lack of it. By far the most consistent cue was the shorter duration of the stressed vowel; this was seen even in the less distinct results of the second set of measurements.

#### 2.2.1.4. Discussion

The findings of the first set of measurements seem to be not just a feature of the citation forms of words, since their general trend is reproduced in the second set, using fluent connected speech. Thus the main points can be regarded as real features of the stressed penult in Welsh; i.e. shorter duration, lower amplitude integral (probably influenced by the shorter duration), and lack of any great consistency in the cues of  $F_0$ , mean amplitude and peak amplitude. This situation is most peculiar when compared with the acoustic features of stress in English, where shorter duration and lower amplitude integral (consistent cues in Welsh) would most often betoken lack of stress, while the cues of  $F_0$ , mean and peak amplitude would be very reliable cues to stress in English, unlike their status in Welsh. It indeed seems that any acoustic cue to stress in Welsh must be sought in a different area entirely.

### 2.2.2 Unrehearsed speech: foot measurements

#### 2.2.2.1. Aim

Since vowel duration (or rather, the lack of it) had been found to be such a reliable cue to stress in Welsh, it was decided to investigate timing phenomena more closely. However, since intrinsic cues had failed so convincingly (in the case of penults) to provide a definition of the realisation of stress in Welsh, the aspect of the speech signal to be examined next was relational rather than inherent. This investigation centred on temporal phenomena in a wider domain than that of the syllable - i.e. rhythmic phenomena.

#### 2.2.2.2. Method

The spectrograms of connected speech described in section 2.2.1. were used, together with the stress judgements made by the native Welsh speaker. The stretches between the onset of each stressed vowel were measured, these stretches being referred to as 'feet'. For each foot, a note was also made of the number of syllables contained in it, and the mean duration of the syllables in a particular foot was calculated by dividing the foot length by the number of syllables in it. Thus the three variables in question were: length of foot, number of syllables in foot, and length of syllable.

Two different kinds of measurement were carried out, corresponding to two different hypotheses concerning the identification of the stressed syllable in Welsh. Under the first hypothesis, the penultimate syllables judged as stressed were counted as stressed for the purposes of foot demarcation; while under the second hypothesis, their corresponding final syllables were counted as stressed. The resulting feet could be termed a 'Welsh speaker's foot' and an 'English speaker's foot' respectively, with stressed monosyllables counting as stressed under both hypotheses.

#### 2.2.2.3. Results

The resulting measurements were analysed using the SAS (Statistical Analysis System) package on the Cambridge University IBM 3081 mainframe computer. This gave the (product-moment) correlation coefficients over both prepausal and non-prepausal feet, as shown in the table below.

### Two types of foot measurements

correlation of:		'Welsh speaker's		'English speaker's
		foot'		foot'
-----	+	-----	+	-----
		coeff. significance		coeff. significance
	+	-----	+	-----
length of foot: no. of		0.90            0.0001		0.93            0.0001
syllables in foot		n = 46		n = 42
length of syllable:		-0.37            0.0116		-0.23            0.1455
no. of syllables		n = 46		n = 42
in foot				
-----	+	-----	+	-----

#### 2.2.2.4. Discussion

The results for the 'Welsh speaker's foot' showed a slight tendency for stressed syllables to occur at approximately equal intervals. Although the length of the foot was partly determined by the number of syllables in the foot, there was also a tendency for the length of each syllable to be negatively correlated with the number of syllables per foot. Thus the more syllables there were in a foot, the shorter each syllable would become, on the whole, so that the interval between stressed syllables was not as long as it would have been were the length of the foot determined solely by the number of syllables in it. In other words, there was a distinct tendency towards isochrony.

This was not the case, however, for the 'English speaker's foot'. As before, the number of syllables per foot was an important influence on foot length, but the negative correlation between syllable length and number of syllables per foot was less marked and had a much lower degree of significance (which was, statistically speaking, not significant at all and therefore worthless).

The overall result is that, if the penult is counted as stressed, a much greater tendency towards isochrony is seen than if the more acoustically prominent ultima is counted as stressed. The 'stressedness' of the penult

in Welsh seems to be due less to its inherent acoustic properties than to its function as the keystone of the rhythmic unit; this finding could well have significant implications for our view of stress in general.

### 2.2.3 Rehearsed speech

#### 2.2.3.1. Aim

Having discovered that the key to stress in Welsh lay in the rhythm of connected speech, the next step was to check this finding for a larger sample of speech. At the same time, it was decided to test both for any difference in rhythmicity there might be between two different styles of presentation of a list of sentences, and also to test for any effects there might be due to 'syllabification' - i.e. measuring not from stressed vowel onset, but from stressed syllable onset, so that initial consonants of stressed syllables would be included in the following foot.

#### 2.2.3.2. Method

A male native Welsh speaker was recorded, under the conditions described in section 2.1.2. above, reading a list of simple Welsh sentences. The recording consisted of two sets of sentences. The first set contained randomly-ordered sentences, while the second set contained most of these sentences, arranged so that similar sentences occurred in groups. Within each group, each successive sentence was longer by the addition of one or two syllables. One such group is the following:

- 1) Mae'r got ar fwrdd mewn tŷ  
/maɪr gɔt ar vɔrð meɔn tɪ:/  
The coat is on a table in a house
- 2) Mae'r got ar y bwrdd mewn tŷ  
/maɪr gɔt ar ə bɔrð meɔn tɪ:/  
The coat is on the table in a house
- 3) Mae'r cotiau ar y bwrdd mewn tŷ  
/maɪr kɔtjaɪ ar ə bɔrð meɔn tɪ:/  
The coats are on the table in a house
- 4) Mae'r cotiau ar y bwrdd yn y tŷ  
/maɪr kɔtjaɪ ar ə bɔrð ən ə tɪ:/  
The coats are on the table in the house
- 5) Mae'r cotiau ar y byrddiau yn y tŷ  
/maɪr kɔtjaɪ ar ə bɔrðjaɪ ən ə tɪ:/  
The coats are on the tables in the house

Such an ordering represented an attempt to test both for the consistency of the tendency towards isochrony, measured over a different set of sentences from before, and also to test for any change in rhythmicality due to these similar sentences being grouped together in this way. The first list contained these sentences in random order, as well as similar 'distractor' sentences, and was presented first, all in order to avoid suggesting to the subject that the sentences in the first set had any rhythmic relationship. A native Welsh speaker (not the recorded subject) then listened to the tape and made stress judgements.

The recordings were then measured using the SEDIT sound editing program on a Computer Automation 'Alpha' LSI-2/40 minicomputer with a sampling rate of 10 kHz after analog low pass filtering at 5 kHz. The speech waveform was displayed on a 26 x 34 cm display scope, and two manually-controlled cursors were used for the demarcation of points on the waveform. A statistical analysis of the measurements was then performed, using the Statistical Analysis System package on the Cambridge University IBM 3081 mainframe computer.

#### 2.2.3.3. Results

It was found that prepausal and non-prepausal feet were not comparable, in that prepausal feet generally contained only one or two syllables, while non-prepausal feet could contain up to six syllables. The number of prepausal feet was insufficient for any firm conclusions to be reached, but in general it seemed that, given an equal number of syllables per foot, the prepausal foot tended to be longer. This observation, if reliable, would tally with similar observations made for prepausal feet in English (see Lehist 1977 for a brief summary of such work). Because of this difference, in subsequent results only the non-prepausal feet are considered, these being more numerous and probably 'unmarked' in their duration.

A second finding was that the two sets of sentences seemed significantly different in rhythmicality, as had been hypothesised; however, the difference lay in the opposite direction from that expected, in that the first set showed a greater tendency towards isochrony for both the syllabified and non-syllabified versions. The relevant correlation

coefficient is that between syllable length and number of syllables per foot; this is displayed below, where 'non-syllabified' refers to measurements made from one stressed vowel onset to the next, and 'syllabified' refers to measurements made from one stressed syllable onset to the next.

For non-prepausal feet only: product-moment correlation coefficients between syllable length and number of syllables per foot

	<u>First set</u>	<u>significance</u>	<u>Second set</u>	<u>significance</u>
non-syllabified	-0.45	0.0001	-0.26	0.0072
syllabified	-0.38	0.0001	-0.13	0.2038

A possible explanation for this result may lie in the fact that the first set was at the same time slightly more 'syllable-timed', in that it showed a stronger correlation between foot length and number of syllables per foot in both the syllabified and non-syllabified versions, as displayed below.

For non-prepausal feet only: product-moment correlation coefficients between foot length and number of syllables per foot

	<u>First set</u>	<u>significance</u>	<u>Second set</u>	<u>significance</u>
non-syllabified	0.83	0.0001	0.78	0.0001
syllabified	0.82	0.0001	0.80	0.0001

The reason for these stronger correlations in the first set could be simply that there was less variability in this set. This interpretation is supported by the fact that the significance of the coefficients for the second set in the first of the two tables above is much lower than the significance of the coefficients for the first set; the 'significance' of 0.2038 for the syllabified version is in fact not statistically significant at all, and so this coefficient is most unreliable. The lower significances of the second set, and the fact that both its correlation coefficients are weaker, point to a greater variability in this set, and thus suggest that these particular findings (as regards the difference in variability between the sets) cannot be regarded as conclusive. It may be that the first set was produced more rhythmically than the second, but on the evidence presented above this cannot be said to have been proved satisfactorily, and the question remains open.

A third finding was that, in both sets, the non-syllabified version showed a greater tendency towards isochrony. This tendency was even greater in the prepausal feet, but this could probably be explained by the different possible numbers of syllables per foot in non-prepausal and prepausal feet - no prepausal foot contained more than two syllables, so isochrony would be more likely to be found in such a context. For non-prepausal feet, the correlation coefficients are as follows:

For non-prepausal feet only: product-moment correlation coefficients

correlation between	foot length: no. of syllables per foot	syllable length: no. of syllables per foot
<u>FIRST SET</u>		
non-syllabified	0.8283	-0.4546
significance	0.0001	0.0001
syllabified	0.8152	-0.3789
significance	0.0001	0.0001
<u>SECOND SET</u>		
non-syllabified	0.7814	-0.2644
significance	0.0001	0.0072
syllabified	0.7975	(-0.1302)
significance	0.0001	(0.2038)

It is clear that the non-syllabified versions of measurements show a much greater tendency towards isochrony/ stress timing (from the correlation between syllable length and number of syllables per foot), while there is very little difference between the two types of measurement in the case of the tendency towards syllable timing (the correlation between foot length and number of syllables per foot). The high significances of nearly all the correlation coefficients lend support to this conclusion.

The preceding table also demonstrates that the tendency towards isochrony found in the previous experiment was also found in this set of measurements, thus lending confidence that the previous finding was significant. The tendency is not as strong as that found in speech produced outside experimental conditions (section 2.2.2.) but is nevertheless clearly there, most reliably in the first set. The coefficient of determination (calculated by squaring the correlation coefficient and multiplying by 100) gives a measure of how much of the variability in one factor can be explained by the variability in the other

factor, and can be used to compare the two sets of measurements, as seen below (where  $r$ =correlation coefficient,  $cd$ =coefficient of determination, and all feet are non-prepausal):

<u>Rehearsed speech (first set only)</u>						
	<u>Unrehearsed speech</u>		<u>non-syllabified</u>		<u>syllabified</u>	
	no. of syllables		no. of syllables		no. of syllables	
	per foot		per foot		per foot	
	$r$	$cd$	$r$	$cd$	$r$	$cd$
foot length	0.85	72%	0.83	69%	0.82	67%
syllable length	-0.55	30%	-0.45	20%	-0.38	14%

The meaning of these figures is, for example, that in the non-syllabified version of the measurements of rehearsed speech, 20% of the variation in syllable length could be explained by the variation in the number of syllables per foot.

#### 2.2.3.4. Discussion

Since prepausal and non-prepausal feet were found not to be comparable, the following remarks apply only to non-prepausal feet, the most numerous type. A tendency towards isochrony was observed in this experiment, thus encouraging the view that this tendency is a real phenomenon and not merely a feature of the first set of this series of measurements. Although the tendency is slightly weaker in the second experiment, it is nevertheless present, especially in the first set of sentences recorded.

The null hypothesis that there was no difference in rhythmicity between the two sets of sentences could not be conclusively proved false, because of the probable different degree of variability in each set. Thus this aspect of the experiment was inconclusive.

Perhaps the most intriguing aspect of this experiment is the finding that the non-syllabified versions of measurements showed a greater tendency towards isochrony. This could mean that, for purposes of rhythm, each consonant 'belongs' to the preceding rather than to the following vowel. Such a conclusion is supported by two observations:

a) A consonant after a stressed vowel is lengthened, as described in section 2.2.4.1. and investigated in section 2.3.1.

b) Phonological vowel length in Welsh is determined by the nature and number of the following consonant(s) - see Rhys Jones (1977:21-

22) for a summary of vowel length rules.

For rhythmic and phonetic purposes, therefore, it seems that feet must be counted as starting at the onset of the stressed vowel. This finding seems to contradict two of R.Jones' rules of syllabification (in a recent auditory study of three Welsh dialects), viz.:

- '(1) Onset zero marginal occurs initially only in the working unit. Final zero marginal can occur in any position.
- (2) Statistical evidence favours a C element as the onset marginal of a syllable rather than as the final marginal of a preceding syllable.' (Jones, R.O. 1967:250)

Jones' syllabification rules for Welsh are based on the statistical approach advocated by O'Connor and Trim (1953) for English, and give comparable results to Selkirk's 'Maximal Syllable Onset Principle' for English, whereby a consonant is assigned to the following vowel if at all possible (Selkirk 1980b:9). It could be that on another level of linguistic analysis the rules of syllabification take a different form, so that for purposes of morphological alternation, and especially for the Welsh system of word-initial consonantal mutations, syllables could be seen as beginning with a consonant wherever possible. However, this is as yet merely speculation. What does seem to be clear is that stress in Welsh, and taking vowel onset as the start of the syllable, are linked by the factor of isochrony; and it could be said that this type of syllable-division is more closely linked to rhythm than is the more conventional type of syllable-division.

#### **2.2.4 Segmental measurements**

##### **2.2.4.1. Consonants**

A list of thirty-two isolated regularly-stressed Welsh disyllables with medial voiceless stop was read by the Welsh speaker of section 2.2.3., and was recorded. Measurements of the consonants were made from the tape using the SEDIT program (see section 2.2.3.2.) and a statistical analysis was conducted.

The main result was that the post-stress consonant was significantly longer than the consonant after the unstressed (final) syllable. The mean length of the post-stress consonant was 167 ms, that of the final consonant was 109 ms, a difference of 58 ms at a significance level of  $p < 0.01$ . The post-stress consonant had far lower variability; it tended not to deviate very much from its mean, unlike the final consonant. This could be explained by the following observation.

The proportion of the word taken up by the post-stress consonant tended to decrease as the overall word length increased (correlation of consonant as percentage of word, and word length =  $-0.82$ , significance level  $p < 0.0001$ ). Thus the post-stress consonant was tending towards a constant absolute length. On the other hand, the proportion of the word taken up by the final consonant tended to increase as the word length increased (correlation of consonant as percentage of word, and word length =  $0.80$ , significance level  $p < 0.0001$ ). Thus the final consonant tended to take up a disproportionate amount of the increase in word length, leaving the medial, post-stress consonant relatively unchanged. This finding fits observations made by workers cited by Lehiste (1977).

Perceptual evidence cited by Lehiste (op. cit.) suggests that a human listener 'expects' final segments to be longer and normalises accordingly. Thus it may be that the over-long final consonant in longer words would not be perceived as very long, and so the medial consonant would have more perceptual length than the physical facts would seem to warrant. In words uttered at normal conversational speed, however, the post-stress consonant is longer than the post-unstress consonant; a finding that agrees with the 'extra duration' observed auditorily for post-stress voiceless stops in certain contexts (Jones, R.O. 1967).

#### 2.2.4.2. Vowels

Measurements were also made of the vowels in the unrehearsed speech data. These were classified along three dimensions: prepausal versus non-prepausal, stressed versus unstressed, and phonologically long versus phonologically short. A statistical analysis was then carried out, and the following findings were made.

Prepausal vowels were found to be significantly longer than non-prepausal vowels ( $p < 0.01$ ), with mean lengths of 126 ms and 77 ms respectively. Because of this difference, and since long vowels may only occur in stressed syllables in Welsh, the effect of stress on vowel length was considered in relation to short non-prepausal vowels only. Stressed vowels were found to be significantly shorter than unstressed vowels ( $p < 0.05$ ), with mean lengths of 59 ms and 73 ms respectively. This finding bears out in a more rigorous way the finding concerning duration of the stressed vowel, in sections 2.1.3. and 2.2.1.3. Because of the restriction on the distribution of long vowels, the effect of phonological length on vowel duration was considered in relation to stressed vowels only. No significant difference could be demonstrated here; however, the number of observations in each category was far lower than in the case of the other comparisons, and thus any difference in length would have been more difficult to prove statistically - the mean durations indicated that phonologically long vowels have a longer phonetic duration. Thus this particular comparison must be regarded as inconclusive.

The interesting finding in this set of measurements is that the stressed vowel of the penult is consistently shorter (note: stressed monosyllables are not being considered). The lengthening of prepausal segments has been observed in other languages (see Lehiste 1977 for a summary), and these findings merely confirm the fact for Welsh. It is also to be expected that phonologically long vowels would be longer in duration than phonologically short vowels; though this could not be demonstrated conclusively in the above measurements, the tendency seems to be in this direction. The conclusion from all segmental measurements is that stress in penults is linked with a shortening of the stressed vowel and a lengthening of the following consonant; this is not at all what would have been predicted from the evidence of stress in other languages.

## 2.3 Perceptual experiments

### **2.3.1 Post-stress consonant**

#### 2.3.1.1. Aim

In a previous set of measurements (section 2.2.4.1.), the duration of a consonant after a stressed vowel had been found to be significantly longer than the duration of a consonant after an unstressed vowel. It was decided to test this discovery about speech production in an experiment concerned with speech perception, in order to find out whether the perceptual facts paralleled those of production.

#### 2.3.1.2. Method

Editing This experiment made use of perhaps the only convincing minimal pair in Welsh differentiated solely by stress: namely, ymladd, /əmlað/ 'to fight', and ymlâdd /əmlá:ð/ 'to tire oneself out'. These words were among a list recorded by a (female) native Welsh speaker, in the same conditions as in section 2.1.2.

The SEDIT sound editing program (see section 2.2.3.2.) was then used to isolate one sample of each of the two words in question, and to make modifications in the length of the /m/ segment. For each source word, the length of this segment was first increased to 255 ms by reduplication of a few cycles of the nasal portion of the waveform, and then reduced to 13 ms in steps of 15 or 16 ms at first (i.e. four cycles) and then in steps of 11 or 12 ms (i.e. three cycles) as the nasal became shorter. The original word was then added in each case, making a total of forty stimuli, each half of this list originating from one of the two source words. The editing-down of the sound wave was done in integral numbers of cycles, as otherwise an audible 'click' spoilt the effect; this is the reason for the slight fluctuation in length of each of the two types of step described above.

The stimuli were then processed using a Computer Automation 'Alpha' LSI-2/40 minicomputer to drive a CED 301 digital speech synthesiser

programmed to implement a linear prediction coding model of speech synthesis. The first step of the processing was to reduce the speech data to linear prediction coefficients. This was accomplished by using a specialised program already supplied. The program had the drawback, however, of beginning its reducing to LPC coefficients a little earlier than the point specified by the operator, while also ending its processing a little later than the point specified; the extent of each discrepancy was not known exactly. The (somewhat crude) solution to this problem was to precede and follow each stimulus with approx. 200 ms of low-amplitude machine noise, with the result that a little of this random noise was included with the stimulus (both before and after each vowel) in the linear prediction coefficients. The random noise had, however, been unaccountably amplified by the program, so that the auditory result was similar to a word spoken at moderate loudness on a bad telephone line.

Re-synthesis The next step in processing was to play the stimuli through the speech synthesiser linked to the minicomputer. This was accomplished using a linear prediction coding program already supplied. However, a modification was made to this program so that the speech output was reduced to a constant monotone on all voiced segments. The monotone frequency chosen in this case was 280 Hz. Words spoken on a high  $F_0$  are somewhat easier to hear than speech at a low  $F_0$ , in conditions of low signal-to-noise ratio. Thus a fairly high  $F_0$  was chosen, to improve the audibility of the stimuli above the ineradicable background hiss mentioned above.

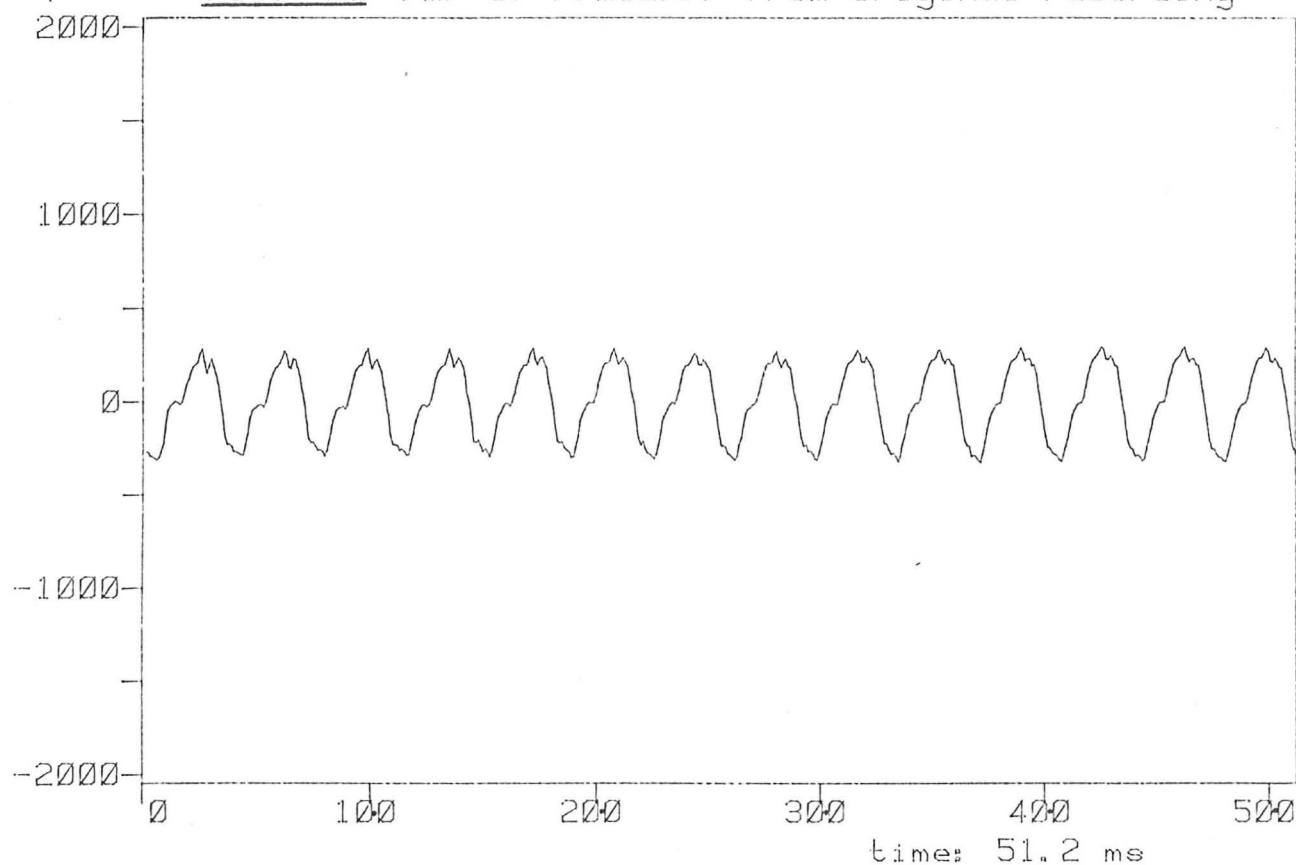
The purpose of reducing the  $F_0$  of the stimuli to a monotone was to factor out all pitch cues to stress perception, so that the experiment would focus only on cues of segmental duration, if any. It was not possible to control the intensity of the stimuli. This had probably been through several transformations in the stages of processing, and it was not immediately obvious how to set about controlling for intensity. The intensity of the stimulus words was probably only indirectly related to the intensity of the source words. However, the measurements described in sections 2.1. and 2.2.1. had indicated that mean and peak amplitude, and the integral of the amplitude, were most unreliable for the purposes of stress identification, and so it was assumed that intensity would not be a

significant factor in stress judgements. Fig. 4 presents a portion of the waveform of the nasal both before and after processing by the synthesiser. It will be noted that although the higher frequencies have been lost and the amplitude reduced, the waveform is a passable one for a nasal. Similarly, Fig. 5 presents a portion of the waveform of the /a/ in ymlâdd before and after synthesis; it will be noted that no great distortion has been introduced.

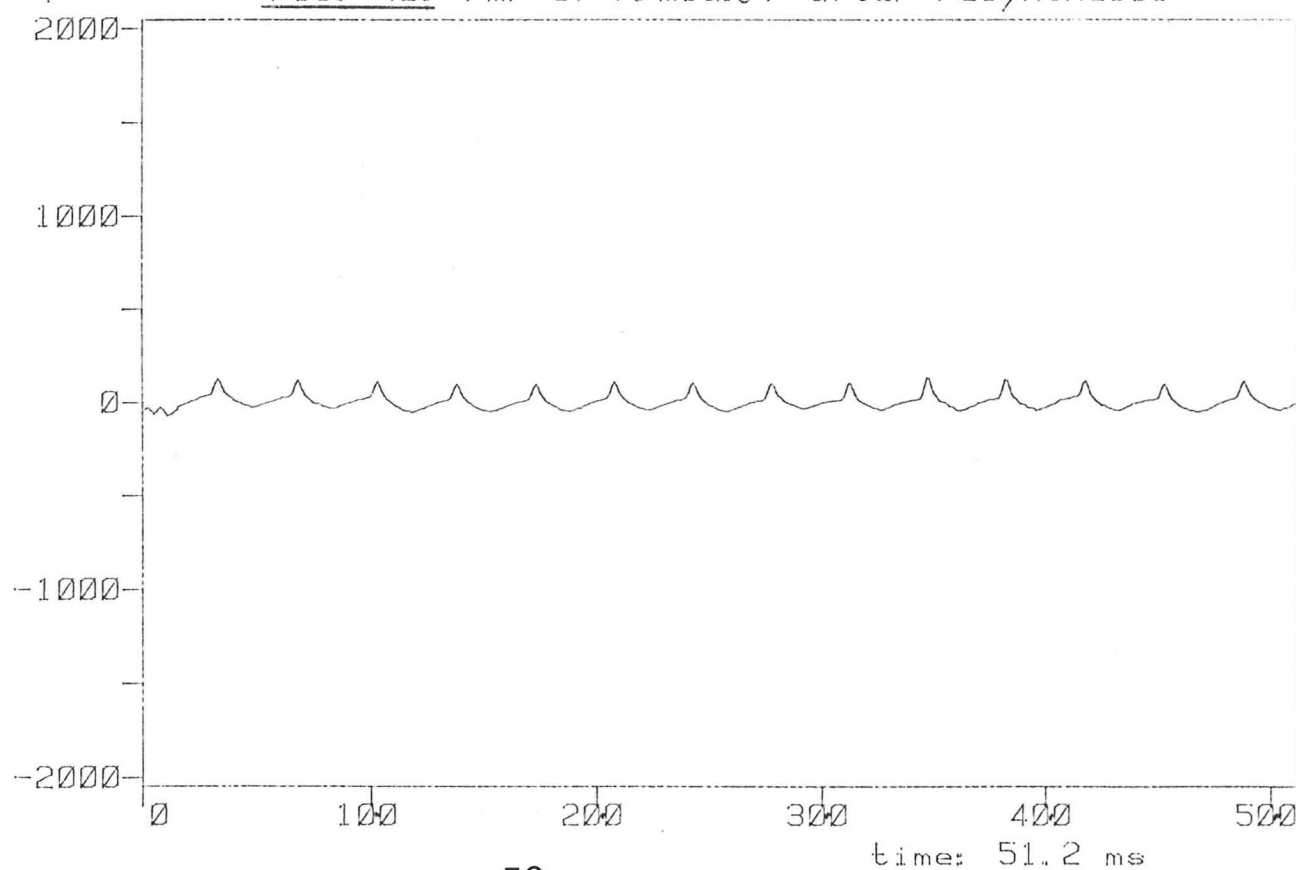
Experimental tape The synthesised monotone stimuli were recorded on an open-reel tape recorder (Ferrograph series seven Mark 2) as they were played out, and were then re-recorded, in random order, on an Ampex AG 440 open-reel tape recorder. The pattern of re-recording was as follows. The list of stimuli was introduced by a few sentences in Welsh describing the nature of the task, but without any 'cue word' or spoken example. Each stimulus was preceded by its number, spoken in Welsh. After the number there was a short pause, and then the first instance of the stimulus, after which about two or three seconds' silence were left on the tape before the second instance of the same stimulus. After this was another three seconds (approx.) of silence before the number of the next stimulus. The silent gaps were to allow the Welsh listeners time to make judgements and note them down. After the list of forty words had been re-recorded in this way, the phrase 'please turn over' was spoken in Welsh, and the next list began (see section 2.3.2.). This list was the first of four similar lists of stimuli.

To accompany the experimental tape, response sheets were prepared for each listener. These included the explanatory Welsh words, typed at the top of the page, with an example. The two alternatives, ymladd and ymlâdd, headed a column each with a paraphrase in Welsh beneath each word, and space was left by each stimulus number for a tick to be made beside the word identified by the listener with the stimulus. The experimental tape was then copied onto a chromium dioxide TDK SA60 cassette tape, and this cassette tape was the one used in the experiment. A group of ten female native Welsh speakers, all students at the Normal (i.e. teacher-training) College at Bangor, in the north of Wales, then listened simultaneously to the tape using headphones in the Language Learning Laboratory of the Normal College, and marked their judgements on the response sheets, having

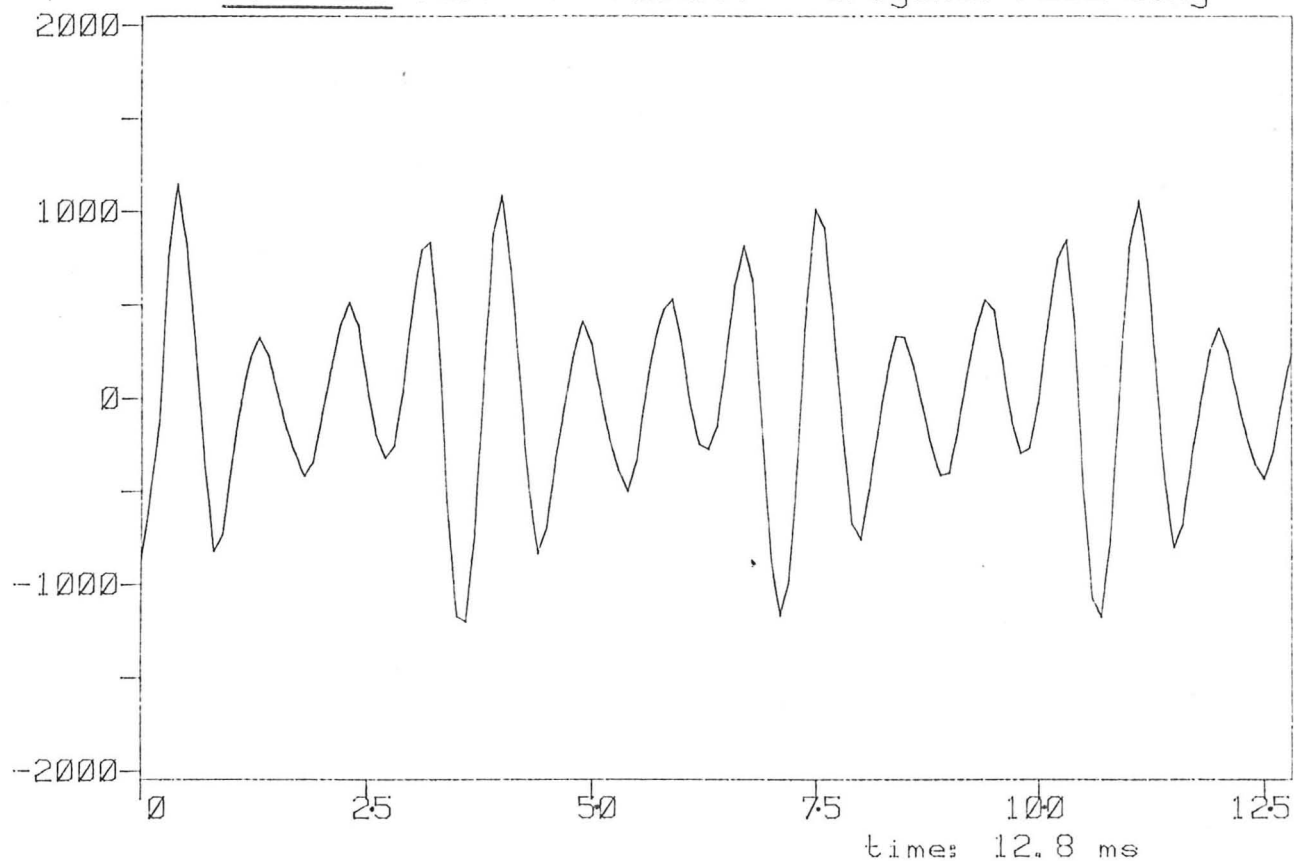
amplitude FIG. 4a: /m/ of /əmlá:ð/ from original recording



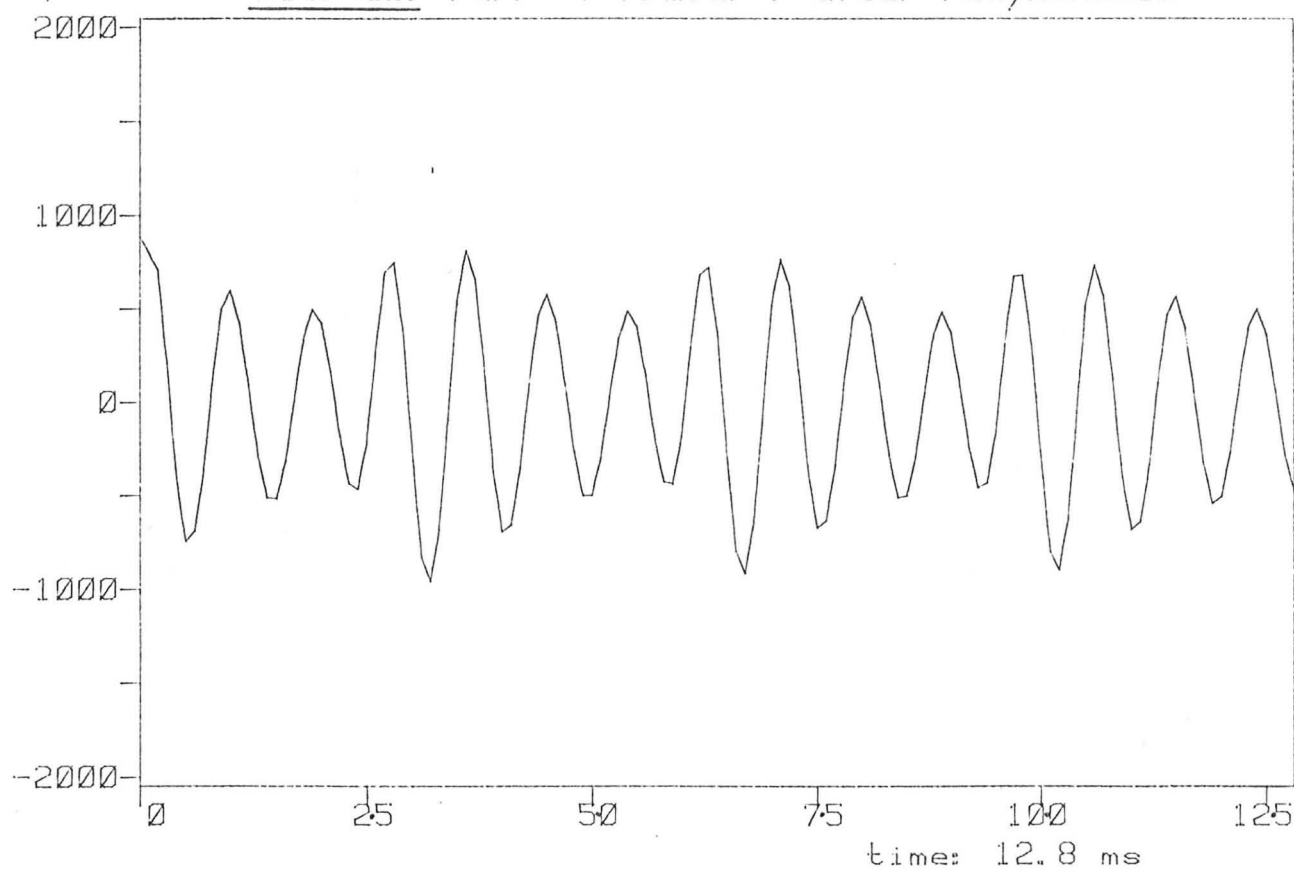
amplitude FIG. 4b: /m/ of /əmlá:ð/ after resynthesis



amplitude FIG. 5a: /a:/ of /əmlā:ð/ - original recording



amplitude FIG. 5b: /a:/ of /əmlā:ð/ after resynthesis



been asked to identify the stimulus they heard as one of the two possibilities. A statistical analysis of the stress judgements was then carried out, using the Statistical Analysis System package on the Cambridge University IBM 3081 mainframe computer.

Variables and indices The length of the nasal was calculated as a percentage of the length of the word and formed one variable. Other variables were: source word (with two possible original words) and the number of subjects who judged each stimulus to be ynlâdd - since this was a forced-choice situation, it follows that the number judging the stimulus to be ynladd would be the remainder out of ten.

The responses of each subject were then checked to test for marked deviation from the other subjects' judgements. Any case of consistent disagreement would have led to the subject(s) in question being discounted in the final statistical analysis. To this end, the lists' 'majority sites' were identified; these were stimuli which at least 50% of the subjects had judged to be ynlâdd. For each subject, three indices were then calculated, as follows:

- a) The first index was gained by dividing the number of stimuli marked by the subject that turned out to be majority sites, by the number of ynlâdd judgements made by the subject. A low figure would indicate that many of the subject's ynlâdd judgements were non-majority sites: in other words, that she tended not to agree with the others. This index gave no information as to whether the subject managed to locate many of the total number of possible majority sites.
- b) The second index was gained by dividing the number of stimuli marked by the subject that turned out to be majority sites, by the number of majority sites contained in the list as a whole. A low figure would indicate that the subject did not locate many of the possible majority sites: in other words, that where she did agree with the others, this could not be to any large extent. This index gave no information as to whether many of the subject's ynlâdd judgements were not majority sites.

c) The third index was gained by dividing the number of ymlâdd judgements made by the subject, by the number of stimuli in the list that were judged ymlâdd by at least one subject. A low figure would indicate that the subject did not hear many items as ymlâdd at all: in other words, that she tended to hear ymladd (stress on penult) whatever the original word. This index gave no information as to what proportion of the subject's ymlâdd judgements turned out to be majority sites.

Fig. 6 shows the resulting indices by subject. The calculations were taken over all four lists of stimuli (see sections 2.3.2. and 2.3.3. for details of the other three lists). It will be noted that there is no great deviation between subjects in mean values over all indices. Because of this reasonable degree of consistency, the judgements of all the subjects were included in the statistical analysis. The mean value of the third index is very slightly above 0.5, which might seem to suggest that, over all the lists taken together, the subjects tended to make slightly more ymlâdd judgements; however, this figure was due to the influence of list 3, as will be seen later (section 2.3.3.). The mean value of the first index is high, which suggests that subjects tended to agree amongst themselves in their ymlâdd judgements. The mean value of the second index is also high, which suggests that individual subjects tended to locate most of the majority sites.

#### 2.3.1.3. Results

Nasal length effect It was found that the length of the post-stress nasal had a significant influence on stress judgements, as had been suggested by the preliminary measurements described in section 2.2.4.1. The (product-moment) correlation coefficient between the length of the nasal (as a percentage of the word length) and the number of ymlâdd judgements made, was  $-0.54$ , at a significance level of  $p < 0.0004$ . This means that the greater the proportion of the word taken up by the nasal, the more likely the word was to be judged as ymladd (stress on the penult): in other words, listeners expect the consonant to be longer when it follows a stressed vowel.

Fig. 6: indices by subject

index	subject no.:-					
	1	2	3	4	5	6
a	0.85	0.84	0.79	0.79	0.87	0.81
b	0.73	0.78	0.78	0.78	0.78	0.68
c	0.51	0.55	0.61	0.59	0.54	0.51
mean a+b	0.79	0.81	0.79	0.78	0.83	0.74
	7	8	9	10	means	
a	0.72	0.88	0.72	0.85	0.81	
b	0.74	0.81	0.70	0.82	0.76	
c	0.61	0.57	0.58	0.58	0.57	
mean a+b	0.73	0.85	0.71	0.84	0.79	

Fig. 9: indices by list no.

index	list no.:-				means
	1	2	3	4	
a	0.4359	0.5714	0.6944	0.7097	0.6029
b	0.4250	0.5714	0.6250	0.5500	0.5429
c	0.9750	1.0000	0.9000	0.7750	0.9125
mean a+b	0.4305	0.5714	0.6597	0.6299	0.5729

see text for description of indices and their calculation:

Fig. 6 - section 2.3.1.2., under 'Variables and indices'

Fig. 9 - section 2.3.3.3., under 'Calculating indices'

Source word effect Simultaneously with the nasal length effect there was also a significant source word effect, whereby the original forms of the stimulus words influenced listeners' stress judgements. This was found by calculating the (product-moment) correlation coefficient of source word number and number of ymlâdd judgements. The coefficient was 0.7136, at a significance level of  $p < 0.0001$ , indicating a moderate source word effect on stress judgements. A possible cause of this source word effect will be discussed in section 2.3.3.3.

#### 2.3.1.4. Discussion

This experiment provided support for the notion that stress in Welsh is accompanied by an increase in the length of the following consonant. The experiment supported in perceptual terms what had previously been discovered only from the speech production point of view. It also brought out a previously unsuspected source word effect which, however, was not strong enough to obliterate the nasal length effect.

### 2.3.2 Phonological segmental length

#### 2.3.2.1. Aim

The second perceptual experiment was of a different kind, being concerned not with stress but with phonological consonantal length. The aim was to discover whether or not the duration of the nasal plays any part in differentiating between the two words tonau /tɔːnaɪ/ 'tunes' and tonnau /tɔːnaɪ/ 'waves'. The first word is traditionally regarded as having long /o/ and short /n/, while the opposite is the case for the second word. The experiment was designed to test for any objective basis for the distinction between long (or 'geminate') and short /n/ in these two words.

#### 2.3.2.2. Method

The procedure was much the same as for the first perceptual experiment (see section 2.3.1.2.). The two words were recorded by the same speaker and under the same conditions as before. They were then edited by the SEDIT program so that the duration of the first vowel of each word was 120 ms - an intermediate stage between the original /ɔ/ of 95 ms and /oː/ of 139 ms.

The length of the nasal in each word was increased to 250 ms and then reduced to 14 ms in steps of 13 or 14 ms at first (i.e. 3 cycles) and then in steps of 9 or 10 ms (i.e. 2 cycles) as the nasal became shorter. The completed stimuli were then reduced to linear prediction coefficients, adding a stretch of machine noise before and after each word, as before. The stimuli were output on the speech synthesiser at a monotone of 280 Hz, and were recorded on open-reel tape. They were then re-recorded in a different random order from the first list, using introductory sentences and numbers, and pauses between each item, as in the first experiment of this kind. Response sheets of the same format as before were prepared, and the ten listeners were asked to identify the stimulus they heard as one of the two possibilities. During the running of the experiment, there was a minute or so of pause between each of the lists on the experimental tape, during which the ten listeners conversed with one another in Welsh.

A statistical analysis of the judgements was carried out as before. The length of the nasal was calculated as a percentage of the length of the word, and formed one variable. Other variables were: source word (two possibilities), and the number of subjects who judged each stimulus to be tonau (with short /n/); since this was a forced-choice situation, it follows that the number judging the stimulus to be tonnau (with long /n:/) would be the remainder out of ten.

#### 2.3.2.3. Results

There was no nasal length effect found at all; the duration of the nasal as a percentage of word length had no correlation with the number of tonau judgements. This is in sharp contrast to the findings of the first perceptual experiment, where a clear and significant nasal length effect was discovered.

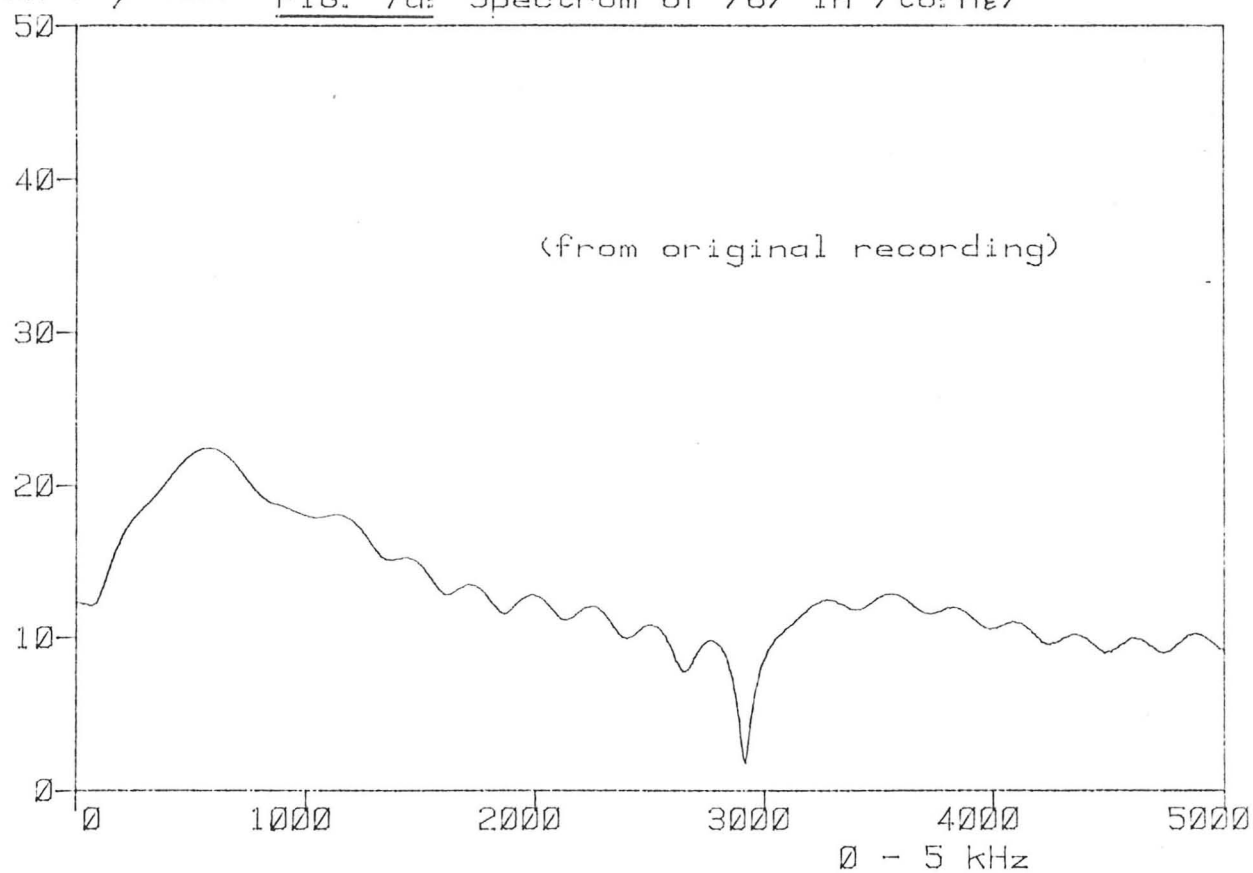
There was a strong source word effect; the source word from which a stimulus originated had a strong influence on the number of tonau judgements. The correlation coefficient for these two variables was 0.7608, at a significance level of  $p < 0.0001$ , indicating that this was a reliable and significant finding.

#### 2.3.2.4. Discussion

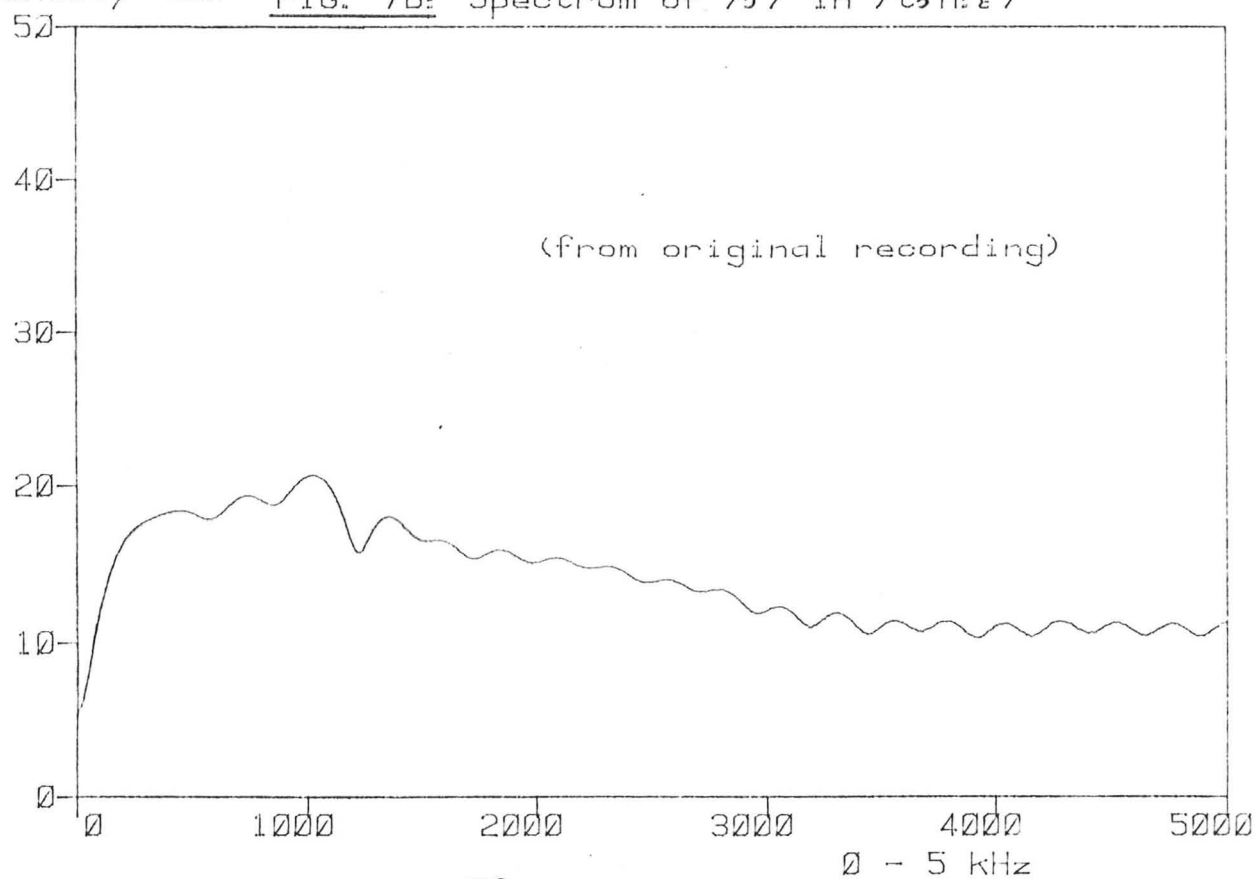
The only factor operating in this second list of stimuli was the source word effect. Clearly, phonological segmental length has no basis in phonetic facts in the case of the nasal in these two words. The cause of the source word effect in this case is probably the different quality of the two 'o' vowels. Spectra were made of the unprocessed forms of the two vowels in question, using an already-provided program on the minicomputer. Fig. 7 gives a comparison of these two spectra; it can be seen that they differ in formant structure and are thus probably perceptually distinct. Such a finding agrees with the present-day situation in the Welsh language, where the old distinction between single -n- and geminate -nn- is dying out. It is never used in colloquial speech, and not always used in careful speech, the quality of the vowel taking over the whole burden of differentiation in such cases. Thus the results of this experiment provide objective support for this auditorily-observed tendency in modern Welsh. Although the results may seem disappointing, in reality they are of great interest, since they indicate that, even given the whole range of nasal length, the vowel quality alone is sufficient to condition judgements. This strength of the vowel quality cue gives a plausible reason why the nasal length cue is dying out in this word-pair - it is simply not needed, as listeners can differentiate sufficiently from vowel quality cues. In any language, what is not needed is soon lost, and Welsh seems to be no exception.

In order to investigate whether vowel quality were also the cause of the source word effect in the first word-pair, spectra were made of the unprocessed forms of the second vowel of ymhladd and ymhlâdd. Fig. 8 shows a comparison of these spectra, and it is clear that there is little, if any, discrepancy between them. To all intents and purposes. they are the same vowel, although articulatorily there is supposedly a difference in closeness - short /a/ being [a<sup>+</sup>] and long /a/ being [a<sup>+</sup>] (Jones, R.O.:279). However, any articulatory difference there may be is obviously insufficient to give rise to an acoustic difference, and so it can be assumed that there was no auditory difference between the two vowels in this experiment. Therefore some other cause of the source word effect would have to be found (see section 2.3.3.3.).

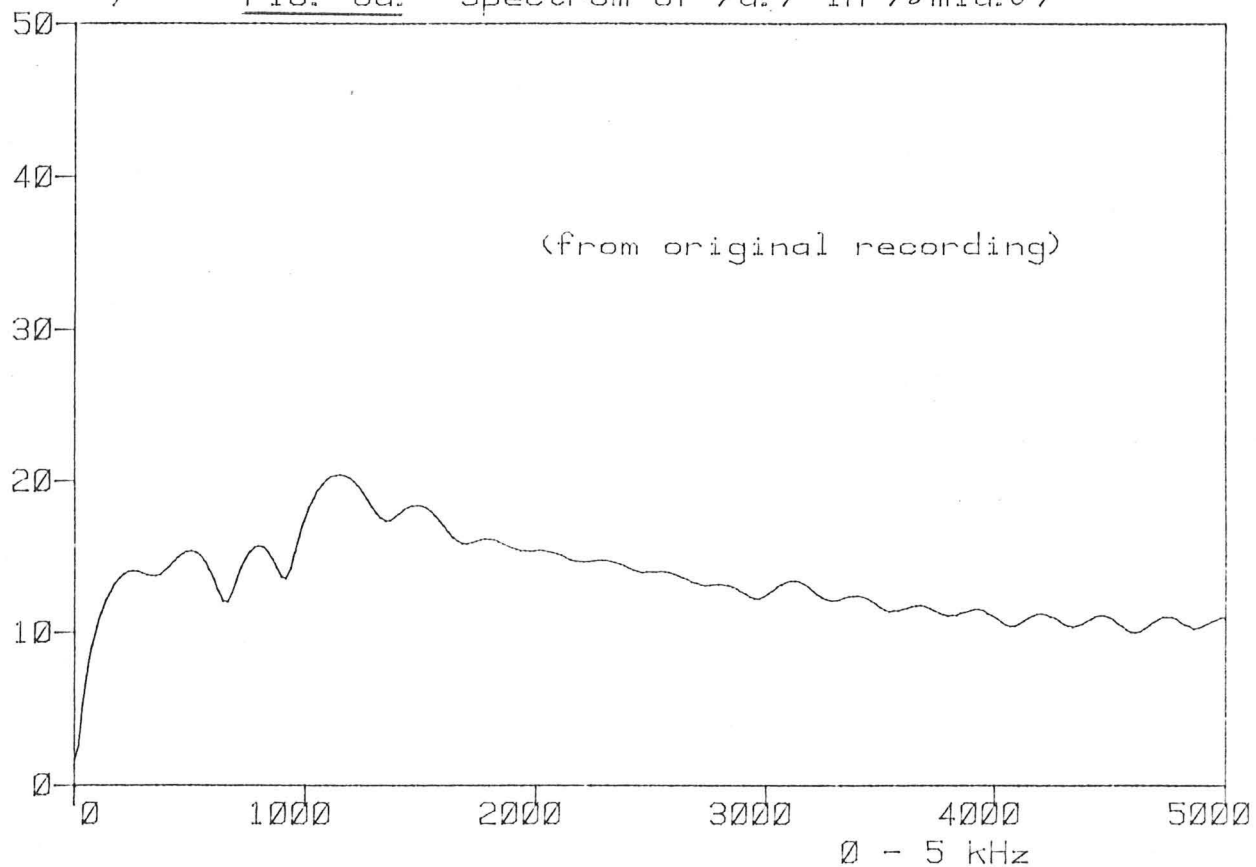
intensity (dB) FIG. 7a: Spectrum of /o/ in /to:nε/



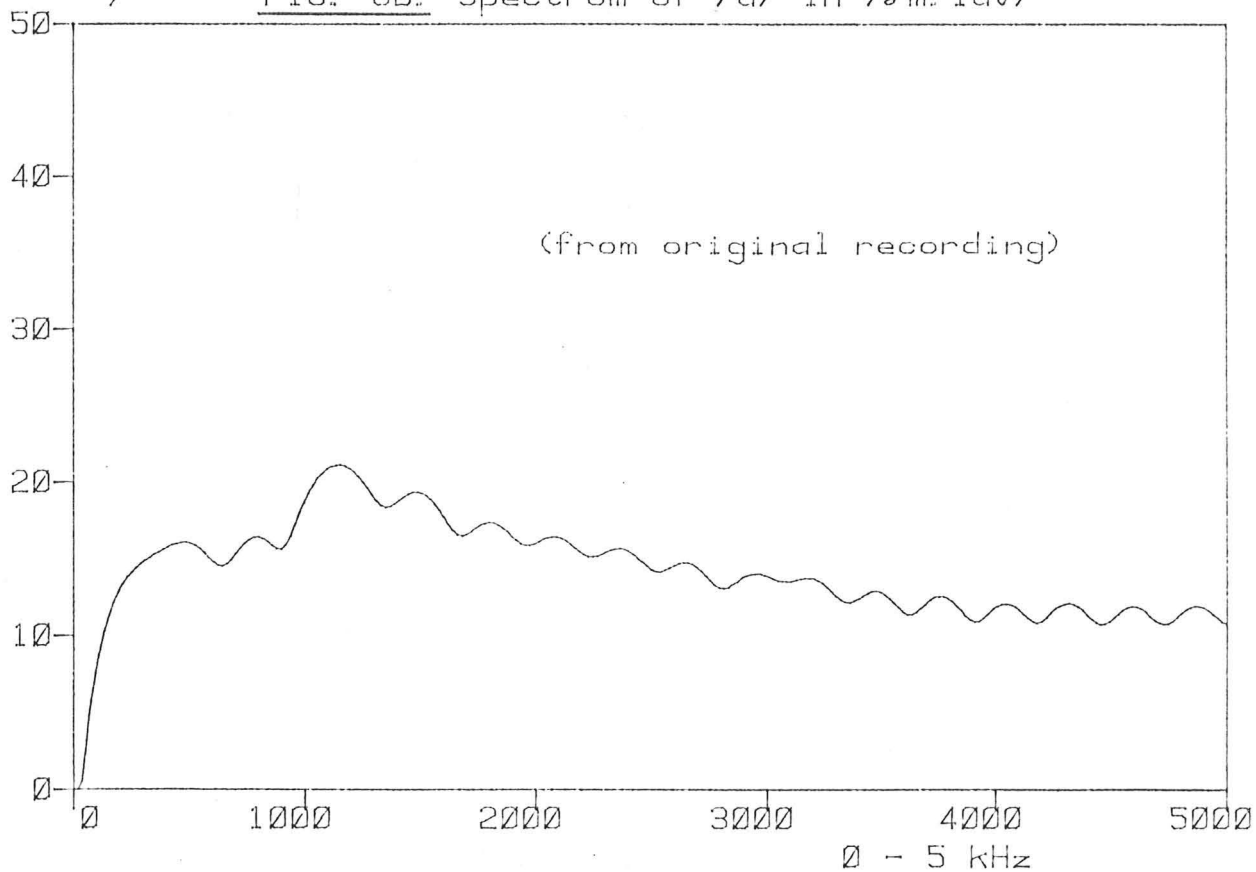
intensity (dB) FIG. 7b: Spectrum of /o/ in /tɔ:nε/



intensity FIG. 8a: Spectrum of /a:/ in /əmlá:ʃ/



intensity (dB) FIG. 8b: Spectrum of /a/ in /əmla:ʃ/



### 2.3.3 Effect of pitch-prominence

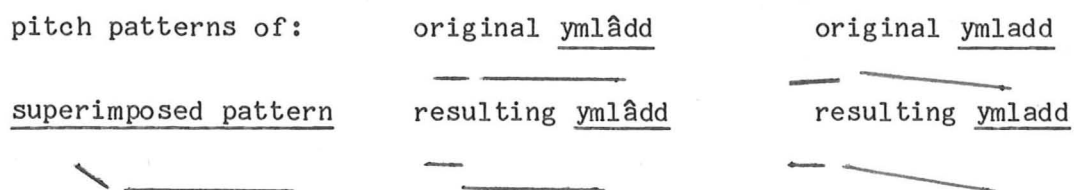
#### 2.3.3.1. Aim

Having established a post-stress consonant length cue to stress in words presented on a monotone, the next step was to test for any variation in this cue due to the pitch-prominence of one or the other of the two syllables. Pitch-prominence was seen as either higher pitch on a syllable, or a pitch glide. This was in accordance with Crystal's definition for English, where any deviation from the 'norm' (which is a slight step downwards in pitch) constitutes pitch-prominence (Crystal 1969).

#### 2.3.3.2. Method

Third list The minimal pair ymlàdd/ ymladd was again used, being edited and reduced to linear prediction coefficients as described in section 2.3.1.2. Upon re-synthesis, however, no monotone was used. Instead, an Fo pattern was superimposed on the synthesiser output using a digitiser pad (Summagraphics 'Bitpad one') linked to the synthesiser via the minicomputer. A manually-operated 'pen' linked to the digitiser tablet was used to superimpose an Fo pattern on the speech as it was played out by the synthesiser, this output being simultaneously recorded on open-reel tape. The digitiser also had the effect of slowing down the speech slightly; however, as the nasal length was calculated in percentage rather than absolute terms, this did not affect the method of analysis.

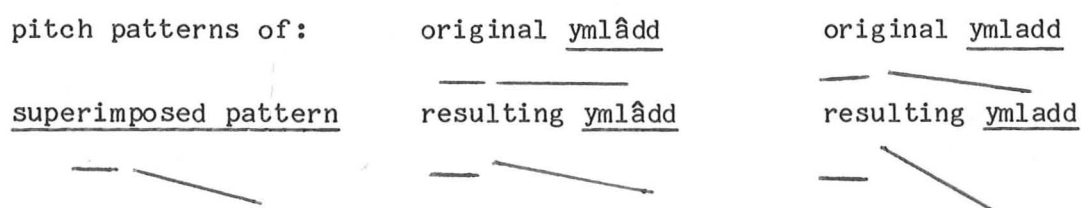
Two types of Fo pattern were superimposed on the speech output. In the first pattern, the superimposed curve was a downglide during the first syllable and a steady lower pitch during the second syllable. Since the source words had different original pitch patterns, the result differed according to source word, as follows:



The two pitch patterns resulting from this process had in common the fact that the second syllable did not begin at a higher pitch than the first. It was possible to distinguish auditorily between the two patterns, but only when they were listened to closely by someone familiar with the tracking of intonation patterns. The stimuli from source word ymlâdd began the first syllable at approx. 315 Hz and kept this fairly constant, while the second syllable began at approx. 310 Hz and stayed level. The stimuli from source word ymladd began the first syllable at approx. 280 Hz and kept this fairly constant, while the second syllable began at approx. 315 Hz and fell to approx. 250 Hz - the only pitch change that could be perceived. The first syllable in each case was too short to carry a change in  $F_0$ . The perceptual effect, on careful listening, was as in the diagrams above.

The list of stimuli was re-recorded in a different random order from the first two lists, in the same format as before. It was presented to the listeners as described earlier (section 2.3.1.2.), with a short break after the preceding list. A statistical analysis was then carried out as before.

Fourth list The same procedure was carried out again with a different superimposed pitch pattern. This second pattern consisted of a level first syllable and a falling ultima. The result differed slightly according to source word, as follows:



The two pitch patterns resulting from this process had in common the fact that the ultima began at a higher pitch than the penult, and then fell in pitch. Again, it was possible to distinguish auditorily between the two patterns, but only on careful listening. The stimuli from source word ymlâdd began the penult at approx. 305 Hz, while the ultima began at approx. 330 Hz and fell to approx. 300 Hz. The stimuli from source word ymladd began the first syllable at approx. 300 Hz, while the ultima began at approx. 350 Hz and fell to approx. 270 Hz. The penult in each case was

too short to carry a change in Fo. The perceptual effect, on careful listening, was as in the diagrams above. The list of stimuli was re-recorded and presented as described earlier, and a statistical analysis was carried out.

### 2.3.3.3. Results

Calculating indices For each of the four lists, indices were calculated based on listeners' judgements. The indices were of two kinds: a) to indicate the degree of agreement between listeners (and thus, indirectly, the difficulty of the task), and b) to indicate the accuracy with which the listeners pinpointed the source word of each stimulus (thus giving a measure of any source word effect there might be). The first group of indices (see Fig. 9 on page 54) were three in number, as follows:

- a) The first index was gained by dividing the number of majority sites in the list by the number of stimuli in the list that had been judged as ymlâdd by at least one listener. A low figure would indicate that many ymlâdd judgements were made on non-majority sites: in other words, that there was a wide degree of 'scatter' of the judgements, and not much agreement. This index gave no information as to whether there were many majority sites in the list.
- b) The second index was gained by dividing the number of majority sites in the list by the number of stimuli in the list. A low figure would indicate that not many stimuli were majority sites: in other words, that there was little agreement among subjects. This index gave no information as to whether there was much 'scatter' of judgements, or as to the total number of ymlâdd judgements made.
- c) The third index was gained by dividing the number of stimuli in the list judged as ymlâdd by at least one subject, by the number of stimuli in the list. A low figure would indicate that only a small proportion of the stimuli were capable of being judged as ymlâdd: in other words, that there could be an unknown biasing factor. A high figure would indicate that a high proportion of the stimuli were capable of being judged as ymlâdd - this turned out to be the case, and thus showed the need to use the 'majority site' system instead of working with absolute numbers of ymlâdd judgements, since only thus could consistencies in the judgements be followed. This index gave no information as to the degree of consistency of judgements, but merely established the validity of using the 'majority site' system.

The second type of index could be termed an 'accuracy index', and was obtained by dividing the number of stimuli that were correctly judged ymlâdd or tonau by each subject, by the total number of ymlâdd or tonau

judgements made by the subject. A low figure would indicate that the subject was not successfully locating the stimuli originating from ymlâdd or tonau: in other words, that she was not being influenced by the source word. This index gave no information either as to whether the subjects agreed consistently with one another in their source word location, or as to whether each subject made many ymlâdd or tonau judgements at all. Fig. 10 shows these indices broken down by subject and list number.

Results of indices From Fig. 9 it can be seen that the first list showed the lowest degree of agreement among subjects, while the third list showed the highest degree. The high values for the third index indicated the need to use the 'majority site' system to track down any consistencies in judgements. The degree of agreement between subjects will later be seen to be important in considering the source word effect.

From Fig. 10 it can be seen that the subjects differed in their susceptibility to the source word effect, but there was no outstanding deviation, so no subject was discounted. More importantly, the source word effect was seen to differ between the lists, such that it was strongest in the fourth list and weakest in the third list. This will later be useful information in considering the cause of the source word effect.

Nasal length effect The length of the medial nasal had a significant influence on stress judgements in both the third and the fourth lists. In the third list the (product-moment) correlation coefficient between the nasal length as a percentage of word length, and the number of ymlâdd judgements made, was  $-0.57$  at a significance level of  $p < 0.0002$ . In the fourth list, the correlation coefficient was  $-0.37$  at a significance level of  $p < 0.0208$ . It will be seen that the correlation in list 4 is weaker and slightly less reliable; however, it is still statistically significant, and a real effect.

Source word effect A moderate source word effect was found for list 3 (though it was at its lowest for this list), as can be seen from the mean 'accuracy index' in Fig. 10 of 0.68. For this list, also, the correlation between source word and number of ymlâdd judgements was 0.65 ( $p < 0.0001$ ). A much greater source word effect was found for list 4, as can be seen from

Fig. 10: 'accuracy indices' by subject and list no.

list no.	subject no.:-					
	1	2	3	4	5	6
1	0.78	0.82	0.65	0.70	0.87	0.77
2	0.53	0.95	0.67	0.81	0.50	0.67
3	0.79	0.57	0.81	0.62	0.70	0.50
4	0.94	0.95	0.95	0.76	0.82	0.67
means of 1,3,and 4	0.84	0.78	0.80	0.69	0.80	0.65

	7	8	9	10	means
1	0.57	0.65	0.60	0.94	0.74
2	0.43	0.93	0.93	0.48	0.69
3	0.67	0.77	0.50	0.83	0.68
4	0.73	0.83	0.63	0.82	0.81
means of 1,3,and 4	0.66	0.75	0.58	0.86	0.74

see text for description of indices and their calculation:

Fig. 10 - section 2.3.3.3., under 'Calculating indices'

the mean 'accuracy index' in Fig. 10 of 0.81. For this list, also, the correlation between source word and number of ymlâdd judgements was 0.85 ( $p < 0.0001$ ). Clearly there is a source word effect for both lists, though far more so in the fourth list.

The cause of this source word effect cannot be differing formant structure of the vowels (see section 2.3.2.4.). The most obvious explanation, at first sight, would seem to be the perceptibly different pitch patterns of the stimuli according to each word. However, this is not such a satisfactory explanation as it may seem. From the indices in Fig. 9 it has been concluded that: a) the third list showed the highest degree of agreement among subjects but the weakest source word effect - in other words, the listeners were consistent in marking the 'wrong' thing; and that b) the fourth list showed a moderate degree of agreement among subjects and the strongest source word effect - in other words, the listeners were marking the 'right' thing, but less consistently. Such a pattern points to the existence of at least one other factor that is 'interfering' with the source word effect and preventing it from becoming the main factor.

Another argument against the 'pitch pattern' explanation of the source word effect would be the fact that lists 3 and 4 are so different in their results. If the pitch patterns were perceptually distinct to an identical degree, it would have been expected that the source word effect would also be present to an identical degree in the two lists. In fact, the judgements show a wholly different pattern, as regards both consistency and accuracy, while the pitch patterns do not show any marked degree of difference according to source word; for each list, the resulting Fo curves have major points in common, chiefly concerned with whether or not the ultima begins at a higher Fo than the penult, and whether the general pattern of the ultima is a low or a high fall/step-down (to use British school terms of intonational analysis). Moreover, the source word effect was also found in list 1, where there was no Fo movement at all. Thus the differing pitch patterns cannot be said to have given rise to the source word effect.

An alternative explanation for the source word effect is provided by features of segmental length, in particular of the vowel /a/ in the two source words. While the segments /ə/, /l/ and /ð/ had much the same duration in the two source words, the vowel /a/ was 201 ms in ymlâdd and 145 ms in ymladd, a difference of 56 ms. While the first vowel of tonau and tonnau had been equalised in duration, this had not been done for the /a/ of ymlâdd and ymladd, and thus all the stimuli showed this difference in vowel length. The conclusion must be that the source word effect was conditioned solely by this difference in vowel length between the two source words.

#### 2.3.3.4. Discussion

The effect of adding a pitch pattern, rather than presenting stimuli on a monotone, seems to be to increase the degree of agreement between subjects (as can be seen from the high values for the first and second indices of lists 3 and 4 in Fig. 9). Whether this increase in agreement is in the same direction for each of the two pitch patterns is another matter. The addition of pitch patterns also seems to polarise the trends of judgements, in that the nasal length effect and the source word effect are at their strongest and weakest in one of the lists with Fo patterns introduced.

The Fo pattern of list 3 seems to shift judgements towards ymlâdd, as can be seen from its high value for the second index in Fig. 9. This pitch pattern also allows the nasal length cue to come into its own, as the correlation between nasal length and ymlâdd judgements is at its highest in list 3 (being -0.57 at a very high significance level of  $p < 0.0002$ ). The degree of consistency between subjects is at its highest, while the source word effect is at its weakest (see section 2.3.3.3.), though it is still significant. These results, taken together, suggest that the pitch pattern of list 3 is in some sense a 'marked' intonation pattern, since it conditions the choice of a particular member of a minimal pair, in a consistent fashion, and forces <sup>s</sup>listeners to make use of all possible cues, e.g. the length of the medial nasal. It seems that a syllable that does not begin at a higher pitch than the preceding syllable is more likely to be judged stressed, and that this tendency is a consistent one.

The Fo pattern of list 4 does not shift judgements towards either ymlâdd or ymladd, as can be seen from its medium value for the second index in Fig. 9. This pitch pattern allows the length of /a/ (source word effect) to come into its own, as the correlation between source word and number of ymlâdd judgements is at its highest in list 4 (being 0.85 at a significance level of  $p < 0.0001$ ). The degree of consistency between listeners is moderately high, while the nasal length effect is at its weakest (see section 2.3.3.3.), though it is still significant. These results, taken together, suggest that the pitch pattern of list 4 is in some sense an 'unmarked' intonation pattern, since it does not condition a choice of either member of a minimal pair, and forces listeners to make use of the cue of /a/ length. In the Welsh vowel system, all long vowels have a short counterpart which differs from them also in quality, except in the case of long and short /a/. Here, it is length alone that differentiates reliably, and thus the listeners are in fact depending on a segmental phonological difference rather than on the pitch pattern in making their judgements.

The fact that a source word effect was found is in fact advantageous from two points of view. First, it makes it even more remarkable that a reliable nasal length cue to stress was found at all, since the source word effect is working against the cue. This increases confidence that the discovered nasal length cue is a real effect and not an artefact of the experiment. Secondly, it makes the experimental stimuli approximate more closely to actual conditions of language use. This means that any effects found (nasal length cue, pitch pattern effects) are more likely to be operative outside the experimental situation. Thus what may have seemed to be a biasing effect in fact confirmed the strength and reliability of the observations made, and related the experimental findings a little more closely to conditions of language use.

## 2.3.4 Comparison of perceptual experiments

### 2.3.4.1. Nasal length effect

In the following discussion, the various effects will be considered only in relation to the ymlâdd/ ymladd pair, i.e. for lists 1,3 and 4 only.

List 2 has been omitted from the discussion because of the anomalous trend of judgements in this case, which showed that only the source word effect was operative (though the cause of the source word effect here was different from its cause in the other three lists, the two causes may be traced back to different realisations of segmental phonological differences according to vowel type - see section 2.3.3.3.). Because of this difference, and also because list 2 showed no nasal length effect, this list will not be further considered.

The nasal length effect was found to be significant in all three lists under consideration, though it varied in strength between the lists. The table below gives the (product-moment) correlation coefficients between nasal length as a percentage of word length and the number of ymlâdd judgements, together with the significance of the correlation. It also gives the coefficient of determination, which is derived by squaring the correlation coefficient and multiplying by 100, and which gives a measure of how much of the variability in one factor can be explained by the variability in the other factor.

Nasal length effect by list number

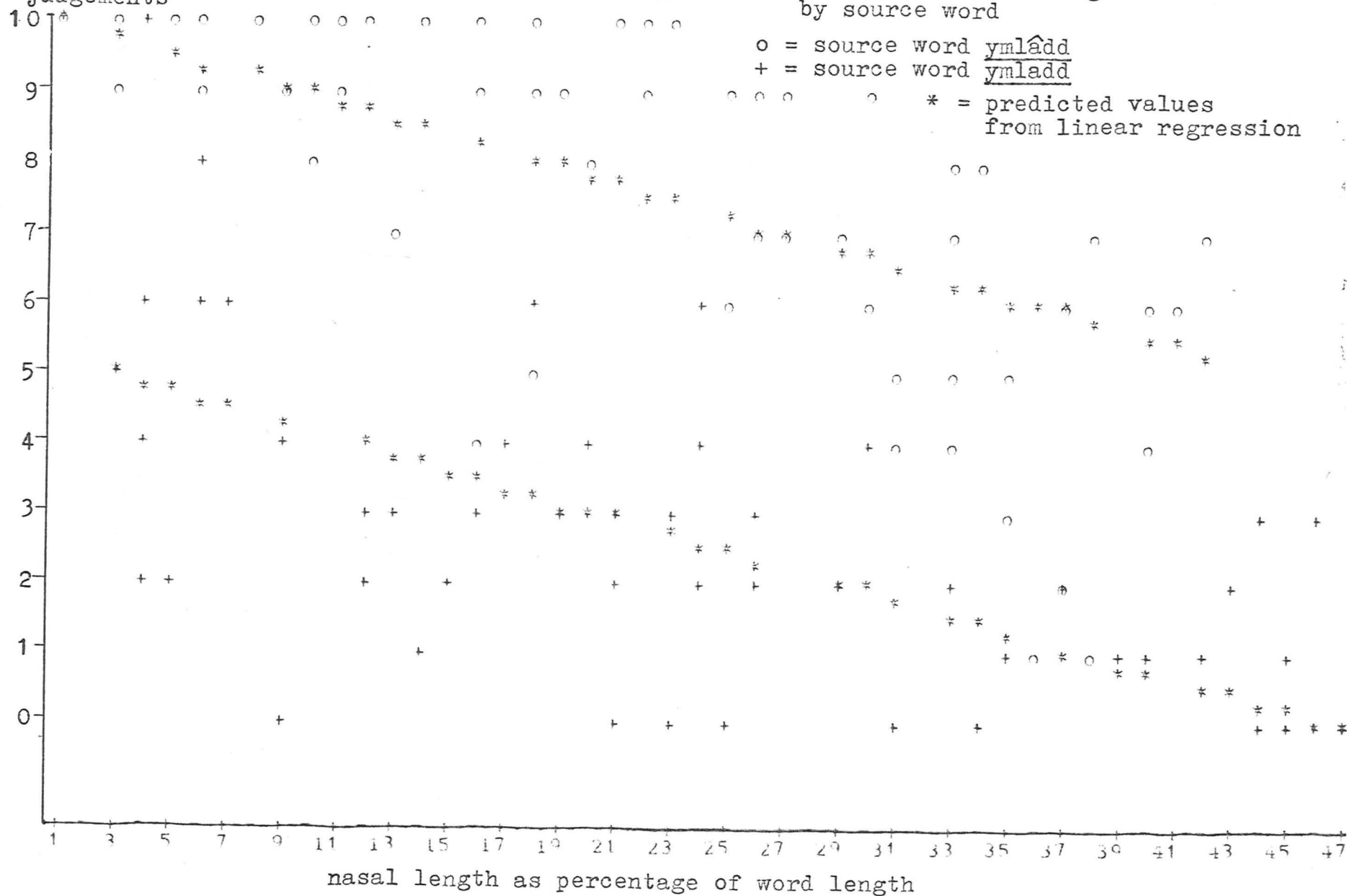
<u>list</u>	<u>correlation</u>	<u>significance</u>	<u>coeff. of determination</u>
1	-0.5444	0.0004	30%
3	-0.5703	0.0002	33%
4	-0.3738	0.0208	14%

It will be noted that the correlation is very similar for lists 1 and 3, but much weaker and a little less significant in list 4. This means that, for the pitch pattern of list 4, it is less likely that the ultima will be more perceptually 'stressed' as the length of the medial nasal decreases. However, the factor is still present in list 4 at an acceptable degree of significance.

From inspection of graphs of number of ymlâdd judgements against nasal length, it seemed that the nasal length effect was only reliable when the length of the nasal was between 5% and 45% of the length of the word (see Fig. 11), since it is at around these points that the correlation breaks down. If the nasal is less than 5% of the word's duration, it is probably

no. of  
ym $\hat{l}$ add  
judgements

Fig. 11: Plot of ym $\hat{l}$ add judgements against nasal length  
by source word



too short to be heard satisfactorily. (In this case, it consisted of about six very short cycles at the point in question - i.e. a stretch of about 20 ms of nasal - in the stimulus of list 1). This would lead to confusion on the part of the subjects, and a more random pattern of stress judgements. If the nasal is more than 45% of the word's duration (the absolute duration varied between the words) the result is rather unnatural in effect and could lead to distorted or random judgements.

#### 2.3.4.2. Source word effect

The source word effect was found to be significant in all three lists under consideration, though it varied in strength between the lists. The table below gives the (product-moment) correlation coefficients between source word (value 1 for original ymladd, and 2 for original ymlâdd) and number of ymlâdd judgements, together with the significance of the correlation and the coefficient of determination.

Source word effect by list number

<u>list</u>	<u>correlation</u>	<u>significance</u>	<u>coeff. of determination</u>
1	0.7136	0.0001	51%
2	0.7608	0.0001	58%
3	0.6499	0.0001	42%
4	0.8480	0.0001	72%

It will be noted that the correlation is very similar for lists 1 and 3, but much higher in list 4. Here, it is even higher than that of list 2, where it was found to be the sole effect operative. This means that the phonological length of the /a/ is of much more importance when the word is subjected to the pitch pattern of list 4. However, the source word effect is still present in the other lists, though to only a minor extent in the first list.

The 'accuracy indices' of Fig. 10 were calculated by dividing the number of stimuli correctly marked as ymlâdd or tonau by the total number of ymlâdd or tonau judgements made, for each subject and each list. This index gives a measure of the subjects' ability to identify accurately the source word of the stimuli - a high value would indicate a strong source word effect. A value of 0.50 could be due to chance alone, as half the

stimuli originated from ymlâdd; thus only values consistently higher than 0.50 should be seen as significant. The table below summarises these indices, giving the mean accuracy index over all ten subjects for each list number.

<u>Source word effect based on accuracy indices</u>				
	<u>list 1</u>	<u>list 2</u>	<u>list 3</u>	<u>list 4</u>
mean index	0.74	0.69	0.68	0.81

This way of measuring the source word effect shows the same pattern as before, in that the value for list 4 is considerably higher than that for the other three lists, which are all similar. This table differs from the previous one in that the value for the first list is now higher than that for lists 2 and 3, whereas in the previous table the value for list 2 was higher than that for lists 1 and 3. However, both tables agree in assigning list 3 the lowest value for source word effect.

The second way of measuring the source word effect is perhaps more accurate than the first, since the first method is based on treating a category variable (source word) as an integer variable, in order to obtain a correlation coefficient. This is a slightly dubious practice, statistically speaking, and is probably best abandoned for these purposes in favour of the second method. Both methods agree on the main trends, however; namely, that the source word effect is strongest in list 4, moderate in list 1, and weakest in list 3.

#### 2.3.4.3. Independence of the two effects

Within each list, the nasal length effect (correlation of nasal length with ymlâdd judgements) was calculated for stimuli from each source word. In each case, it was not possible to show that the source word had any effect on the strength of the correlation. The conclusion is that the different source words did not give rise to different correlation coefficients. This means that the source word had no effect on the strength of the nasal length cue, and thus that the two effects were completely independent of each other.

Having established the independence of the two effects, a linear regression was carried out for lists 1, 3 and 4 together, the dependent variable being the number of ymlâdd judgements, and the independent variable being the length of the nasal as a percentage of the word length. This model had an f-ratio significant at  $p < 0.001$ , and thus its description could be relied upon. However, the coefficient of determination was low at 23%, and the coefficient of variation was high at 61%, indicating high variability and the presence of factors unaccounted for. This model was clearly not satisfactory.

A similar linear regression was then carried out for lists 1, 3 and 4 together, the dependent variable being the number of ymlâdd judgements, and the two independent variables being the nasal length and the source word (here treated of necessity as an integer variable, rather than as the category variable it actually was). This model had an f-ratio significant at  $p < 0.001$ , and thus its description could be relied upon. The coefficient of determination was high at 71%, and the coefficient of variation was lower at 38%, indicating lower variability and a greater number of factors accounted for. Thus the second model was far more satisfactory than the first, indicating, as expected, that both the nasal length effect and the source word effect need to be taken into consideration when predicting the number of ymlâdd judgements likely to be made.

The output of this model was plotted as Fig. 11, where the values of 'number of ymlâdd judgements' predicted by the model are superimposed on a plot of number of ymlâdd judgements against nasal length, by source word. It is clear from this plot that the predicted values are a reasonably good fit to the trend of the observed values. The source word effect is seen in the vertical separation of the two main blocs of observations; its lack of dominance is seen in the fact that some observations for each source word encroach on the territory of the other bloc. The nasal length effect is seen in the negative slope of the observations and predicted values; its lack of dominance is seen in the fact that there is a small degree of scatter around the two regression lines, some of this scatter probably being due to the source word effect.

The unaccounted-for variability in the improved model is probably at least partly due to the different  $F_0$  patterns of the three lists considered. However, it was not possible to break down the analysis still further by the addition of the variable 'list number', since the number of observations in each category would then have been unacceptably small. In addition, while a case could be made for treating 'source word' as an integer variable, there is no basis for doing this with the uncompromisingly category variable 'list number', and so this type of linear regression could not be performed in these conditions.

Despite the above limitations on the development of a model, a very rough notion of the contribution of 'list number' to the model may be obtained by considering the coefficients of determination for each of the two effects, for each of the three lists. These were shown in sections 2.3.4.1. and 2.3.4.2., and are summarised in the table below, where they are also summed for each list.

Contribution of the two effects to the model, by list number

coefficients of determination for:			
<u>list</u>	<u>nasal length</u>	<u>source word</u>	<u>sum of both coeffs.</u>
1	0.2963	0.5092	0.8055 i.e. 81%
3	0.3252	0.4158	0.7410 i.e. 74%
4	0.1397	0.7191	0.8588 i.e. 86%

It has been established that the two effects are wholly independent in their operation (see start of this section) and so their coefficients of determination do not overlap. This means that it is a valid procedure to sum their coefficients of determination and treat the result as the total variability accounted for by the model. The remainder constitutes the error, i.e. variability not accounted for. This error could be due to random fluctuation and/or to one or more undiscovered variables. However, where there is a high value for the summed coefficients, it is more likely that there is no third variable, and that the remaining variability is due solely to chance.

It is clear from the table above that there is most remaining variation in list 3, and least in list 4. This means that the two effects discovered

can account for nearly all the variation in ymlâdd judgements for list 4, and that this pitch pattern does not in itself have any marked effect on listeners' judgements. On the other hand, the summed coefficients for list 3 leave a sizeable amount of variation unaccounted for, and it is quite probable that this large error represents the operation of a third factor (as well as random variation, naturally). In this case, the third factor would be the pitch pattern of list 3, which shifts judgements more towards ymlâdd (see section 2.3.3.4.) and thus has a significant effect in itself. The position of list 1 is intermediate, and thus inconclusive on the evidence available here; while it is possible that the error here disguises a third factor (i.e. the monotone pattern), it is also possible that the error is due solely to random fluctuation. The effect of the pitch pattern of list 1 must be estimated from other evidence, as follows.

From sections 2.3.4.1. and 2.3.4.2. it can be seen that lists 1 and 3 tend to function together in opposition to list 4, in that the former two lists show a strong nasal length effect but weaker source word effect, and vice-versa for list 4. This on its own would serve to assign also to list 1 the above conclusions for list 3. Against this, however, should be set the finding, from Fig. 9, that lists 3 and 4 tend to show a similar high degree of agreement between subjects, while list 1 is alone in showing a low degree of consistency in judgements between subjects. The conclusion would appear to be that the pitch pattern of list 1 (i.e. a monotone) is like that of list 3 in that it has a marked effect on stress judgements in its own right. (Whether this effect is in the same direction as that of list 3 is a different matter). At the same time, however, the pitch pattern of list 1 gives rise to more uncertainty in judgements, and thus its effect cannot be said to be as strong as that of list 3. The nature of the effect of a monotone thus established will be discussed in section 2.3.4.4. below.

#### 2.3.4.4. Pitch pattern effects

Accuracy of judgements A new accuracy index was calculated for each list, based on the source word of those stimuli judged to be ymlâdd. The number of correct ymlâdd judgements was divided by the total number of ymlâdd judgements, for each list. A low figure would indicate difficulty

in picking out those stimuli originating from ymlâdd - in other words, a weak source word effect. The figures gave no information as to consistency of agreement between subjects, unlike the indices in Fig. 10. The new accuracy indices are shown in the table below.

Accuracy of listeners' "ymlâdd" judgements

<u>list</u>	<u>from original</u>	<u>from original</u>	<u>total</u>	<u>index of</u>
<u>no.</u>	<u>"ymlâdd"</u>	<u>"ymladd"</u>	<u>judgements</u>	<u>accuracy</u>
1	136	52	188	0.7234
3	154	75	229	0.6725
4	164	39	203	0.8079

The same procedure was repeated using the listeners' ymladd judgements (stress on penult). The number of correct ymladd judgements was divided by the total number of ymladd judgements for each list. A low figure would indicate difficulty in picking out those stimuli originating from ymladd, and thus these figures supply the other aspect of the source word effect. These accuracy indices are shown in the table below.

Accuracy of listeners' "ymladd" judgements

<u>list</u>	<u>from original</u>	<u>from original</u>	<u>total</u>	<u>index of</u>
<u>no.</u>	<u>"ymlâdd"</u>	<u>"ymladd"</u>	<u>judgements</u>	<u>accuracy</u>
1	64	148	212	0.6981
3	46	125	171	0.7310
4	36	161	197	0.8173

From a comparison of the above two tables it will be seen that the greatest accuracy (i.e. greatest source word effect) was seen in list 4, for both ymlâdd and ymladd judgements. The other two lists, however, are not consistent between the two tables. List 3 shows much greater accuracy in ymladd judgements than in ymlâdd judgements. This is due to the fact that list 3 received a far greater absolute number of ymlâdd judgements than ymladd judgements, as seen from a comparison of the 'total judgements' for list 3 in each table. The other two lists did not have such a great discrepancy between the two types of judgement. The result is that list 3 received a large number of ymlâdd judgements, both accurate and inaccurate, and this led to a low index for accuracy of ymlâdd judgements

but a moderately high index for accuracy of ymladd judgements. Thus the pitch pattern of list 3 tended to influence judgements towards ymlâdd, as noted in sections 2.3.3.4. and 2.3.4.2.

List 1, on the other hand, shows slightly more accuracy for ymlâdd judgements than for ymladd judgements, due to the fact that it received rather more ymladd judgements in absolute terms, both accurate and inaccurate. The discrepancy between types of judgement is not as great as for list 3, but is still markedly greater than that for list 4, which is so slight as to suggest a mere random fluctuation. Thus the pitch pattern of list 1 (a monotone) tended to influence judgements towards ymladd. In section 2.3.4.3. it was established that the pitch pattern of list 1 exerted an influence on judgements in its own right; it is now plain that this influence was in the direction of ymladd.

Summary of pitch effects The situation for the three lists under consideration is summarised below.

list 1      shifts judgements towards ymladd  
              much disagreement between listeners  
              nasal length effect moderate to strong  
              source word effect moderate

list 3      shifts judgements towards ymlâdd  
              much consistency between listeners  
              nasal length effect strong  
              source word effect weak

list 4      does not shift judgements  
              moderate consistency between listeners  
              nasal length effect weak  
              source word effect strong

These findings are relevant to the intonational system of Welsh. On balance, it could be said that the pitch pattern of list 4 constitutes an 'unmarked' intonational pattern, while those of the other two lists constitute 'marked' intonational patterns. This is because the Fo pattern in list 4 did not influence judgements in its own right, and allowed the

cue of segmental phonological length (of the /a/ vowel) to be taken into account more fully in judgements, due to the lack of 'interfering' cues.

On the other hand, list 3 consistently conditioned ymlâdd judgements. Since the form ymlâdd is an example of an irregular and therefore 'marked' stress pattern in Welsh, it can be argued that the listeners were less likely to be expecting this form and perhaps more reluctant to mark it. Thus they would pay more attention to subsidiary cues such as the nasal length effect (at its strongest in list 3) in order to be certain in their judgements. This, of course, is merely speculation; but there does not seem to be any other obvious explanation for the facts summarised above. Thus the pitch pattern of list 3 may be regarded as 'marked'.

The pitch pattern of list 1 (a monotone) conditioned a shift towards ymladd, though not as consistently as the opposite shift in list 3. The other features of list 1 are inconclusive, and thus no more can be said with certainty about this pitch pattern beyond the fact that it is 'marked'.

Discussion of pitch effects The list 4 pitch pattern, that seems to be 'unmarked', has the penult on a lower pitch than the start of the ultima, while the ultima soon falls to a lower pitch. This is, in fact, the pitch pattern most often seen in prenuclear position for Welsh, where the characteristic 'head' takes the form of a gentle downward slope of stressed syllables (both penults and stressed monosyllables) with unstressed syllables bobbing above this line in pitch (see Thomas, C.H. 1967 and Williams 1980 for a description of this 'saw-toothed' head). The pitch pattern of list 4 could also be interpreted as a 'nuclear tone'; with stress on the first syllable, it would constitute a 'rise-fall', and with stress on the ultima, a 'high fall' (see O'Connor and Arnold 1961 for an introduction to these terms). The fact that it could be interpreted as nuclear whichever syllable is heard as stressed may help to explain why this pitch pattern does not shift judgements either way; both the above-mentioned nuclear tones are found in Welsh (Thomas, C.H. 1967, Williams 1980). An important point to note is that under the 'rise-fall' interpretation, the unstressed ultima would carry the pitch glide; in other words, it would be pitch-prominent by the criteria generally used

(see O'Connor and Arnold 1961, Crystal 1969, for definitions of pitch-prominence). The implication is that pitch-prominence, as such, is no cue at all to stress in Welsh - and neither is lack of pitch-prominence, since under the high fall interpretation the stressed penult would also be pitch-prominent. Thus there is no consistent Fo cue to stress in terms of pitch-prominence of the syllable in question. Where intonation does influence stress judgements, this is not by virtue of a simple mapping of pitch-prominence with stress, but by virtue of the existence of particular categories of possible intonational structures, the influence of which can restrict the choices open to listeners in certain specialised conditions.

The pitch pattern of list 3 is not found in the 'head' portion of the tone-unit in Welsh, since the great majority of penults are not higher in pitch than their following syllable (Williams 1980). Thus it is not as 'unmarked' a pattern as that of list 4, being less common. When considered from the point of view of nuclear tones, moreover, it could only be seen as nuclear if the stress were assigned to the second syllable, in which case it would constitute a low fall. If the stress were seen as lying on the first syllable, the resultant pitch pattern would not be one commonly found in Welsh at all. Herein lies the reason, most probably, for the tendency of the list 3 pitch pattern to shift judgements so consistently towards ynlâdd; the listeners are influenced by what they take to be a nuclear pitch pattern, and apply the only possible structure, in nuclear tone terms, to what they hear. It should be noted, however, that even here the nasal length effect is not obliterated. On the contrary, it is very strong, demonstrating that, strong as the effects of intonation in Welsh may be, they are certainly not strong enough to override the post-stress consonant cue to stress in all cases.

The pitch pattern of list 1 (monotone) is found only rarely in the head portion of a tone-unit in Welsh, and thus is more 'marked' and uncommon a pattern than that of list 4. Considered from the point of view of head contours, each of the two syllables is equally likely to be the stressed one (unlike the situation in list 3) and so there is no positive intonational cue to stress placement. There is also no nuclear tone cue, since there are no level nuclei in Welsh, and so the listeners are not given much help by this pattern; a fact reflected by the high degree of

inconsistency between listeners. Since the form ymladd is much the most common of the minimal pair in its stress pattern, it could be that this form is preferred by default, leading to a preponderance of ymladd judgements for want of any cue at all to the contrary. The nasal length effect is moderate to strong and the /a/ length cue is moderate, indicating that listeners are being forced to make use of all possible cues. The pitch pattern of list 1 is the least helpful of the three in intonational terms, since a monotone is both the rarest pattern found of the three and also, where it is in fact found, gives no intonational cue at all as to the stress pattern. In the other two lists, the pitch patterns are more common (particularly for that of list 4) and/or give cues to stress pattern from an intonational point of view (in the case of list 3). The specialised, uncommon, and unhelpful nature of the pitch pattern of list 1 is reflected in the low consistency of judgements here, the need to rely more on the two minor cues, and the need to have recourse to the default form ymladd for want of any strong cues. It should also be noted that, where Fo influences stress judgements, this is not in terms of pitch-prominence (which is the same for all languages and forms the raw material for intonation) but in terms of intonational categories (which are specific to a particular language and not subject to modification) - this may well be the reason why pitch-prominence per se was found in the earlier sets of measurements not to be a reliable cue to stress in Welsh.

## 2.4 Experimental findings

### **2.4.1 Summary of experimental results**

The first set of measurements (section 2.1.) was by nature a preliminary experiment, serving to delineate the problem rather than to offer a solution. It was found that the stressed syllable in Welsh was quite unlike the stressed syllable in English from the point of view of acoustic cues, and that the most consistent cue to stress in Welsh, in the case of regularly-stressed disyllables, was the shorter duration of the stressed vowel. Since this set of measurements concerned only regularly-stressed disyllables, this conclusion should be taken to mean only that stressed

penults are consistently shorter than unstressed ultimas in the case of words spoken in isolation. (Stressed monosyllables, on the other hand, contain indisputably long vowels phonetically that often contain a pitch glide, as is the case for unstressed ultimas). The most intriguing finding of this set of measurements is that the usual reliable cues to stress in English and other languages (high  $F_0$ , high amplitude) are not at all reliable in the case of Welsh. Even in the absence of an identifiable positive cue to stress in Welsh, this finding would be an interesting one and a spur to further investigation.

The second set of measurements (section 2.2.1.) checked the findings of the first set in the case of unrehearsed, connected speech. The same general trend was indeed found; namely, that the stressed syllable in Welsh has shorter duration, but cannot with certainty be identified with any other acoustic cue. The fact that the same trend was found in this type of speech increases the probability that the trend was a real aspect of spoken Welsh and not merely an experimental artefact; thus more confidence can be placed in these findings.

The third set of measurements (section 2.2.2.) attempted to discover what it was that cued stress in Welsh, since inherent acoustic cues had been shown to be of little help. It was found that rhythm is the foundation of stress in Welsh, and that the stressed penult retains its status of 'stressed' not in virtue of any acoustic prominence it may have in its own right, but because of its function as the keystone of the rhythmic unit. This finding is perhaps the most theoretically significant of all the experimental results described here (see section 2.4.2. for a discussion of its implications).

The fourth set of measurements (section 2.2.3.) was designed to check the findings of the previous measurements, to test for any variation in rhythmicity between two types of list, and to investigate the effect of syllabification. The results supported the earlier finding of a tendency to isochrony in Welsh, thus increasing confidence in this finding. The two types of list could not conclusively be shown to differ in rhythmicity, owing to the high degree of variability in the statistics for each list. It was found that there was a greater tendency towards

isochrony if the rhythmic feet were counted from the onset of the stressed vowel rather than including any initial consonants of the stressed syllable. This finding indicates a possible need to recognise two types of syllabification for different levels of analysis in Welsh; the more 'abstract' and 'morphological' syllable division would include initial consonants within the stressed syllable and would be used for purposes of morphological processes, while the more 'concrete' and 'phonetic' system would count the stressed syllable only from the onset of the vowel, and would be used for purposes of rhythm and stress.

The fifth set of measurements (section 2.2.4.1.) reverted to measurements of segments - in this case, of consonants. It was found that, in regularly-stressed disyllables with medial voiceless stop, the length of the post-stress stop was greater than the length of the consonant after the unstressed ultima. This discovery of greater length in the post-stress consonant agrees with observations made auditorily in the past.

The sixth set of measurements (section 2.2.4.2.) again concentrated on segments - in this case, vowels. The prepausal lengthening seen in other languages was also demonstrated here for Welsh. Less predictably, it was also found that the stressed vowel was significantly shorter than the unstressed vowel, when phonologically short vowels only were considered. This bears out one of the findings of the first and second sets of measurements. The segmental measurements thus lead to the conclusion that stress in Welsh is identified with shorter duration of the stressed vowel and longer duration of its following consonant. The finding is all the more interesting for not being the obvious one, and reinforces the belief that stress in Welsh is based on factors of timing rather than on pitch or intensity. What is not revealed is the reason why this relationship between vowel and consonant should be linked with phonological stress.

The next experiments (section 2.3.) were of a perceptual nature, seeking to test for speech perception the findings that had been obtained for speech production in the earlier sets of measurements. The results of these perceptual experiments can be summarised as follows:

- a) The duration of the following consonant is a cue to stress in speech perception, as it had been found to be for speech production.
- b) The length of the vowel of the ultima can also influence perception. This means that the word-pair ymlâdd and ymladd is not a true minimal pair, since the words are differentiated not only by stress placement but also by the length of the /a/. However, they were the nearest pair obtainable to a stress-related minimal pair in Welsh, and thus this 'source word' effect could not be avoided.
- c) Intonation can influence stress judgements. This influence, however, is not a direct mapping of pitch-prominence with stress, but is a more indirect matter of restriction of choice, under certain conditions, in accordance with the recognised intonational categories of Welsh, i.e. the possible nuclear tones and head patterns. However, intonational considerations never completely override the post-stress consonant cue to stress.
- d) Where the intonational pattern is unmarked, the stage is left clear for the operation of cues of segmental length; in particular, the vowel length cue in the case of this word-pair.
- e) Where the intonational pattern is marked and biassed towards a particular interpretation according to the recognised intonational categories of Welsh, then this interpretation will influence stress judgements but will by no means outweigh the post-stress consonant cue to stress.
- f) Where the intonational pattern is marked but not biassed to any interpretation according to intonational structure, then stress judgements are made (albeit less consistently) on the basis of segmental cues of duration (of the post-stress consonant, and of /a/ length in this pair), and also on the basis of the default or unmarked member of the word-pair.

The general conclusion to be drawn from the perceptual experiments is that the perception of stress in Welsh is a complex composite of interdependent factors, and that stress and intonation, though different

in their basis, nevertheless interact in ways which the above experiments have only begun to investigate.

## **2.4.2 Implications of the results**

### **2.4.2.1. Stress as separate from intonation**

The above experimental findings contain two types of conclusion that are of relevance to the theory of stress in general; first, that stress and intonation are separate phenomena, and secondly, that stress is defined in terms of rhythm.

The first conclusion accords with both the 'British' and the 'American' views of stress, which define it separately from intonation, though naturally conceding that stress and intonation can interact to a large extent. This view is at odds with the view of stress as 'accent' (see section 1.2.4.) propounded by Bolinger, who defined stress (his 'accent') solely in terms of pitch variation, relegating intonation to the area of degrees of pitch variation. Such a view treats phenomena traditionally regarded as part of intonation as being instead part of stress (or 'accent'). It is therefore open to the charge of confusing stress and intonation, although, as has been mentioned, it makes allowance for intonational phenomena - it merely draws the line between stress and intonation in a very different place from other workers and from the present investigation.

The experimental results presented here give concrete evidence, for Welsh, that stress must be defined in entirely different terms from intonation (conceived of as pitch variation, or degree of pitch variation, of any kind). The significance of these results is not that they reiterate what better scholars have said many times in the past on theoretical grounds - namely, that stress and intonation are separate - but that they provide concrete phonetic evidence for this assertion - namely, the lack of inherent acoustic cues to stress, and particularly the remarkable unreliability of pitch-prominence as a cue to stress.

#### 2.4.2.2. Stress as defined in terms of rhythm

The second conclusion to be drawn from these experiments is that stress in Welsh is defined in terms of rhythm. Having established that stress is to be defined independently of intonation, it remains to find this independent basis of stress. For the traditional 'British' and 'American' schools, the independent basis was seen as increased intensity or loudness on the stressed syllable (section 1.2.2.), or else was seen as increase in  $F_0$  accompanied by increased articulatory effort and thus increased loudness (section 1.2.3.1.). Even the acoustic experiments into inherent cues to English stress (section 1.2.3.2.) concluded that the stressed syllable, though pitch-prominent, was also fairly consistently accompanied by increased intensity which could, if need be, override a contradictory pitch cue (Lieberman 1960). The integral of the amplitude with respect to time was found to be practically as reliable a stress cue as pitch-prominence (Lieberman, *op. cit.*). Thus the consensus view was that stress in English is based on intensity or loudness, though intonational effects can interfere with this pattern.

Such a view of stress is not valid for Welsh on the evidence presented here, as intensity was found to be at best a most unreliable cue, while pitch patterns were no help at all (from a direct-mapping point of view). The basis of stress in Welsh was found to be rhythm, and reliably so; while the only other consistent factor was also durational in nature (i.e. the length of the post-stress consonant), and could conceivably thus play a part in the rhythmic patterning.

#### 2.4.2.3. Linking the theoretical with the physical

These two conclusions, if valid, could lead to a fundamental reappraisal of the nature and functioning of stress. The separation of stress and intonation would exclude the 'stress as accent' analyses, while the definition of stress in terms of rhythm would exclude all other past analyses. The only theory so far formulated that is not excluded by these conclusions is the new 'metrical' theory of stress (see section 1.2.6. for a brief introduction to this theory). This theory defines stress in terms of the 'metrical grid' (loosely related to the 'prosodic structure'), while intonation is characterised independently by linking certain recognised

functional 'tunes' to the 'prosodic structure'. The metrical theory also sees stress as based on rhythm, which is defined on the metrical grid and functions alongside, but separately from, the categories imposed by intonation.

The metrical theory has, to date, been elaborated and justified largely on grounds of theoretical elegance and abstract phonological considerations, such as Liberman's criticism of Chomsky and Halle's 'stress cycle' (see section 1.2.6.1.). It is here that the real significance of the present experimental findings is seen; namely, in providing a concrete, phonetic basis for categories and definitions so far advocated in abstract terms only. The linking of theoretical and experimental research is a notoriously difficult process, since the assumptions and methodology of the one aspect so often run counter to those of the other; to such an extent that there may sometimes seem to be no basis for comparison at all. In the niceties of theoretical elaboration this may well be true for the metrical theory also - the present study has nowhere near enough detail to test this - but in the fundamental definitions and assumptions of the theory a clear link can be discerned between abstract system-building and concrete experimentation. It is true that this link (so far, at least) has been established only in the case of Welsh; however, it is at least possible that the same conclusions will prove to be valid for English also.

#### 2.4.2.4. Extending the physical to the theoretical

Having established a firm empirical basis for the separation of stress and intonation, and for the definition of stress in terms of rhythm (in the case of the Welsh language, at least), it now remains to investigate the theoretical implications of these two fundamental findings. This will involve a discussion and modification of the metrical theory of stress. The presentation will be, first, in terms of English, in order to make use of a language that is the most familiar to workers in this field, and also in order to show that the revised theory is at least descriptively adequate for English. The theory will be extended to describe certain problems in English intonation, in order to demonstrate its superiority in this field. The presentation will then be applied to Welsh - for which, it

will be remembered, there is now an established link between the metrical theory and the concrete facts of stress - and will then be extended to characterise the patterns of Welsh intonation, in an attempt to capture in theoretical terms something of the interaction of stress and intonation that was discovered on the physical level.

The theoretical presentation will make no direct reference to the physical facts described earlier, since the link between the two aspects is at the fundamental level of assumptions and definitions, and as such does not itself enter into the details of the theory. Such a disregard of the physical facts, however, in no way implies that they are of no importance; on the contrary, without this demonstrated link the theoretical presentation that follows would be merely an alternative and more elegant theory and not, as is the case, a simpler and also more explanatorily adequate theory, firmly rooted in the physical facts of language use and branching out into the wider realms of meaning and communication.

### 3. THEORETICAL CONSIDERATIONS

#### 3.1. Syllable level

At the start of the derivation, orthographic words and location of syllable nuclei are 'given'. The first stage is to mark each vowel/syllable nucleus as [+full] or [-full]. Liberman marks each vowel for [ $\pm$ stress], where [+stress] vowels are tense vowels or vowels before 'strong clusters'. He sees this as a segmental feature, interacting with the metrical grid as follows:

1)initially, there is an assignment of the segmental feature [ $\pm$ stress].

2)second, there is a metrical bracketing of the word.

3)third, there is an assignment of metrical node labels, on various principles.

4)fourth, there is an adjustment of the distribution of the segmental feature [ $\pm$ stress]. (L:139)

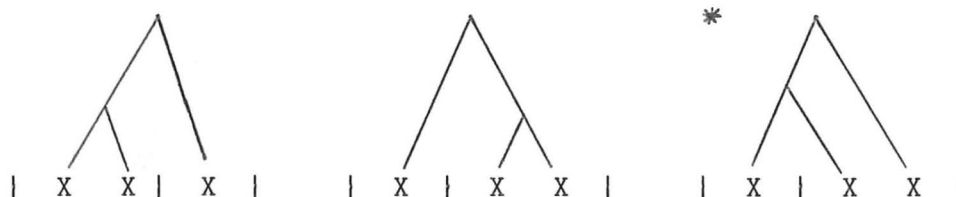
It is this feature which makes the distinction in the following, which have identical metrical structure but different syllable types.



The initial assignment of [ $\pm$ stress] is done on the basis of vowel tenseness, or the strength of the following consonant cluster, respecting word boundaries. The term 'stress' is here made to apply to a purely segmental feature conditioned by segmental features of the vowel itself or of following consonants. At this stage of the derivation it has no relation to syntagmatic aspects of the utterance, such as rhythm. This is all the more surprising in that stress comes to be seen later as dependent on the syntagmatic feature of metrical structure. In their discussion on the 'metrical grid', which is determined by the metrical structure, LP

write '...the "degree of stress" of a given syllable is represented by the height of the column of marks that stands over it.' (LP:312) 'Stress' is being used in two different ways here, unless the derivation is entirely circular (the syllable 'nucleus' feature [ $\pm$ stress] determining metrical bracketing, which determines the 'metrical grid', which determines 'stress' on each syllable). Clearly, it is better to avoid using 'stress' for a purely segmental feature. This is what has been done under the proposed analysis, where the feature [ $\pm$ full] applies only to vowel segments and avoids any suggestion of circularity.

In Liberman's theory, the metrical bracketing of the word follows. There are certain conditions on bracketing, stated in the form of rules inserting a 'foot boundary', symbolised as '|', into the string of segments. These 'foot boundaries' divide the string into groups which the metrical bracketing cannot violate. The first two examples below are permissible, but in the third the metrical bracketing violates the foot boundaries, and so is not permissible.



There are two 'foot boundary insertion rules', or conditions on well-formed bracketings:

- 'a) All # boundaries are foot boundaries
- b) A foot boundary is inserted in the environment  $\_ / [ + ] [ - ]'$  (L:141)

The next process is one of node-labelling:

'In any lexical metrical constituent [MN], N is strong if and only if it is complex.' (L:142) 'Complex' means 'dominating non-terminal material'; this is the most general form of the rule.

The initial values of [ $\pm$ stress] are then adjusted where appropriate, the most important redistribution rule being as follows:

S                      S  
 |              ----> |  
 -                      +

The reason for this is that no 's' may dominate a [-stress] vowel. Other stress redistribution rules are:

Case 1 (for nouns) [-stress] [+stress]/\_[-stress][-stress]#

Case 2 (for verbs) [-stress] [+stress]/\_[-stress]#

Naturally, there are exceptions to all these rules, but this is the basic process in Liberman's original work. An example follows:

1) clas si fi ca tion              initial [ $\pm$ stress] assignment  
      -    -    -    +    -

2) | clas si fi | ca tion |              foot boundary insertion  
      -    -    -    +    -

3) | clas si fi | ca tion |              bracketing  
      -    -    -    +    -

4) | clas si fi | ca tion |              labelling  
      -    -    -    +    -

5) | + |              [ $\pm$ stress] adjustment:    s                      s  
      -                      |              ----> |  
      -                      -                      +

Liberman's use of [ $\pm$ stress] is therefore intimately connected with word stress, which is what his 'metrical bracketing' is also attempting to

account for: there seems to be unnecessary duplication here. Since the original idea behind metrical bracketing was to remove the need for an SPE-type [ $\pm$ stress] feature assigned cyclically, this seems self-defeating.

So instead of Liberman's [ $\pm$ stress], a different vowel feature will be proposed: [ $\pm$ full]. This distinguishes between [ə] and [ɜ] on the one hand ( $[-full]$ ), and all other vowels on the other hand ( $[+full]$ ). What this feature is intended to express is the fact that, in English, [ə] and [ɜ] are the only vowels which cannot be stressed. All other vowels – even short ones – can conceivably be stressed; only  $[-full]$  vowels never are. If a syllable with a  $[-full]$  vowel is stressed (e.g. for purposes of contrast), then the vowel quality is changed to one that is  $[+full]$ , as in the following:

/rəli:v/	--->	/rili:v/		/ʌnde/	--->	/ʌndɜ:/
- +		+ +		+ -		+ +

Mostly, the new  $[+full]$  vowel fits in with the orthography. In the absence of any other clues (e.g. morphological alternations), it would seem that speakers are guided by spelling.

The value  $[-full]$  applies to vowels which are stressless even in citation form, needing especial emphasis and a change in quality and/or quantity before they can receive sentence stress. Syllabic consonants would also be  $[-full]$ , being inherently stressless. The definite article, normally 'ðə/' and therefore  $[-full]$ , becomes 'ði:/' or 'ðɪ/' before a word beginning with a vowel, and is then  $[+full]$ . This is also the form it takes when given emphasis.

The feature [ $\pm$ full], unlike Liberman's [ $\pm$ stress], is concerned not with sentence stress but purely with the prerequisite for stress – i.e., a non-reduced vowel. An example follows:

+	+	+	+	-	+	-	+	+	-	+	-
John gave the information to Susan's uncle											

### 3.2. WORD LEVEL

At this stage of the derivation, syllables marked for [ $\pm$ full] are 'given'. The next step is to form the metrical structure of each word.

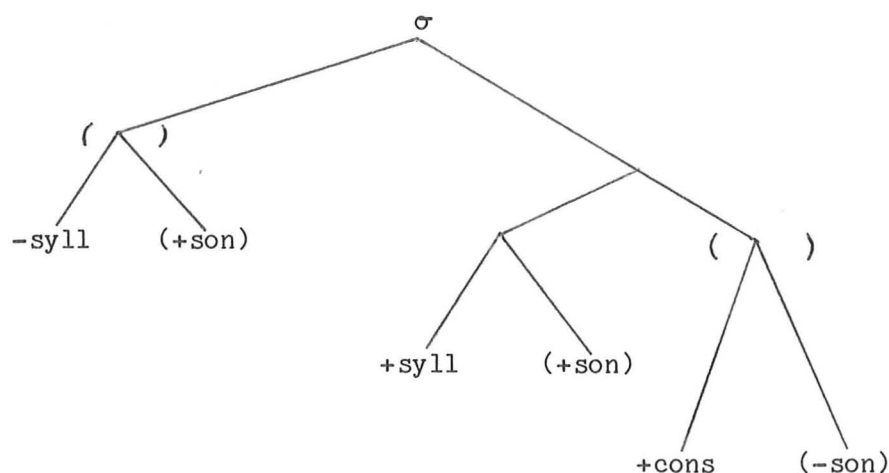
L and LP contain no formal distinction between word-level and phrasal level, and yet it is often necessary to make this distinction. L notes three differences between word stress and phrasal stress, pointing out that word stress assignment is subject to some arbitrary exceptions, does not always result in a structure suitable as input to metrical annotation, and is not entirely relational. (L:138)

Two other differences may be noted: the stress pattern of a phrase or sentence has a fair degree of optionality, while word stress is usually unalterable; also, if word stress should be changed, what is emphasised is the newly-stressed syllable, not, as in the case of altered phrasal stress, the object in the real world to which the newly-stressed unit refers.

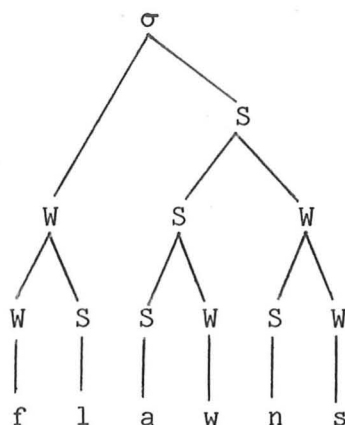
So there is a need to recognise a distinct word level, with its own units and rules, but which nevertheless follows the general principles of metrical bracketing. Such a system has been proposed by Elizabeth Selkirk (see Selkirk 1980 a and b). She sets up three units below the phrasal level: the syllable, 'stress foot', and prosodic word.

#### 3.2.1 Syllable, or 'σ'

Most workers have accepted the existence of the syllable as a useful and indeed a necessary unit for suprasegmental description. Selkirk defines the syllable, for English, by means of the 'English syllable template', where brackets indicate optionality (see Selkirk 1980b:8).

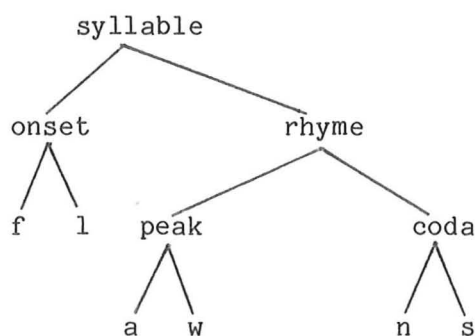


Her 'Maximal Syllable Onset Principle' formalises the finding that: 'In general, when a medial consonant or consonant cluster may be analyzed as either a coda [i.e. syllable-final] or an onset [i.e. syllable-initial] according to syllabification principles, [i.e. according to initial and final consonants and clusters permitted in the language] it is the onset analysis which prevails.' (Selkirk 1980b:9) Selkirk draws a parallel between the relation strong/weak and the relation more/less sonorous, with respect to segments. Vowels are more sonorant than liquids, which in turn are more sonorant than fricatives or stops. Thus she represents 'flounce' as follows:



She writes: 'The peak is of course strong (s), i.e. more sonorous, than the onset. And within each of the other constituents, the s has been assigned to the more sonorous element.' (op. cit.:5) She also quotes Pike's suggestion that w/s assignment depends on 'a universally defined sonority hierarchy' (op. cit.:6) Selkirk's representation is merely a reformulation,

in metrical terms, of the pattern developed by earlier scholars. She represents the earlier pattern as follows:

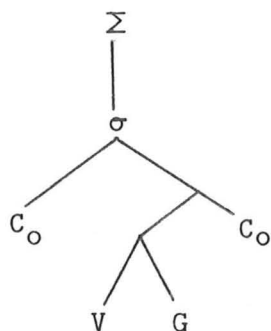


Selkirk notes that, while the prosodic category 'syllable' is required in linguistic theory, there is no support for naming its internal constituents in the above way, and so she rejects them in favour of what she refers to as the 'IC analysis of the syllable', citing past work which can be seen to support this analysis. This analysis also makes possible a unified treatment of the heavy/light syllable distinction, which plays an important part in stress systems. 'The light syllable CV can be characterized as one whose rhyme is simple, non-branching, while the heavy syllable, be it CVC or CVV (a long vowel) is one with a complex, or branching, rhyme.' (Selkirk 1980b:6)

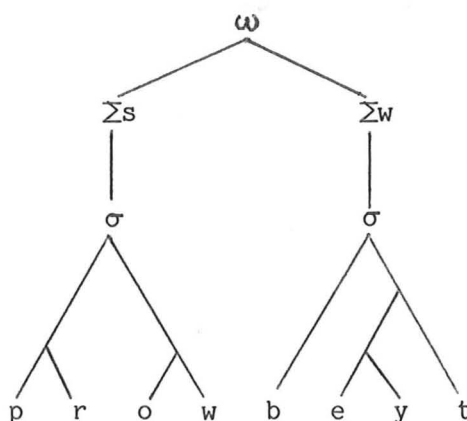
### 3.2.2 Stress foot, or 'Σ'

The 'next' type of prosodic unit in the hierarchy is the (stress) foot, which Selkirk outlines for English by means of a set of templates for 'basic feet', grouped into 'prosodic words', which are symbolised by 'ω' (see Selkirk 1980b:12).

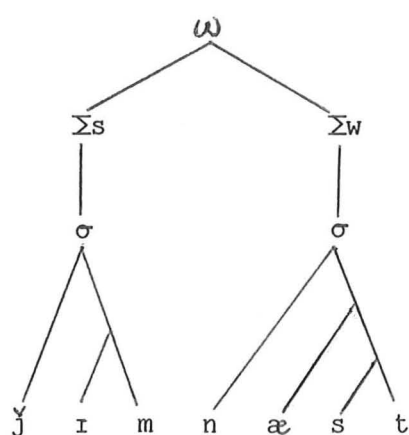
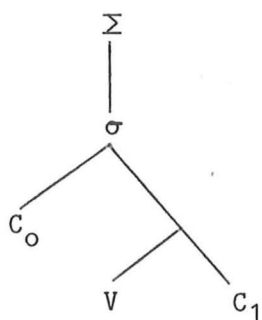
i)



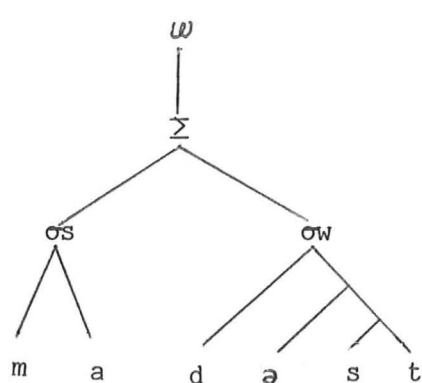
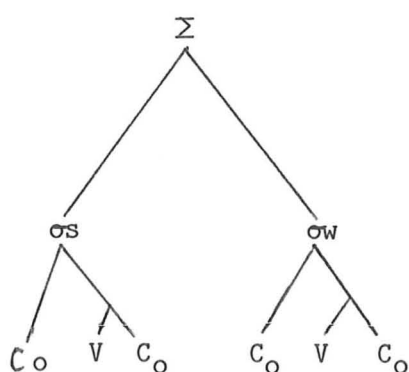
(VG = tense vowel or diphthong)



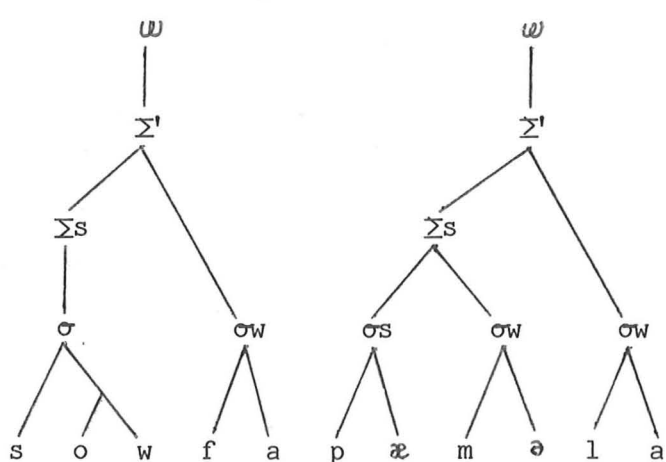
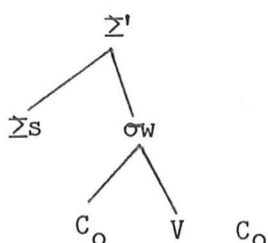
ii)



iii)

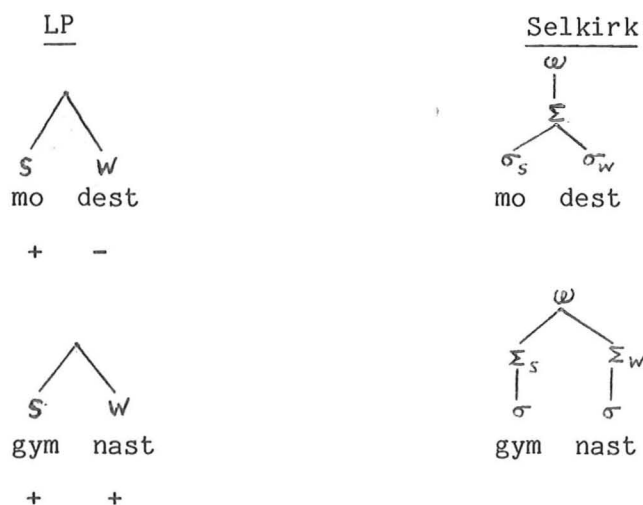


iv) the 'Super Foot'



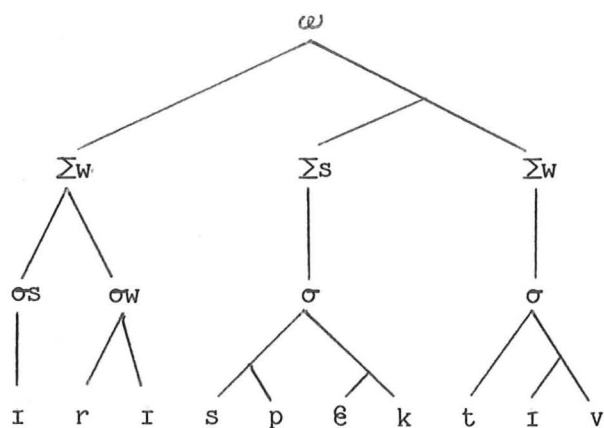
Selkirk sees these templates as functioning as well-formedness conditions on underlying representations within the syntactic domain of the simple (i.e. non-branching) stem, in English. If a syllable is itself a foot, it will receive some degree of stress. 'The claim is thus that being a foot will give an individual syllable a status in prosodic structure that is distinct from that of a syllable which is not a foot,

that a syllable will be phonetically 'interpreted' differently depending on its footness.' (Selkirk 1980b:13) It is in terms of the foot that she explains the distinction between the stress patterns of 'modest' and 'gymnast' that is noted by LP:



Selkirk's analysis has no need of the feature [ $\pm$ stress], and is therefore to be preferred, since the feature was (re-)introduced merely to deal with the fact that these two words, though stressed differently, had identical prosodic structure. In Selkirk's analysis, each has a different prosodic structure, reflecting directly the difference in stressing. However, this does not militate against the feature [ $\pm$ full], since a ' $\sigma_w$ ' (metrically weak syllable which is not itself a foot) may contain either a [+full] or a [-full] vowel, and so there is still a need for this distinction between inherently stressless and circumstantially stressless syllables. Representations of metrical structure in this study will include a Selkirk-style analysis at the word level, but will also have [ $\pm$ full] at the syllable level. The only condition is that a [-full] vowel must be part of a ' $\sigma_w$ ' - i.e. a weak syllable that is not itself a foot.

A point to note about Selkirk's analysis into 'stress feet' is that it does not necessarily follow morphological structure; c.f. her analysis of 'irrespective' (op. cit.:14):



The structure is determined by the following principles:

'The Prosodic Word: Constituency

The  $\Sigma$  are joined in a right branching structure.

The Prosodic Word: Prominence

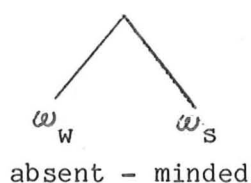
Given a pair of sister nodes  $[N_1 N_2]$ ,  $N_2$  is s iff it branches. (This is the Liberman and Prince Lexical Category Prominence Rule)' (op. cit.:14)

Unlike the syllable, the 'stress foot' was first described formally by Selkirk, who does not enter into an extensive justification of it. However, it seems a useful concept, from the examples given, and there seems to be a need for a unit intermediate between the syllable and the word.

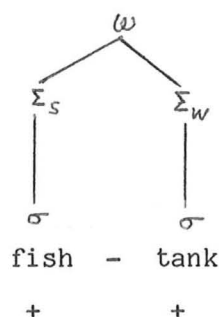
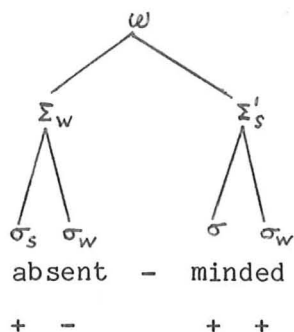
### 3.2.3 Prosodic word, or ' $\omega$ '

Selkirk has little description of this unit, merely noting briefly the findings of other workers and stating that there seems to be 'a fair amount of evidence that the  $\omega$  is a privileged domain of phonological processes.' -(op. cit.:14). The prosodic word is formed out of stress feet according to the principles given above, and the syntactic domain of well-formedness is identical for the syllable, foot, and word in English, being the simple stem. In other languages, the syntactic domain may be something other than the simple stem, but Selkirk postulates as a universal constraint that it is the same for syllable, foot, and word.

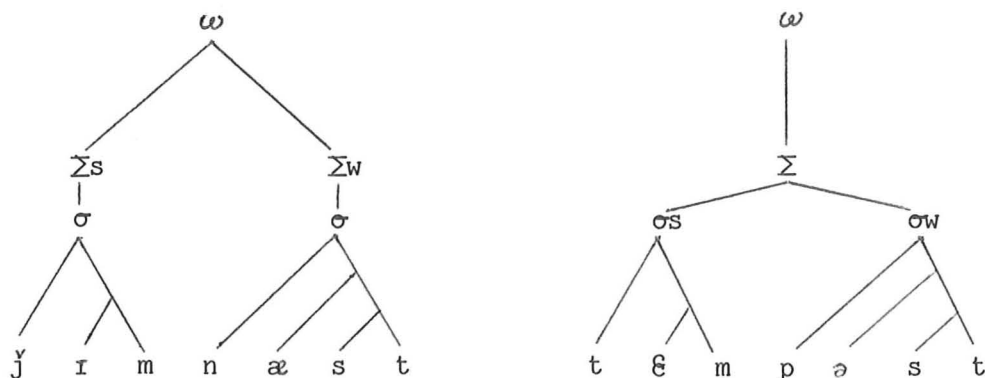
A point worth noting is that the lack of any clear definition of a prosodic word can lead to disagreement. For example, Selkirk analyses 'absent-minded' as being made up of two prosodic words, where the higher node in this example does not constitute a named unit (Selkirk 1980b:16).



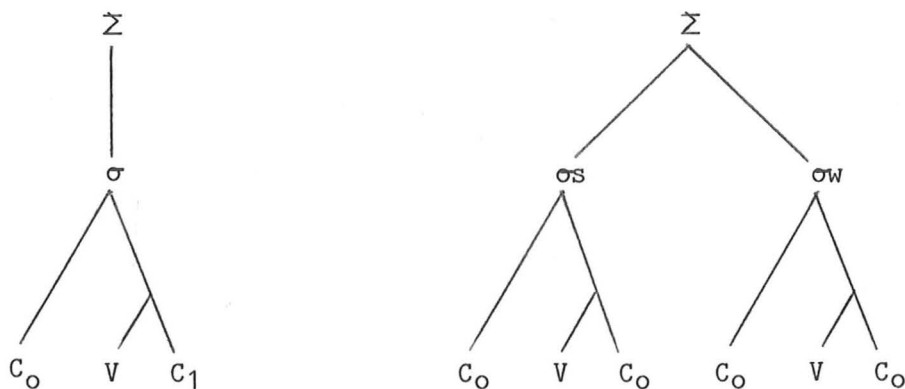
However, I would prefer to analyse this as being just one prosodic word, along with other compound adjectives and also compound nouns, thus avoiding the use of nameless units and corresponding more closely to the semantic facts:



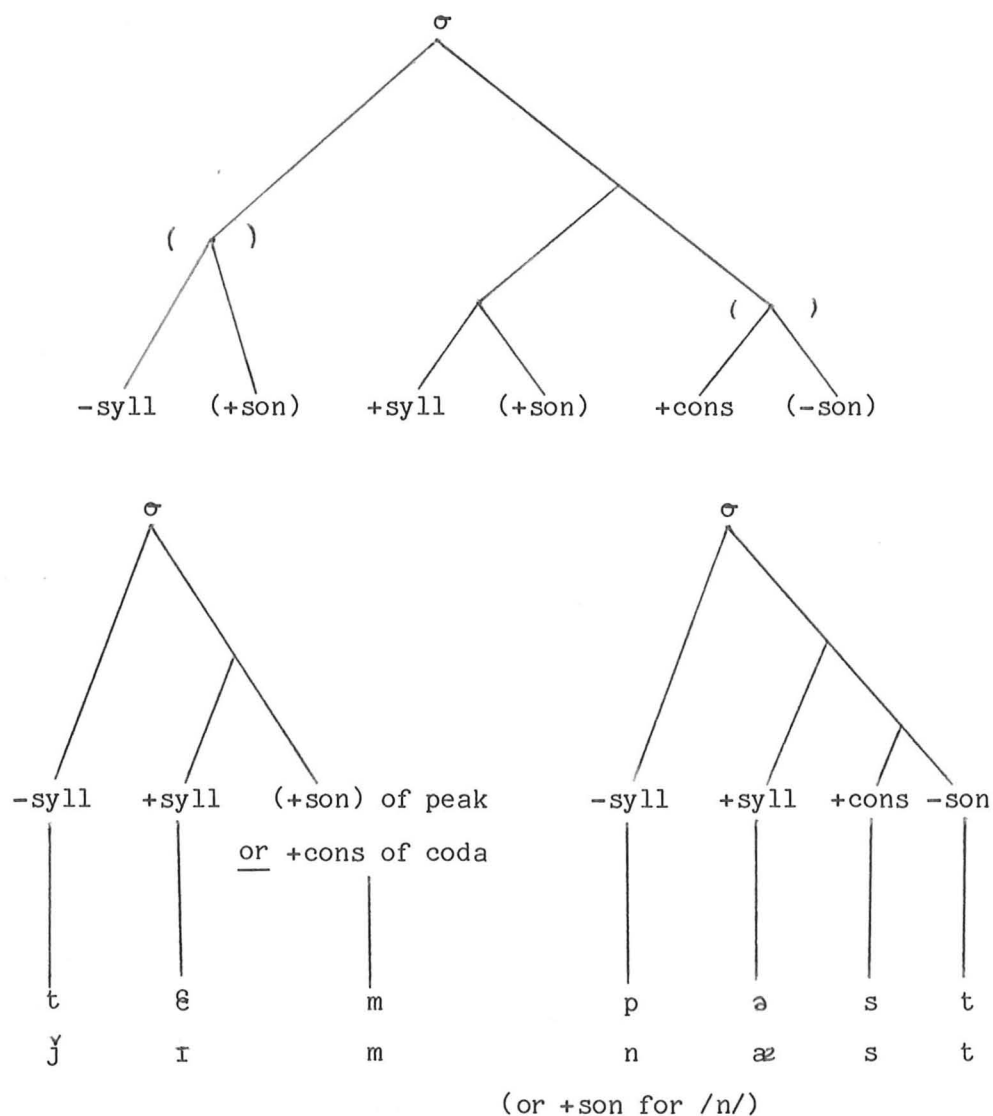
Another point worthy of mention in connexion with the templates for 'basic feet' is Selkirk's analysis of the words 'gymnast' and 'tempest', as follows (op. cit.:12):



Selkirk analyses 'gymnast' as two stress feet, to distinguish it from 'modest', which, while having the same s-w structure, is made up of only one (bisyllabic) foot, of the same pattern as 'tempest' above. The basic templates in question are as follows, where the first is a monosyllabic foot containing a non-tense vowel, and the second is a bisyllabic foot with non-tense vowels:



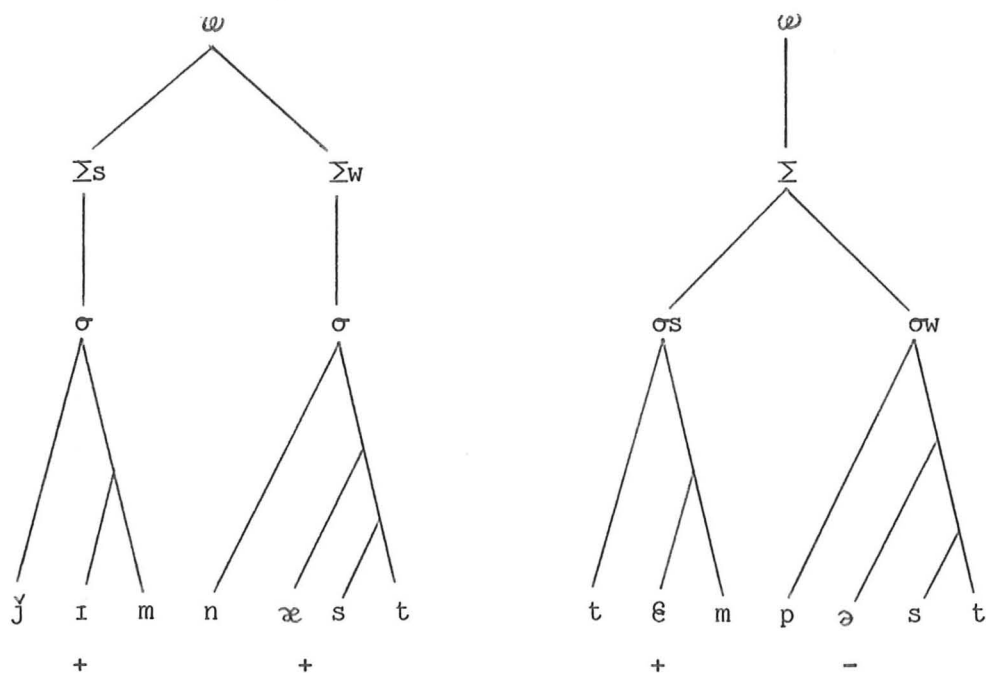
It will be noted that, according to these templates, 'gymnast' could just as well be analysed as an example of the second pattern above. The same applies to 'tempest'. Indeed, the two words are very similar in segmental composition, and have practically the same relationship to Selkirk's 'English Syllable Template' below (op.cit.:8):



Clearly, it is not the hierarchical structure of the syllables that determines their different stress patterning into feet, since this structure is identical. The only difference which seems relevant is that, in the second syllable, 'tempest' has /ə/ (in this dialect) while 'gymnast' has /æ/. There is no formal distinction made between different non-tense vowels in this analysis, which leads to the above identical patterns in terms of syllable structure. However, the two words are different in terms of foot structure, which Selkirk claims is based solely on hierarchical syllable structure. The only solution appears to be to allow the feature [ $\pm$ full] into the description, with the constraint that a syllable with a [-full] vowel can never be a foot in its own right; in fact, this means that a [-full] vowel will always be analysed as 'ow', i.e., a weak syllable in a bisyllabic foot or in a superfoot, because of the

prescribed system of bracketing into feet. This condition is also relevant to Selkirk's 'Monosyllable Rule' (see start of 3.4.5.).

The structure of the two words would then be as follows:



Thus it follows that [-full] vowels are seen only in 'ow' units, while [+full] vowels are also seen in 'os' or 'o $\Sigma$ ' units. This is so in the other examples given by Selkirk with the 'basic foot' templates. The various positions taken in this matter can be stated simply:

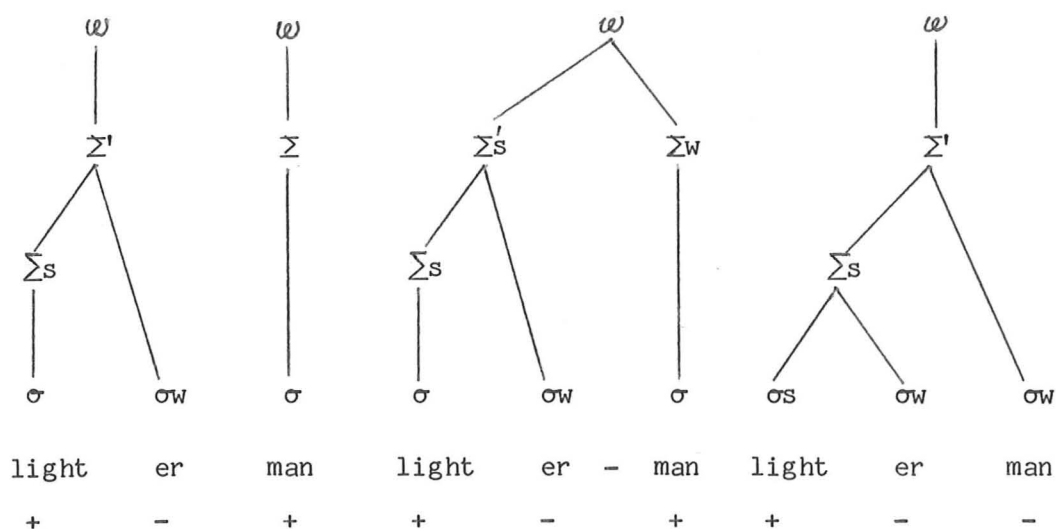
LP regard 'modest' and 'gymnast' as having the same prosodic structure but different segmental structure, as expressed by their use of [ $\pm$ stress] to mark vowels.

Selkirk regards the words as having different prosodic structure but the same segmental structure (insofar as she is interested in the nature of segments at all).

I regard them as having both different prosodic structure and different segmental structure (though there is no need to posit any sort of causality here). This view is necessary for a correct account of the similarities and differences between words of the type 'modest', 'gymnast', and 'tempest'.

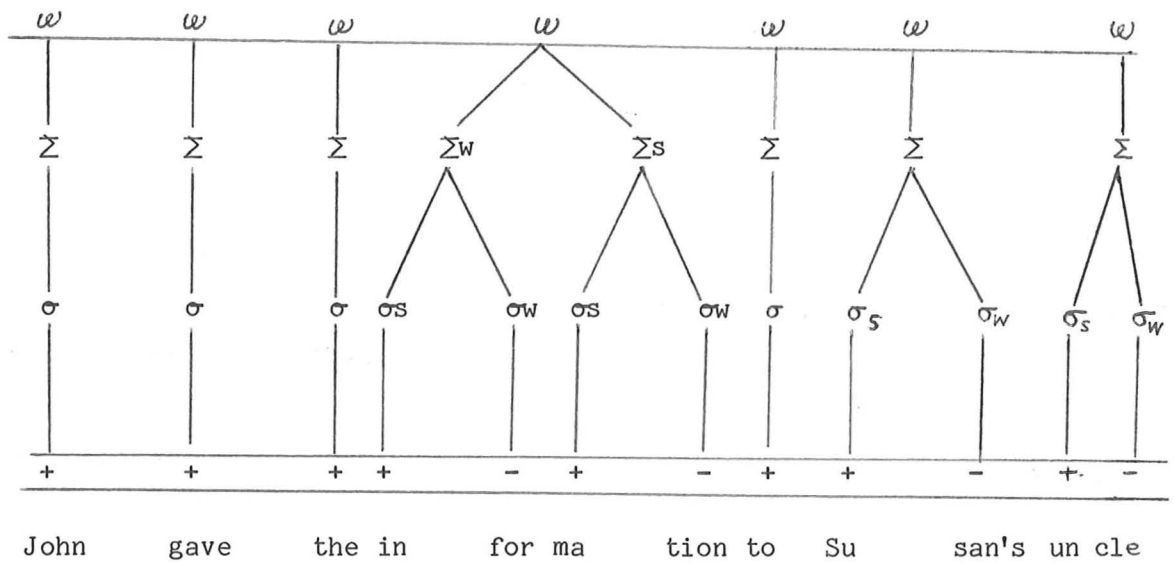
To return to our consideration of the 'prosodic word' as a unit: the recognition of a separate word level, distinct from the phrasal level and with its own units and well-formedness conditions, seems to be necessary. The definition of what constitutes a 'word' has always been fraught with problems, and this analysis likewise has scope for disagreement. But the use of 'stress feet' in well-defined patterns seems to offer at least an approximation to a solution, in that they allow for a level intermediate between word and syllable which could be the location of units which, for want of such a level, have been variously assigned to the syllable or the word level.

The stress foot also makes clear the progress from two lexical words to loose compound to close compound, as in the following:



In the first example, the units are two separate prosodic words. In the second example, the units are separate 'stress feet' within the same prosodic word - this could be termed a 'loose compound'. In the third example, the units are subsumed into the same stress foot and therefore of necessity are in the same prosodic word - this could be termed a 'close compound'. (In the last example, the vowel of 'man' is a syllabic /n/, and so the stress foot conforms to the template.) This process is one of reduction, both semantically and phonologically, and has two stages. Use of the 'stress foot' enables these stages to be characterised in a clear way.

The diagram below gives an example of structure at the 'word level', the horizontal lines being added merely to make the structure easier to discern.



### 3.3. SEMANTIC LEVEL

#### 3.3.1 Thompson's features

In L and LP, there is no consideration of the semantic content of words. In Thompson (1980) words are marked for semantic features, as follows:

##### i) [ $\pm$ SALIENT]

'Feature of words, determined by the grammatical category thereof,...' (Thompson 1980:138).

+salient words are: adjective, adverb, demonstrative pronoun, intransitive preposition (e.g. the/night be/fore), noun, negative auxiliary (e.g. that/weren't/opened), negative (e.g. not, never), number, participle (past, passive, present), quantifier, verb, verbal particle.

-salient words are: article, auxiliary, complementiser, conjunction (though there is uncertainty with 'since', 'because', etc.), copula ('be', 'become', also main verbs functioning as copula, e.g. 'grow' + adj.), determiner, dummy ('there', 'it'), particle ('of' in e.g. /all of/those, 'to' in e.g. /went to/sleep, have to/clean), possessive pronoun, preposition, pronoun, relative pronoun.

The distinction between +SALIENT and -SALIENT seems to correspond, though only very roughly, to the well-known distinction between lexical and nonlexical words, respectively - i.e. members of open and of closed sets respectively. Lexical words refer to objects or actions in the real world, or else have some kind of semantic content, while nonlexical (or 'grammatical') words are devoid of semantic content or reference to the real world, and possess only a grammatical or logical function. The distinction, however, is not perfect in all cases.

##### ii) $\pm$ CONTENT

'Feature of groups or words or rarely syllables (+CONTENT only), assigned prior to the footmaker on a semantic basis, basically to reverse the effects of  $\pm$ SALIENT.' (loc.cit.)

iii) +HIGHLIGHT

'Feature of groups or words or rarely syllables, assigned prior to the footmaker on a semantic basis, for emphasis or contrast.'(loc.cit.) This feature never seems to take a negative value, in Thompson's use, and so is not of the same order as the other features.

iv) ±FOCUS

'Feature of intonational sub-constituents, groups, words, or syllables. +FOCUS is assigned prior to the footmaker on a semantic basis, to indicate the locus of principal communicative content within each constituent.'

Thompson's features are used to locate accented syllables and main stress, drawing on the LP 'metrical' theory of stress. It seems curious to have a feature, ±CONTENT, merely to reverse the effects of another feature, ±SALIENT, and there is no very satisfactory explanation of how exactly the features ±CONTENT, +HIGHLIGHT, and ±FOCUS are assigned in the first place. Also, the use of binary features is limiting, necessitating the splitting of the notion of semantic weight into two or more features, when a scalar type of representation would be less clumsy. In fact, something like Thompson's feature system can be used as the basis for an n-ary description of semantic weight, as follows.

The feature ±SALIENT is based on the notion that certain syntactic word classes nearly always have at least some semantic content or reference to the real world, while other word classes nearly always have none. This insight is a valuable one, since it allows for especial characterisation of any words which depart from the norm, and which therefore should not be indiscriminately classed together with words that have that value as their norm. This primary division can be seen as analogous to the feature [±full] for vowels, which distinguishes vowels which are at least potentially stressable from vowels which are non-starters as far as stress is concerned. When one of the latter class of vowels departs from the norm, the vowel's quality is changed, and also there is likely to be a great increase in duration and intensity, which does not occur in syllables normally containing a vowel of this quality. Similarly, a normally -SALIENT, -CONTENT word, when given especial emphasis so that it becomes +CONTENT, has extra intensity and duration not found in words that are normally +CONTENT. Also, a normally +SALIENT word can become completely

devoid of semantic weight. This is the phenomenon of 'anaphoric destressing', whereby some material in the utterance, being in some sense 'old information' (e.g. mentioned previously in discourse, understood by both speaker and hearer), becomes completely redundant semantically and is consequently destressed. For example, a possible answer to the question; 'When did you see my husband?' would be 'I saw your husband yesterday afternoon', in which the words 'I saw your husband' are old or 'given' information, and therefore non-prominent, while 'yesterday afternoon' is the 'new' information, and receives main stress. In the case of anaphoric destressing, under Thompson's system, a normally +SALIENT, +CONTENT word would become +SALIENT, -CONTENT (or perhaps even -SALIENT, -CONTENT, though he does not go into this).

### 3.3.2 Matrix versus grid

One drawback of using features such as these is that a full feature specification will have the structure of a matrix. That is, every cell in the matrix may be occupied by either a plus or a minus value; although there may be a degree of redundancy or predictability stemming from the nature of the features, there is none at all stemming from the nature of the matrix framework, since this does not constrain formally the permissible values at various positions. To take an example from segmental phonology: a segment which is [-continuant] must also be [-syllabic] (in English at least), and a segment which is [+syllabic] must also be [+voice] (again, in English at least). There is no particular directionality in redundancy for these features. Features dealing with place of articulation have only a limited degree of directionality, and even this is based on the configuration of the vocal tract, rather than on the formal nature of the features themselves. The features are usually arranged in a list beside the matrix, beginning with the major class features. However, this is only by convention, and there is no justification in formal terms for not listing the features in random order.

It is at this point that a feature matrix differs from a 'metrical grid' of the type formulated by LP (see section 3.5.). In a metrical grid, a constituent can be present at a given level only if there is a constituent present in the same column at the previous level. Thus the order of the levels, or rows, is critical, and what happens at one level has an automatic effect on later levels of the same column. Redundancy rules can therefore be subsumed into the general principle that a place-holder in a given column at one level assumes the presence of place-holders in the same column at all previous levels. This principle is not based on the inherent nature of each row (with segmental phonology this nature would differ from row to row, e.g. [ $\pm$ continuant], [ $\pm$ low]), since there is no such difference in kind between rows, rather a difference in degree. The principle is based on purely formal properties of the metrical grid and of the hierarchical prosodic structure that it represents in serial form; i.e. the preservation of relative prominence under embedding and the consequent need to recognise distinct levels of relative prominence, symbolised in the grid as the different rows.

The idea of a grid construct seems a useful one in connexion with semantics also. Thompson's features, like any other type of feature, have no formal directionality of the type outlined above, and any tendency to such directionality would have to be handled by ad hoc constraints.

### 3.3.3 A semantic grid

Instead, I will describe a kind of 'semantic grid' to deal with the phenomenon of relative semantic weight which Thompson's features are intended to describe. The grid has an inbuilt directionality of the kind outlined above, and each successive row is to be seen not as a particular feature, but as a new degree of the same feature, [content], meaning 'relative semantic weight'. It will turn out to be n-ary, taking values from 0 to 3. However, it is not merely a disguised simple n-ary feature, as will be seen, since it takes account not merely of inherent semantic weight but also of relative, or syntagmatic, semantic weight in a particular utterance.

For the first level, there is a representation of the phenomenon captured by Thompson's  $\pm$ SALIENT, analogous to  $[\pm$ full] for vowels; i.e. the distinction between words that are potentially stressable and words that do not have even the prerequisite semantic content to be stressable, at least in the particular utterance under discussion. This distinction can be represented as follows:

First the word class of each word (or ' $\emptyset$ ') is determined, and on the basis of this, a place-holder assigned to every word having the first degree of semantic weight. These are more or less the word classes designated as +SALIENT by Thompson, with the exception that prepositions, marked by him as -SALIENT, are accorded the first degree of [content]. All other classes have no place-holder: a dash indicates this, while 'o' symbolises the place-holder. An example of the first row follows:

o	o	-	o	o	o	o
N	V	det	N	P	N	N

John gave the information to Susan's uncle

The second level is applicable only to those words which already have the first degree of [content]. Under normal circumstances, this level will replicate the assignment of place-holders on the first level. Exceptions are prepositions and negatives, which have been noted as having a level of semantic content intermediate between other grammatical words and lexical words. L writes: 'Some nonlexical words (e.g. not) seem normally to have a certain prominence, so that in the usual case they do not undergo...attachment, although it is certainly possible for them to 'glom on' [sic] intonationally.' (Lieberman 1975:110) Thus these words will not receive a place-holder at the second level. Similarly, words which are not unexpected in the context will also fail to receive a second 'o', since they have a reduced degree of semantic weight. In the case of anaphoric destressing, the words which constitute 'old' information do not even receive a place-holder at the first level, and so are not represented at the second level. Most utterances, however, will contain neither of these cases of semantic reduction.

o	o		o		o	o
o	o	-	o	o	o	o
N	V	det	N	P	N	N
John gave the information to Susan's uncle						

The third and final degree of [content] is concerned with what Thompson refers to as +FOCUS: 'the locus of principal communicative content within each constituent'. This is normally assigned to one column only; naturally, this must already have two place-holders. It is usually the last such column in the utterance.

						o
o	o		o		o	o
o	o	-	o	o	o	o
N	V	det	N	P	N	N
John gave the information to Susan's uncle						
2	2	0	2	1	2	3

The numbers underneath the grid refer to the height of each column, and are the values of [content] which are then transferred to the full prosodic representation.

### 3.3.4 Semantic grid versus n-ary feature alone

#### **3.3.4.1 Importance of the syntagmatic dimension**

At first sight, it might seem that this is merely an over-elaborate system for determining the values of [content], which could be seen as no more than a simple n-ary paradigmatic feature on the lines of a phonetic feature such as [vowel height]. Since a value of two implies that the value 'one' has been reached and passed, or a value of three implies that the values 'one' and 'two' have been reached and passed, the 'semantic grid' is quite unnecessary if semantic content is seen as a purely paradigmatic phenomenon. The value of a grid construct lies in its formal property of allowing relationships within each level to be represented in an explicit

fashion. Naturally, this is only relevant if syntagmatic relationships are seen to have a bearing on a word's semantic content.

I would argue that they do indeed have an influence, and that this influence is most clearly represented by the semantic grid. The semantic prominence of a word is affected by the other words present in the utterance, as well as by any inherent properties of the word. This can be seen in the case of a preposition, e.g. 'underneath'. In most cases this will have the value [1 content], as in the following:

						o
						o
						o
	-	-	o	o	-	o
	N <sub>pron</sub>	aux	V	P	det	N
It was placed underneath a newspaper						
	0	0	2	1	0	3

In utterances with few lexical words, however, this word is capable of carrying greater weight semantically, as in the following:

						o
						o
						o
	-	-	-	o	-	o
	N <sub>pron</sub>	aux	V	P	det	N
It was placed underneath a newspaper						
	0	0	0	1	0	3

In this case, the last word is the only one to have more than the first degree of [content], while, of the remaining words, the preposition alone has any value of [content] greater than nought. Thus it assumes greater relative prominence, while remaining subordinate to the last word in semantic weight. Under certain circumstances - such as emphatic speech, where the speaker intends to make every word count - this configuration will result in the preposition becoming prominent, for want of any other non-final stressable word. In such a case, the preposition becomes [2 content] by a rule applying to the second level only, attempting to minimise the intervals between place-holders.

In some utterances, the preposition may be the only word with any positive [content] value, as in the following, where the second pronoun is not the emphatic pronoun:

-	-	0	-
N <sub>pron</sub>	aux	P	N <sub>pron</sub>
It	was	underneath	it
0	0	1	0

Similarly, the preposition here would become [3 content], by a rule scanning the second level and inserting a place-holder if there are none to begin with, or if an unacceptably long stretch must be broken up. The third-level place-holder is inserted on the last 2-column, as usual.

These examples show that an identical [content] value can have varying semantic weight, relatively speaking. This is shown most clearly by a grid, since what determines semantic weight is the relationship between words at each level; this is something that cannot be expressed by the simple number values of [content]. The principle holds even at the first level, as seen in the case of the following, where neither pronoun is of the 'emphatic', 'pointing-out' type, that can be stressed:

-	-	-
N <sub>pron</sub>	cop	N <sub>pron</sub>
It	was	it
0	0	0

It is very difficult to produce this utterance without either a) making one or both of the pronouns into the emphatic type that insists on the identity of a particular thing as distinct from anything else in the world, or b) giving emphatic stress to the copula as if to deny an earlier negative assertion or presupposition. On the basis of word class, all the words are semantically empty. However, a rule of the type mentioned previously would also apply to the first level and, finding no place-holders at all, would proceed to insert one. As there is no clue as to which word would receive one, this would be done either a) on the word that the speaker wishes to emphasise, or b) on the last word, if there is no

emphasised word. Rules at the second and third levels, following the usual process, would assign [3 content] to the word thus marked out. If there were no emphasised word, and the last word thus became [3 content] by default, the result of putting the focus of the utterance at this point would be to make it seem as if that word were emphasised, and not merely the unemphatic bearer of main stress. Even without special devices (such as a pitch glide or a jump to higher pitch on that word), it would strike a native speaker as sounding emphasised in some way, simply because such a word does not normally receive any stress at all. This is the reason why the procedure at the first level is slightly different (in that the assignment of place-holders can depend on speaker intention, not on previously-existing place-holders). At the second or third levels, the same rule need only 'look for' a previously-existing column or columns to break up an unacceptably long stretch. The result of this operation will be not to produce the emphatic type of stress, but simply to produce an utterance with unemphatic stress on a word which, though it has some positive value of [content], does not normally have this amount of stress in the utterance.

### 3.3.4.2 Inability of an n-ary feature to deal with the syntagmatic dimension

It may be argued that the same results can be obtained more simply by using numbers from the start. It is probably true that the same results can be obtained- in most cases, at least - but the process would be far less simple. One would begin by assigning values of [content] to all words; 0 for non-lexical words, 1 for prepositions, negatives and any words with reduced semantic weight in context, 2 for all lexical words but the last (excluding those mentioned under value 1), and 3 for the final (or the focussed) lexical word (where this is not one of those mentioned under value 1). This would present few problems in most utterances; but in some cases, such as those discussed earlier, there would need to be revision. For example, in the sentence

It was underneath a newspaper  
0   0   1   0   3

given the difficulty of producing the non-focussed material with the requisite low amount of stress, one could apply a rule roughly as follows: -

[1 content] ---> [2 content] / Ø Ø (Ø) \_ Ø

Similarly, in the sentence

It was underneath it  
0   0   1   0

this rule applies, together with a rule changing [2 content] to [3 content] in the absence of any original [3 content] word. In the sentence

It was it  
0   0   0

there would be a need for a different rule, such as the following -

[0 content] ---> [1 content] / \_##

(condition: no original [1, 2 or 3 content])

This would only hold for those utterances of this type where no one word receives especial emphasis. The other two rules would then apply.

### 3.3.4.2.1. Thompson's algorithm

The process outlined above is an awkward process, and would probably require several more rules in order to cover all the possible configurations. More promising is the type of algorithm suggested by Thompson for the metrical grid (of which more later). The algorithm is as follows: -

F1 For all words in the input constituent, normalize and adjust their metrical grid as follows:

1) If the word is +SALIENT, make the height of the column over the DTE of the word equal to 4. Take the DTE of a one syllable word to be that syllable. [DTE = 'designated terminal element': that syllable of a word which is reached by following only 's' nodes down the metrical tree of the word.]

2) Make the height of all columns which are neither over the DTE nor of height 1 equal to 2.

3) For each syllable in the word which is marked -FOCUS, decrease the height of its column by 1, unless it is already a 1-column.

4) If any syllable of the word is +CONTENT [i.e. Thompson's version of +CONTENT], set the height of its column to 4. 5) For each syllable in the word which is -CONTENT, decrease the height of its column by 1, unless it's already a 1-column.

F2A Proceeding from right to left through the metrical grid of the

constituent, identify a column as clashing on the basis of the membership of the Clash Set at that point.

The Clash Set is initially empty (at the right end of the constituent), and its membership is updated at each grid position as follows on the basis of the column at that position:

1) If the column is a 1-column, there is no clash. Remove 2 from the Clash Set and carry on.

2) Otherwise, if the column is a 2-column, then if 2 is a member of the Clash Set, there is a clash. Otherwise, remove 4 from the Clash Set, add 2 to the Clash Set, and carry on.

3) Otherwise, the column is a 4-column. If either 4 or 2 is a member of the Clash Set, there is a clash. Otherwise, add 2 and 4 to the Clash Set and carry on.' (Thompson 1980:139)

Thompson's algorithm assigns a single value to each word, on the basis of its semantic features; this is roughly equivalent to assigning each word a value of [content] on the basis of word class membership, etc. from the outset. The algorithm then applies to the sentence in a linear fashion, identifying clashes, which would need to be resolved by means of a rhythm rule where possible. Thompson's 'metrical grid' is rather different from that of LP; for although he assigns column heights to words according to LP's principles (following metrical structure, i.e. prosodic/syntactic structure), the first part of this algorithm is concerned with radically modifying - or 'normalizing' - this grid on the basis of the semantic features assigned by Thompson. Therefore his grid cannot truly be called a metrical grid, since its composition, and the relationships between its constituents, are determined mostly by semantic considerations. It is in fact closer in character to the 'semantic grid' proposed here, though there are important differences between the two approaches.

### **3.3.4.2.2. Thompson's algorithm really an n-ary feature: a modified version**

The algorithm presented by Thompson in 'F2A' (for the identification of clashes) applies to each unit as a whole and does not break up the grid into its component levels. Therefore, there is little point in using a grid construct; single numerical values would do this job far more

simply. Possibly a Thompson-type algorithm, adapted to deal with 'columns' represented explicitly as what they essentially are, i.e. numerical values, could come nearer to solving the sort of problem presented in section 33.4.2.0., where each value applies to one unit (i.e. word), and is processed as a whole. This algorithm would not primarily be a process of locating clashes, but would be a series of rules assigning values of [content] to words according to context as well as according to word class. For instance, given an input of a string of words with utterance-final boundary and word class specifications, the sentence

It was placed underneath a newspaper

0 0 2 1 0 3

would be generated with these values of [content] by an algorithm that simply took into account word class and utterance-final position. In the sentence

It was underneath a newspaper

0 0 1 0 3

the same rules would apply, but also there might be something analogous to Thompson's 'Clash Set' (only this would store the value of a variable rather than, as in Thompson's algorithm, constant numerical values). The procedure would be something like this: -

At the start of the utterance, set the variable, or 'threshold', to a certain value, and process the first word. This is assigned a certain value by the word class membership rules. If the assigned value is less than the threshold value, subtract one from the latter and move to the next word. If the assigned value is equal to or greater than the threshold value, add one to the assigned value, reset the threshold value to its original value, and move to the next word.

The advantage of this method of description, limited though it may be, is that the original value may be altered to correspond to changes in the overall emphasis of the utterance. A low threshold will ensure that more words have values of 2 or perhaps even 3 as well, while a high threshold will prevent many high values of [content] from appearing. Thus a low threshold corresponds to emphatic speech and a high threshold to unemphatic speech, where few words have much semantic weight.

### 3.3.5 A possible solution

### 3.3.5.1 The problem

Although this algorithm is an improvement, it is still not ideal, since it deals with each word only once and therefore cannot take account of the different degrees of semantic weight in themselves. There is no method of controlling the number or distribution of the various values of [content], since these depend ultimately on word class. Yet in English it seems to be a general fact that there will not often be long uninterrupted strings of lexical words or of nonlexical words; if these do occur, the lexical words will be greatly lengthened, the nonlexical words much shortened - or, alternatively, some lexical words will be reduced in stress, some nonlexical words increased in stress. The most common pattern is a fairly even alternation between lexical and nonlexical words. This kind of fact is difficult to express in terms of the algorithm outlined above, which has no way of dealing with restrictions such as the following (which are geared towards short sentences uttered at medium tempo):

i) there will usually be not more than two words (of either kind) between a [1 content] word and the next [1, 2 or 3 content] word; that is, it is unusual for more than two [0 content] words to occur together in the type of fairly short sentence under consideration.

ii) there will usually be no more than three words (of either kind) between a [2 content] word and the next [2 or 3 content] word.

iii) there will usually be no more than five words (of either kind) between a [3 content] word and the next [3 content] word.

Some examples follow: -

					0
	o	o			o
-	o	o	o	-	o
<hr/>					
The cat sat on the mat					
0	2	2	1	0	3

$$\begin{array}{r} \phantom{0} \phantom{0} \phantom{0} \phantom{0} \phantom{0} \\ \phantom{0} \phantom{0} \phantom{0} \phantom{0} \phantom{0} \\ - \phantom{0} \phantom{0} \phantom{0} \phantom{0} \phantom{0} \\ \hline \text{The dog bit the man} \\ \hline 0 \phantom{0} 2 \phantom{0} 2 \phantom{0} 0 \phantom{0} 3 \end{array}$$

			o					o
		o	o		o	o		o
o	o	o	-	-	o	o	-	o
P	N	N	N <sub>pron</sub>	aux	V	adj	det	N
By Tuesday afternoon, he had done all the work								
1	2	3	0	0	2	2	0	3

### 3.3.5.2 The algorithm

It can be seen that there is a reasonably even distribution of [content] values, but not a random distribution: [3 content] is relatively rare, and [2 content] relatively common. The following is a tentative solution to the problem of describing these facts, and is by no means a fully elaborated theory.

The 'semantic grid', under the proposed approach, is dealt with level by level, starting at the first degree of [content] and working from 'left' to 'right' through the utterance. As before, there is a threshold variable, with a reset value of 1 at this level. The algorithm takes the form outlined earlier, with the difference that this is now formulated in terms of assigning rather than locating and modifying. (However, unlike Thompson's proposals, this is not intended as a production model, but merely as a set of well-formedness conditions on the final representation.) As the algorithm proceeds along the string, for each word encountered and assigned no value, a value of 1 is subtracted from the variable. As the latter reaches a lower value, it becomes more likely that a word will be assigned a place-holder (i.e. will take at least this degree of [content]), if it can possibly have this (i.e. if it is either lexical or especially emphasised). After assigning a place-holder, the threshold value is reset and the procedure continues as before to the end of the string. Naturally, when the threshold value is as low as 1, it is quite likely that several successive words will be assigned a place-holder at this level. There is no absolute necessity for one to be assigned as soon as the threshold reaches 0 (on any level); it is merely more likely to be assigned as the threshold decreases. This aspect of the algorithm is intended as a crude approximation to the high degree of optionality

present in language; speakers may vary the relative semantic content of words according to the information they wish to communicate, and many statements, in semantics particularly, can only be framed in roughly statistical terms based on the most frequently-occurring patterns.

The procedure is the same at the second and third levels, the only difference being the reset value of the threshold. On the second level, for the type of simple short sentences exemplified earlier, the reset value would be 2, while on the third level it would be 4. The values are changed in accordance with the overall length and/or emphaticness of the utterance, a low threshold corresponding to a short sentence or emphatically-spoken long sentence, and a high threshold to a longer sentence or unemphatic short sentence. The difference between the reset values of the second and third levels need bear no relation to that between the reset values of the first and second levels, and will in most cases be substantially greater. These features also are intended to reflect the optionality inherent in language use.

#### **3.3.5.3 Flexibility in the algorithm**

There is another point concerning the reset values of the threshold at each level. At the first level, this would almost always be 1; the assignment of place-holders is largely dependent on whether the word is lexical or otherwise eligible at this level. It is only in cases of anaphoric distressing, contrastive stress, or unaccustomed emphasis of nonlexical words, that the speaker's intention overrides the normal pattern. On the second level, there is slightly more freedom; although the assignment of place-holders may be partly dependent on word class, there is also more scope for processes such as a slight reduction in semantic weight for a word which is not unexpected in context (whether discourse setting or situational context), or a slight increase in stress for prepositions and negatives that take on greater importance in this particular utterance. At the third level, there is even more freedom, the main condition being the general principle, applicable at every level, that a place-holder may only be assigned to a word at a given level if a word already has one at the next lowest level. To reflect this state of affairs, one could propose a condition on the reset values such that at the

first level the value is practically always 1, at the second level it is usually 2 but may be 3, and at the third level it may be 3, 4 or 5. Naturally, these values are only applicable to shorter sentences, and would need to be set higher, especially at the higher level (which would also have a wider spread of permissible reset values) for more complicated utterances.

This proposed algorithm is intended purely as part of a formal description; no claims at all are made for psychological reality. The algorithm is based on the notion that it is necessary to take into account not merely the overall semantic weight of a word, but also the relationships holding between words at each level of semantic weight, in order to give some sort of account of the patterns of distribution of the different values of [content] (corresponding to relative semantic weight) in utterances of varying lengths.

### 3.3.6 Semantic and metrical grids

At this point it should perhaps be noted that the 'semantic grid' is not completely identical in kind to an LP-type 'metrical grid'. The metrical grid is related to a hierarchical prosodic structure, or 'metrical tree', by means of a 'Relative Prominence Projection Rule' or equivalent. The metrical grid will be described more fully in section 3.5.: the important feature to note in the present connexion is that one of its functions is to represent relative prominence under embedding. That is, it has an intimate link with a previously-defined hierarchical prosodic structure, which in its turn is loosely linked with the (surface) syntactic structure of the utterance.

The semantic grid on the other hand has been described without reference to any type of previously-defined hierarchical structure, and is based solely on word class membership and what I have loosely referred to as 'speaker's intention'. According to some recent theories of semantics, such as generative semantics, the semantic representation of a sentence may be in the form of a hierarchical structure similar to syntactic structure. If this is so, then there could be a case for linking semantic

structure with a 'semantic grid' in a parallel fashion to the linking of metrical structure with the metrical grid. However, I do not propose to discuss theories of semantic representation here, and so the semantic grid will remain defined without reference to any possible semantic structure. This, in fact, is the approach taken by Liberman towards defining the metrical grid: although the grid is intimately connected to a metrical tree, it is not itself such a tree structure, and can be defined in terms which make no reference to tree structure, i.e. as an 'ordered set of ordered sets'. Liberman writes; 'The primary purpose of this formalization we have just sketched is to make it clear that the notion of a METRICAL GRID can be defined in a way which does not make it a tree structure (i.e., it is not necessarily 'rooted', no constituent structure is established, etc.). The particular kind of formalization we have chosen (a mapping among elements of ordered sets) is simply an attempt to render formally the intuition of hierarchically-related periodicities.' (L:186)

The semantic grid is therefore similar to the metrical grid in that it is not a tree structure, though capable of being related to one, and can be defined in terms which do not presuppose such a tree structure. Also, like the metrical grid, the semantic grid is intended as a formal means of representing 'hierarchically-related periodicities'. In the case of semantics, this refers to the phenomenon whereby words that are 'strongest' (i.e. have greater semantic weight) occur less frequently and are evenly distributed in the sentence. The nature of 'rhythm' is still far from settled, but one could see it in very general terms as a repetition in time of units that are (or tend towards being) equivalent in length, the boundaries between such units being signalled in a consistent fashion. If one takes this view of rhythm, then the semantic as well as the metrical representation of the utterance could be described as rhythmical, in view of the evenness of distribution of semantically important elements.

### 3.3.7 Some advantages of this type of analysis

#### **3.3.7.1 Unitary treatment of anaphoric destressing and nonlexical words**

One advantage of an analysis that incorporates semantic considerations is the fact that there can be a unitary treatment of anaphoric destressing and nonlexical words. In assigning both nonlexical words and anaphorically destressed words the same (zero) value of [content], the claim being made is that an anaphorically destressed word is a lexical word reduced to the status of a nonlexical word because of the context of the particular utterance. Such destressed words refer to objects, etc., that have already been mentioned, or which are in some sense present in the minds of both speaker and hearer. The same could be said of many nonlexical words, such as pronouns, which are only interpretable if one knows the particular object in the real world that is to be taken as the referent of 'he', 'they', etc., in the particular utterance. That this is so is indicated by the common rejoinder, 'Who's we?', requesting an identification of the persons characterised as 'we' when this has not been made sufficiently clear. (Other nonlexical words do not even have this indirect relation to the real world, and exist solely in the language system: e.g. articles, particles, auxiliary verbs.) So there would seem to be some support for grouping nonlexical words together with anaphorically destressed words, and this is done in the proposed analysis by denying them even the first degree of [content]. Also, on the phonological level, both these types of words, unlike lexical words, are likely to appear with reduced vowels and will not be intonationally prominent.

#### **3.3.7.2 Different functioning of lexical and nonlexical words**

Another difference between nonlexical and lexical words, brought out by this type of analysis, is seen in utterances where a nonlexical word is emphasised for a particular purpose. If a lexical word is given such 'contrastive' or 'emphatic' stress, what is focussed on is an object, action, characteristic, relationship, etc., that is seen as existing in the real world. On the other hand, when a nonlexical word is given this type of

stress (and, unlike lexical words, it cannot be stressed without being 'contrastive' or 'emphatic'), then what is focussed on is not something seen as existing in the real world, but something closer to a logical operator, a more abstract entity expressing logical relationships between other entities without itself having any independent value. The following is an example of the use of stressed nonlexical words: -

"'You ARE Zaphod Beeblebrox?' it squeaked.

'Yeah,' said Zaphod, 'but don't shout it out or they'll all want one.'

'THE Zaphod Beeblebrox?'

'No, just A Zaphod Beeblebrox, didn't you hear I come in six packs?'"

(Adams 1980:37)

In the first line of this example, the copula receives stress, even though it is a nonlexical word. This stress is of the emphatic type, rather than the contrastive. Contrastive stress points out an explicit antithesis between two particular words, while emphatic stress makes a distinction between a particular word and any other word which could have occurred at that point - one might characterise their domains of operation as syntagmatic and paradigmatic respectively. In the case of the emphasised 'ARE' in the first line of the example, what is being focussed upon is the function of the copula: the function of establishing identity of reference (when positive) between two stated arguments. The fact that the overall utterance functions as a question (with rising intonation, symbolised in print by a question mark), means that what is being questioned is the validity of identifying the 'you' with the referent of 'Zaphod Beeblebrox'. The focus is not on one or both of the arguments, but upon the logical operation of expressing identity between two or more arguments. If this type of stress (i.e. 'emphatic') had been on a lexical word, the focus would have been not on a logical relation but on an object, attribute, action, etc., in the real world.

The second stressed nonlexical word in the example is 'THE', and this likewise is a case of emphatic stress, asserting the occurrence of the definite article rather than any other word which could have appeared at this point. The function of the definite article is to indicate definiteness of reference, so that the following noun is taken to refer to a particular entity (where singular) in the real world, known to both

speaker and hearer. The utterance in this case functions as a question, and the overall effect is to question the proposition that the entity claiming to be Zaphod Beeblebrox is in fact the particular (and well-known) being of that name, known to both speaker and hearer. Were the proper name, and not the definite article, to receive emphatic stress in this utterance, the effect would be quite different. The effect would then be to question the proposition that the entity claiming the name 'Zaphod Beeblebrox' was in fact entitled to the name in the first place; that is, similarly to the first line of the passage, what is being questioned is the validity of identifying the entity addressed with the entity known as 'Zaphod Beeblebrox'. As in the first line, this questioning can be done explicitly by using an emphatically stressed copula; but it can also be done implicitly by emphatically stressing the proper name itself, since a name or noun, being lexical, makes implicit reference to an object in the world whether it is stressed or not. A nonlexical word, on the other hand, makes no such reference independently, and so the effect of emphatic stress on 'the' is to focus on its function as indicator of definiteness, not on the identification of the hearer with the named entity. In other words, the speaker does not question that the hearer can correctly be designated 'Zaphod Beeblebrox'; he is merely asking whether a relationship of definiteness of reference holds between this Zaphod Beeblebrox and the well-known being of the same name. Thus lexical and nonlexical words behave in a different manner under emphatic stress, because the nonlexical word has no inherent property of picking out a certain object, action, etc. (or a set of such objects, actions, etc.) in the real world. The property which is in fact possessed by this kind of nonlexical word is the property of indicating logical or abstract relations between lexical words. Even pronouns, which do not merely express abstract relations, are nevertheless dependent on other information for full interpretation.

The third stressed nonlexical word in the passage is 'a'. This is given contrastive stress, as the speaker wishes to make an explicit antithesis between this word and the preceding stressed 'the'. The function of the indefinite article is to indicate indefiniteness of reference - that is, an identity of reference with a named set of entities is asserted, but there is no clue given as to correspondence with a particular member of

that named set. In this utterance, what the speaker is emphasising is this property of indefiniteness; he is making no statement at all about identity of reference. The explanatory rhetorical question at the end of the passage confirms both that he is not questioning the identity between himself and the named set of entities (hence the use of 'I' to refer to the set designated 'Zaphod Beeblebrox') and also that what he is denying is the validity of picking out a particular member of that set: the question makes it plain that he sees the set as consisting of more than one member, and that he himself, though a member of that set, is not further specified in terms of definiteness of reference.

The point of this example has been to show that there are important differences between lexical and nonlexical words, which must be accounted for by any theory which includes some consideration of semantics. The analysis that has been put forward here does indeed make the distinction, in that nonlexical words are nearly always [0 content], while lexical words nearly always take a positive value of [content]. This distinction is based on aspects of language use which are related not to structure but to word class; it corresponds to the first level of the semantic grid, where the primary distinction is made between words that can take some value of [content] and words which can only be [0 content]. Subsequent levels are increasingly less dependent on word class, and more dependent on semantic weight relative to other words in the utterance. Thus there is room for considerations of both relative and inherent semantic content in the analysis proposed here. The formulation in terms of a semantic grid enables the relationship between inherent and relative aspects of semantic content to be characterised in a purely formal way that takes due account of both aspects.

### 3.3.8 [content] and the prosodic representation

The end product of the semantic grid is a series of values of [content], one for each prosodic word. These can be transferred to the full prosodic representation, forming a kind of 'semantic level' as in the following example, where the values of [content] were obtained from the grid at the end of section 3.3.3.:



phenomena is another question; the preceding has only been a rough outline of what a fully-developed theory along these lines would look like, and I have no doubt that a great deal has been left unexplained. Nevertheless, even at this stage, an analysis that takes account of semantics seems preferable to one that does not. This point was made by Bolinger, who wrote: 'The distribution of sentence accents is not determined by syntactic structure but by semantic and emotional highlighting. Syntax is relevant indirectly in that some structures are more likely to be highlighted than others.' (Bolinger, D. 1972:644) The same author also points out that semantic content cannot be reduced to a binary feature assigned mainly on the basis of word class (as Thompson's features are), when he writes (of utterances where the verb is highly predictable from the preceding noun); 'It is not necessary for the verb to be fully predictable from the noun; what counts is RELATIVE semantic weight. For this it is necessary to take account of the entire context, including the context of situation.' (op. cit.:635) The assignment of place-holders at the first level of the semantic grid on the basis of word class reflects the fact that some syntactic categories are more likely to be 'highlighted' than others, but the subsequent levels of the semantic grid take more account of the 'entire context', expressing relative semantic weight when reduced to numerical values of [content].

The analysis proposed here provides a means for accounting for phenomena of increase or decrease in stress under special conditions (see section 3.6.). As Bolinger wrote: 'The error of attributing to syntax what belongs to semantics comes from concentrating on the commonplace.' (op. cit.:634) It is precisely through not concentrating on the commonplace that the proposed analysis can account for a wider range of phenomena than can either Thompson's or Liberman's analyses.

### 3.4. METRICAL STRUCTURE (PHRASAL LEVEL)

#### 3.4.1 Selkirk's units

This can be dealt with more or less as LP do, assigning 's' and 'w' (for 'strong' and 'weak' respectively) to nodes of metrical structure. However, Selkirk makes some proposals for this level which seem to be a definite improvement on the original theory.

Selkirk (1980b) recognises three types of prosodic unit of a level higher than that of the prosodic word, namely: the phonological phrase ( $\phi$ ), the intonational phrase (I), and the utterance (U).

##### **3.4.1.1 Phonological phrase**

According to Selkirk, one or more prosodic words ('w' units) may form a phonological phrase, the order of nodes dominated by the  $\phi$  being w-s. The length of a  $\phi$  is optional, as is its precise constituency in terms of number of w units. Selkirk introduces the Rhythm Rule, which adjusts metrical structure to avoid stress clashes, and the Monosyllable Rule, which reduces nonlexical words to the status of weak syllable from the status of weak w unit. She writes: 'There is abundant evidence that a weak monosyllabic non-lexical item will not reduce to  $\sigma$  [syllable] unless it is weak with respect to a strong CONTAINED WITHIN THE SAME  $\phi$ .' (Selkirk 1980b:19) She also writes that her 'Rhythm Rule' likewise operates only within the domain of the  $\phi$  - it is blocked by the boundary between subject and verb phrase. She gives examples from French also, showing that the segmental phenomenon of liaison operates only within a phonological phrase, and writes, 'What we are calling the phonological phrase is thus the privileged domain of two rules of English which apply to and modify prosodic structure. In other languages such as French and Modern Greek there is evidence from rules of the SEGMENTAL phonology that what we define as the phonological phrase constitutes a special domain.' (op.

cit.:20) The principles she gives as governing the internal structure of the phonological phrase refer directly to the syntax of the sentence. There are two such principles regarding the constituency of the  $\phi$ , as follows:

i) An item which is the specifier of a syntactic phrase joins with the head of the phrase.

ii) An item belonging to a 'non-lexical' category....such as Det, Prep, Comp, Verb<sub>aux</sub>, Conjunction, joins with its sister constituent.' (op. cit.:15)

Within the  $\phi$ , Selkirk specifies prominence relations by what is essentially the LP formulation of the Nuclear Stress Rule:

'Given the two sister nodes of prosodic structure [ $N_1$   $N_2$ ] within  $\phi$ ,  $N_2$  is s (and  $N_1$  hence weak).' (op. cit.:18)

However, Selkirk's analysis has no way of explicitly characterising a nonlexical word. This contrasts with the proposed analysis, where nonlexical words are those marked [0 content] (originally, at least). This assigning of [content] values depends on a list of the word classes that may receive a place-holder on the first level of the semantic grid. Thus the proposed analysis is explicit where Selkirk's is vague.

### 3.4.1.2 Intonational phrase

Selkirk's 'intonational phrase', or 'I', is made up of one or more phonological phrases, freely grouped together. This hypothesis is meant to account for the possibilities due to variable phrasing in English: in a given sentence, each  $\phi$  may constitute an intonational phrase, or the whole utterance may be one, or else there may be several degrees in between these two configurations. Selkirk characterises the I as follows:

'Constituency. The I is composed of  $\phi$  joined in a right-branching structure.

Prominence. In I, the nodes [ $N_1$   $N_2$ ] are in the relation w/s.

Syntactic Domain. a) Parentheticals, preposed adverbials, non-restrictive relative clauses, etc., are I's. b) Otherwise, the choice is free.' (op. cit.:26)

Selkirk notes that it is also possible to view the I as being not a binary-branching tree, but a multiple-branching tree, with no s/w prominence relations defined. However, she opts for the binary-branching analysis, 'insofar as it is consistent with the overall Liberman and Prince-type approach we have been taking to prosodic structure.' (op. cit.:24) There is no hard evidence for this position, and indeed syntactic structure, on which her trees are loosely based, does not always have a binary-branching form. However, neither Selkirk nor Liberman give any substantial thought to the question of the precise relations between prosodic structure and syntactic structure, though it is clear in both works that the two types of structure are not meant to be identical. The prosodic structure is based on the surface syntactic structure, yet requirement a) of Selkirk's definition of the syntactic domain of the I seems to presuppose some knowledge of deep structure, or at least of derivational history. This is because it is necessary for the conditions defining well-formed I units to 'know' when certain units have been relocated in syntactic structure (preposed adverbials), when they derive from a completely different 'S' in syntax (parentheticals), and their precise deep structure (non-restrictive relative clauses, as against restrictive relatives). Yet Selkirk gives no hint as to how this kind of phenomenon is to be incorporated into the theory, and merely presents examples with unexplained, 'ready-made', I units.

The examples she gives show that the boundaries between I units are the places where it is natural for a native speaker to insert a pause; also, the I units could be seen as corresponding to the 'minor tone-units' of the British school of suprasegmental analysis.

### 3.4.1.3 Utterance

This is the highest-order category of prosodic structure, consisting of one or more intonational phrases. These are 'not obviously in any relation of subordination to each other (though, in the spirit of Liberman and Prince, may be represented as such), and....the utterance usually coincides with the 'highest' sentence in syntactic structure.' (op. cit.:27) Thus the 'utterance' may be seen as equivalent to the 'major tone-unit' of the British school of suprasegmental analysis, made up of one or more

'minor tone-units'. Selkirk has little more to say about the utterance, or 'U' unit.

### 3.4.2 Discussion

Those prosodic units of Selkirk's that are higher than the word - the  $\phi$ , I and U units - seem to be units of the same general kind. Their prominence relations are defined in the order w-s, and each unit may be made up of one or more units of the next level down, in the relation  $w < \phi < I < U$ , so that it is perfectly possible for an utterance to consist in a single word, which would then be described as being also a  $\phi$  and an I unit. One could easily extend the series to include a unit made up of one or more utterances (a 'paragraph'?) which could be useful for describing phenomena such as the well-known feature of 'declination'. This refers to the gradual lowering of the absolute locations of the relative pitch levels through a longer stretch of speech such that, for example, a pitch which counted as 'mid' at the outset would count as 'high' at the end, the pitch levels being reset at the start of a new stretch of speech.

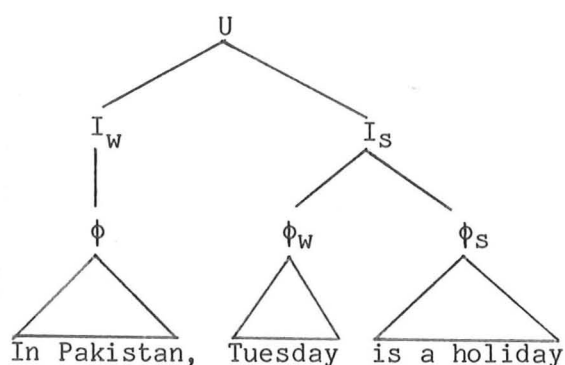
This feature of transitivity in Selkirk's higher-level prosodic units leads me to see them as being essentially the same type of unit, but different in degree. They seem to play a part in intonational patterns, rather than stress or rhythmic patterns, and are largely dependent on the (surface) syntactic structure. From an examination of Selkirk's examples, the following equivalences strike the reader:

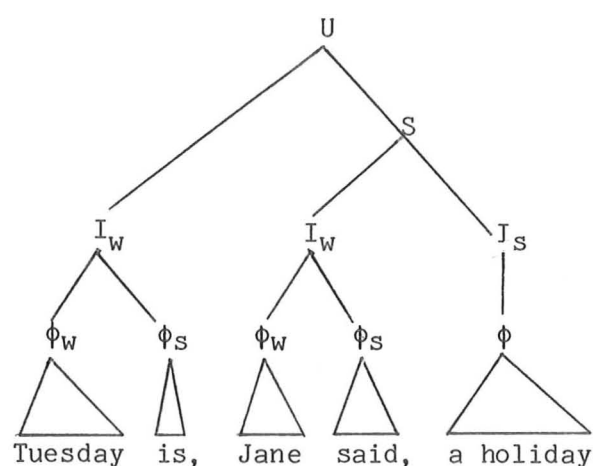
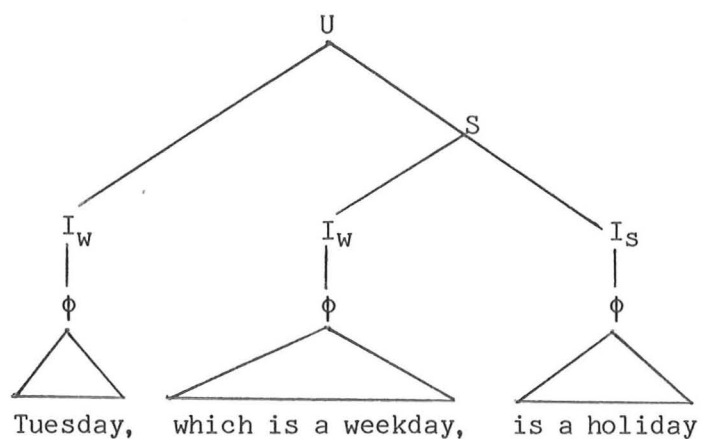
- $\phi \equiv$  'contour', i.e. stressed and pitch-prominent syllable, with any associated non-pitch-prominent syllables. The latter may be stressed or unstressed, as defined by the metrical grid, etc., but are usually unstressed in shorter sentences.
- I  $\equiv$  'minor tone-unit', i.e. one or more 'contours', the last one containing a nuclear tone (usually non-falling in non-final I units), together with any non-pitch-prominent syllables (whether stressed or unstressed, as defined by the metrical grid, etc.) before the first contour.
- U  $\equiv$  'major tone-unit', i.e. one or more 'minor tone-units', ending in a nuclear tone which is particularly linked to sentence-type (e.g.,

falling tone for statements, rising tone for yes/no questions).

It will be noted that 'prominence' in terms of these units is signalled in the actual utterance by pitch-prominence, not by 'stress'. A pitch-prominent syllable will very often be stressed as well, and this could lead to a false equation between the two types of 'prominence'. When the two do not coincide, however, it is the pitch-prominent syllable that counts as 'prominent' for the purposes of defining the above prosodic units. A stressed but non-pitch-prominent syllable, on the other hand, is never the DTE ('designated terminal element', or loosely 'most prominent syllable') of a prosodic unit. However, it may well count as stressed, after the application of the algorithm associated with the metrical grid (see section 3.3.5.1), and would be marked with a dot or line that indicated only stress rather than pitch movement, in the tonetic stress mark systems used by O'Connor and Arnold and other British workers.

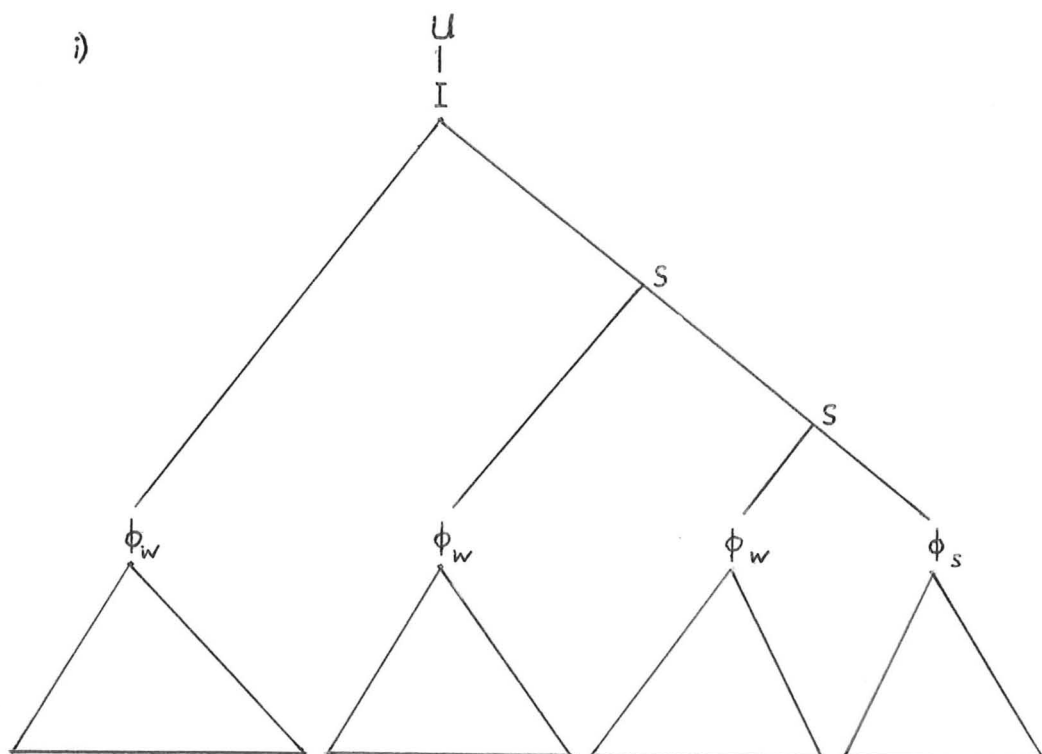
I give examples from Selkirk (1980b:25). The preposed adverbial phrase in the first example, the non-restrictive relative clause in the second example, and the parenthetical clause in the third example, are each necessarily an intonational phrase.





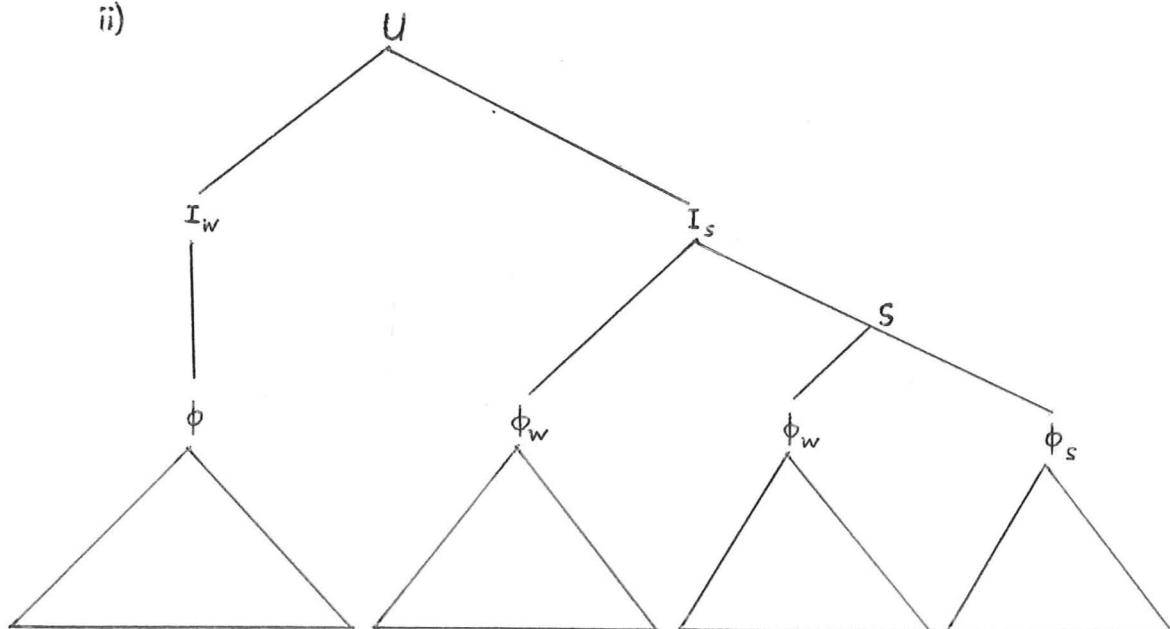
One 's' node in each of the two latter examples is given no designation in terms of type of unit. Selkirk does not mention why her sharply-differentiated prosodic units can be dispensed with so easily, with no discussion; this could support a view of the units as being different only in degree, not in kind, as proposed earlier. (The appended examples of variable phrasing are taken from Selkirk 1980b:22-23.)

i)



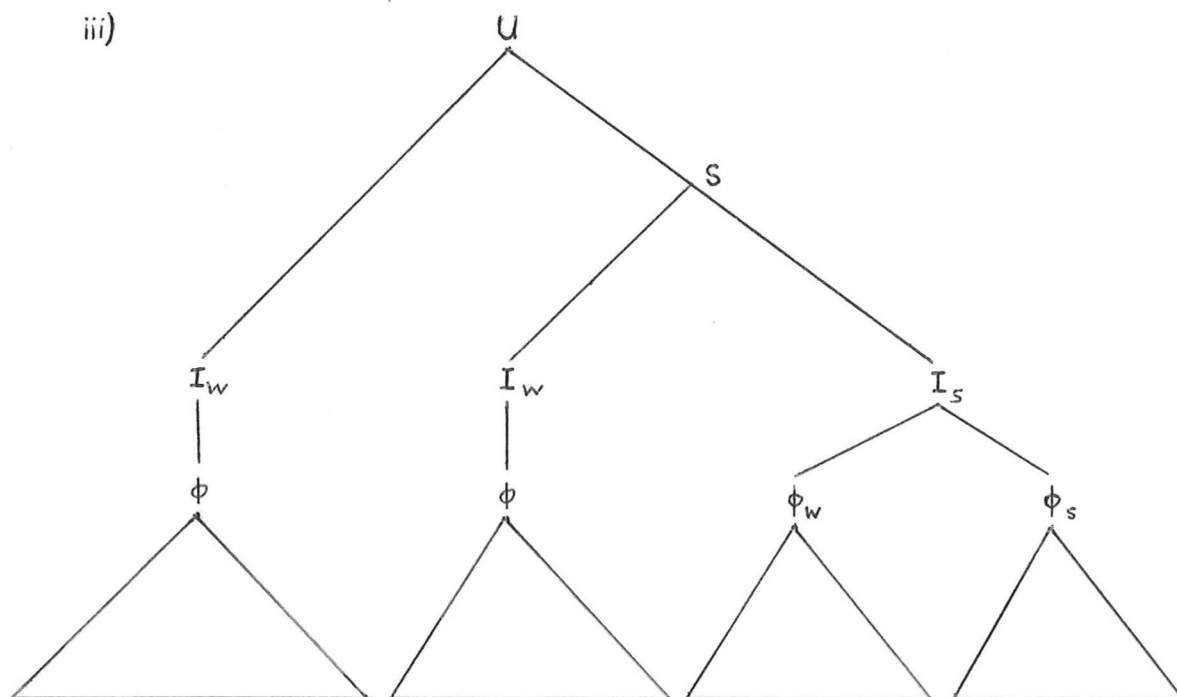
The absent-minded professor has been avidly reading the latest biography of Marcel Proust

ii)



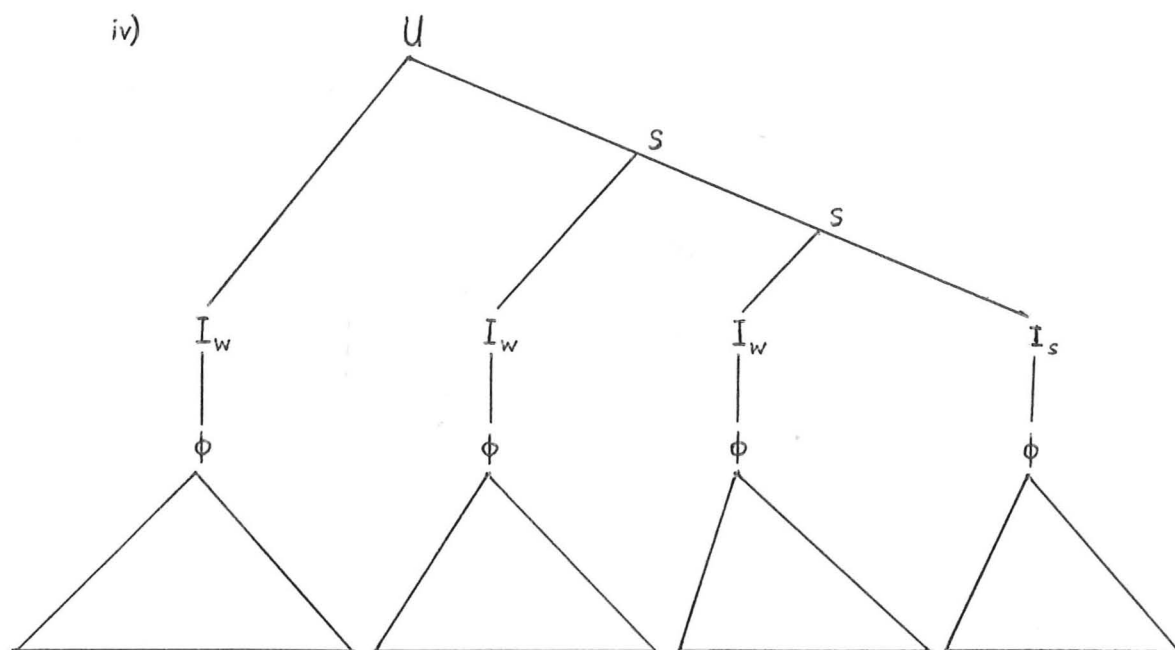
The absent-minded professor has been avidly reading the latest biography of Marcel Proust

iii)



*The absent-minded professor has been avidly reading the latest biography of Marcel Proust*

iv)



*The absent-minded professor has been avidly reading the latest biography of Marcel Proust*

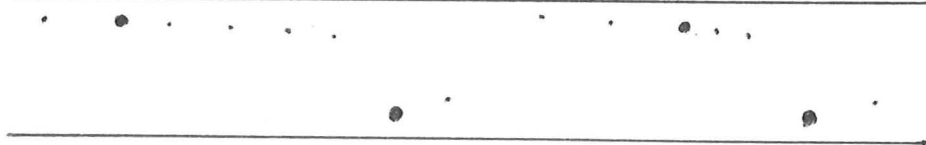
### 3.4.3 Intonation

The appended examples show the variations in phrasing possible in a fairly long sentence, given Selkirk's prosodic units. It would be possible to transcribe each example using tonetic stress marks, as follows - where the nuclear tone types are not the only possible ones:

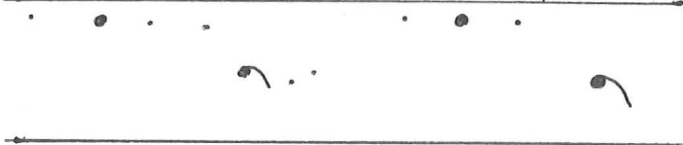
- i) The ·absent-minded pro'fessor has been ·avidly 'reading the ·latest bi'ography of ·Marcel \Proust||
- ii) The ·absent-minded pro<sup>∇</sup>fessor | has been ·avidly 'reading the ·latest bi'ography of ·Marcel \Proust||
- iii) The ·absent-minded pro,fessor | has been ·avidly <sup>∇</sup>reading | the ·latest bi'ography of ·Marcel \Proust||
- iv) The ·absent-minded pro,fessor | has been ·avidly ,reading | the ·latest bi<sup>∇</sup>ography | of ·Marcel \Proust||

Unstressed syllables at the beginning of each I are incorporated into the pitch movement at the end of the preceding I in most cases, but this is not obligatory; normally-unstressed, non-pitch-prominent nonlexical words such as 'has' and 'been' can become pitch-prominent (though not stressed, usually), when at the beginning of a major tone-unit (or 'I' unit), whereas neither unstressed, non-I-unit-initial syllables within a word, nor nonlexical monosyllables not beginning an 'I' unit, may be pitch-prominent - at least not without sounding a little bizarre. This raises the possibility of patterns such as the following, where '·' refers to an unstressed syllable and '●' to a stressed syllable, a tail showing any pitch glide.

The ·absent-minded pro,fessor | has been ·avidly ,reading |



the ·latest bi<sup>∇</sup>ography | of ·Marcel \Proust||



Admittedly, this example exploits the maximum number of possibilities and sounds rather unnatural and stilted. Yet such a style of utterance is sometimes heard, for example from adults speaking to children or from certain television presenters, such as Alan Whicker. An alternative style for the above examples would allow the unstressed syllables in question to continue the preceding pitch movement, as in the following, which sounds a little more commonplace;-

The ·absent-minded pro,fessor | has been ·avidly ,reading |

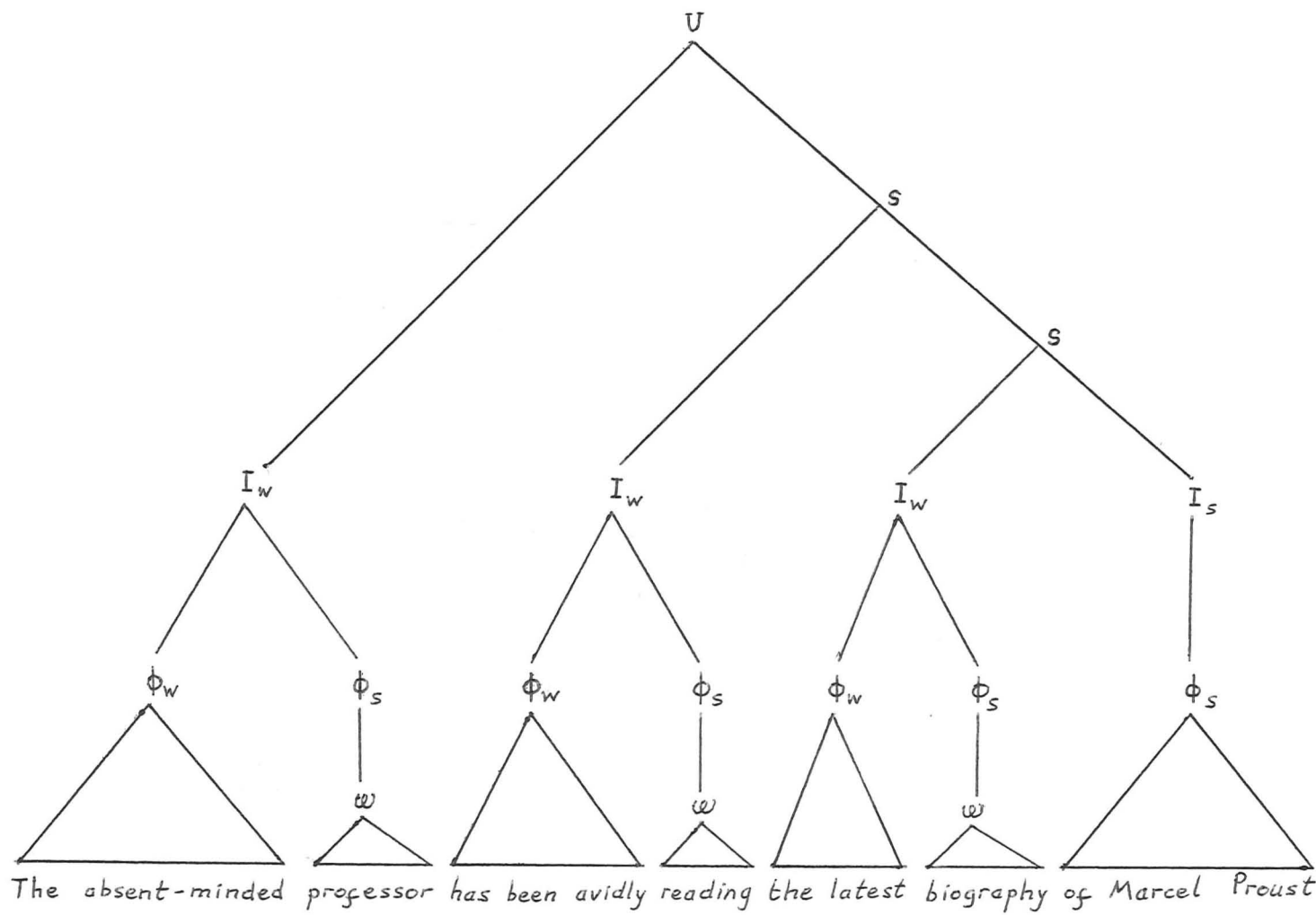


the ·latest bi<sup>V</sup>ography | of ·Marcel \Proust||



These examples have been transcribed as composites of units made up of 'prehead' and 'nucleus', since that is the pattern suggested by Selkirk's assignment of prosodic structure, where the  $\phi$  units are all fairly long. However, since even a single prosodic word can constitute a phonological phrase, an alternative analysis would be as shown in the next diagram (on separate pages).

The contours in this diagram have been 'promoted' to the status of separate  $\phi$  units, whereas in the previous examples they merely formed part of larger  $\phi$  units and therefore could not be pitch-prominent. This example represents a slower, more deliberate style of speech, similar to the exaggeratedly emphatic style of the well-known phrase at the top of page 139, where '↑' represents a high-pitched unstressed syllable.



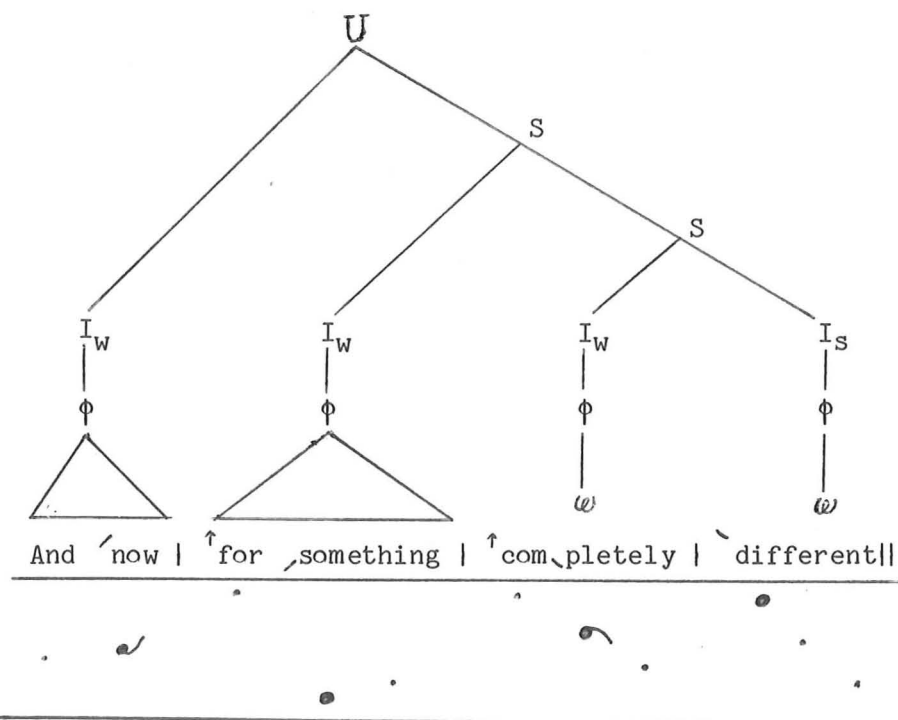
[p.t.o.]

The 'absent-minded professor' has been 'avidly' reading the latest biography of Marcel Proust.

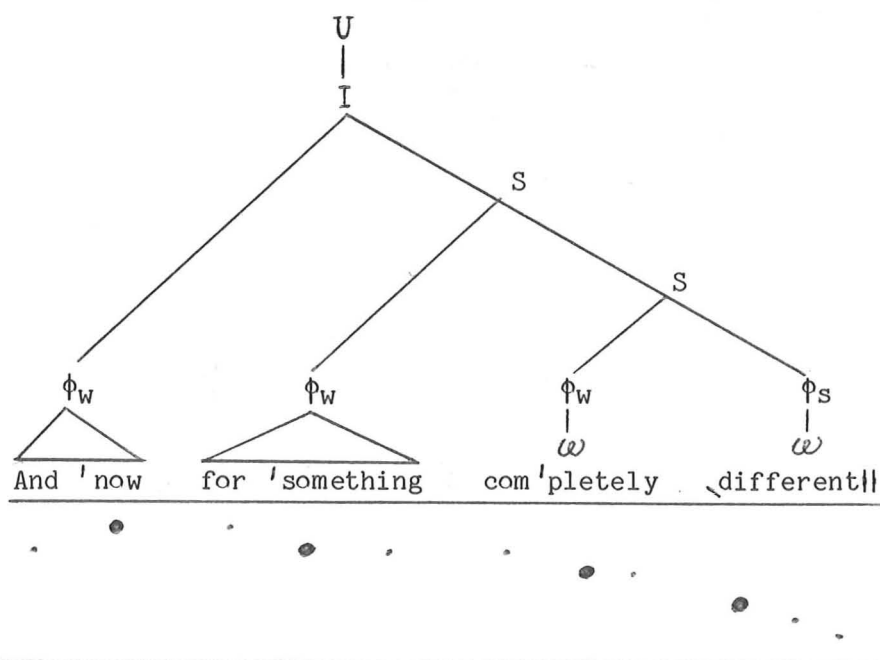


(a possible intonation pattern)

[continued from preceding page - to show effect of this phrasing]



The 'I' units in this example are even smaller than those in the previous example, since they contain only one 'φ' unit each. A more normal style of utterance might perhaps 'demote' these 'I' units to intermediate or prosodic units, giving something like the following:



In the above example, the DTE of each 'φ' corresponds to a stressed syllable in the head, except for the last 'φ', which constitutes the nucleus. The entire utterance is only one 'I' unit, therefore there is only one nuclear tone, and no perceptible pause.

This method of indicating intonational sub-units has the advantage, compared to the tonetic stress-mark system, that it is hierarchical; that is, what is realised as a head contour in one rendition of an utterance may become a nuclear tone in a different rendition of the same utterance, and only the 'metrical' system makes explicit the formal similarities between such units that enable the units to be realised as one or the other type of intonational pattern. In the tonetic stress-mark system, there is no formal similarity in definition between a head unit and a nucleus, and consequently such phenomena as 'variable phrasing' cannot be accounted for, the system being useful mainly for describing the one most common or expected type of phrasing. This is quite adequate for most purposes, and indeed is preferable in cases where all that is needed is a quick and convenient transcription system, but the proposed system seems more theoretically satisfying.

It would be possible, given a tonetic stress-mark system, to formulate a rule allowing the nucleus to be shifted to any syllable bearing a tonetic stress mark. However, this approach would be unable to account for the free use made in everyday speech of the various possibilities of phrasing, since the rule concerns only the nucleus. Also, the stress marks do not make explicit the structural similarity of the units in question, and thus such a rule would have no obvious motivation. The proposed analysis, then, while not so convenient for purposes of transcription, provides a fuller and better-motivated theoretical description.

#### 3.4.4 Building up the tree

Above the level of the prosodic word, metrical structure is defined, at the lowest level, on the basis of each word's [content] value. Words that are [0 content] are invariably weak; words that are [1 content] are nearly always weak. Words which are [3 content] are strong except in a few abnormal cases, while words that are [2 content] may be weak or strong according to context.

The basic ordering of nodes is w-s, as in the formulations of L, LP, and Selkirk for structure above the word level. The bracketing is loosely based upon syntactic surface structure wherever possible, though the metrical structure is by no means the same thing as the syntactic structure of an utterance. The main insight behind using the [content] values in this way is the view that the higher the value for [content], the more likely that word is to be strong; there are no absolute values in this approach, since 'prominence' is crucially dependent upon context. If a normally [0 content] word becomes stressed (e.g., for reasons of contrast), it then takes some positive value of [content] (usually 3) to reflect its greater semantic weight, and it is defined as a relatively strong word in the metrical tree and grid. Where the word with highest [content] value is not the last lexical word in the utterance, an anomalous bracketing order of s-w will be necessary at some point in the tree, and there will be a condition that any material following it cannot form a separate 'φ' unit, since the following syllables (even if they contain a stressed syllable) cannot be pitch-prominent. Given this condition, one can state a common definition of the nucleus that is applicable both to normal and to abnormal nucleus position: the start of the nuclear tone is located on the DTE of the last φ unit in the intonational phrase. This generalisation cannot be stated at all in the tonetic stress-mark system, which is unable to define units where only one syllable is both stressed and pitch-prominent - that is, phonological phrases. This is another area where the proposed analysis seems more satisfactory.

### 3.4.5 Selkirk's Monosyllable Rule

One of the two rules having the phonological phrase as their domain, the 'Monosyllable Rule', is described by Selkirk as follows:-

'In earlier discussions, this rule was seen as one which destressed monosyllabic words belonging to non-lexical categories, when they appeared in certain syntactic contexts. Given the theory of prosodic structure now being entertained, the rule can be given a simpler...formulation:

#### Monosyllable Rule

$$\begin{array}{c} \omega_w \\ | \\ \Sigma \quad \longrightarrow \quad \sigma_w, \text{ if } \omega \text{ dominates a non-lexical item.} \\ | \\ \sigma \end{array}$$

The rule says that a monosyllabic prosodic word  $\omega$  is 'deworded' (and hence 'defooted') if it is weak and corresponds to a non-lexical item in syntactic structure.' (Selkirk 1980b:19)

The result of applying this rule to words such as 'the', 'has', 'been', 'of', etc., is that these words may undergo vowel reduction 'and other such rules which apply only to weak syllables.' (loc. cit.) Under Selkirk's analysis, there is no way of indicating whether or not a word corresponds to a nonlexical item. Under the proposed analysis, the condition on the rule can be stated as 'if  $\omega$  is [0 content]', which amounts to the same thing but with one important consequence: by the principles stated earlier, [0 content] words can only ever be weak, and so this aspect of the node need not be specified in a rule stating only that the word would need to be [0 content]. Also, vowels which are [-full] are always part of a weak  $\sigma$  unit (syllable), while [+full] vowels forming part of a weak syllable are very likely to reduce to [-full], especially at a faster tempo. Thus Selkirk's rule, condition, and subsequent segmental process of vowel reduction can all be subsumed into one process, making use of units and features already shown to be necessary, and without having to resort to conditions peculiar to this rule. Also, Selkirk has no method of characterising reduction to schwa, or of linking this schwa with original schwa in other words; but this can be done by the feature [ $\pm$ full].

### 3.4.5.1 A reformulation of the 'Monosyllable Rule'

$$\begin{array}{c}
 \omega \\
 [0 \text{ content}] \\
 | \\
 \Sigma \quad \text{--->} \quad \sigma \text{ (or sometimes } \sigma_w) \\
 | \quad \quad \quad [-\text{full}] \quad \quad \quad [+full] \\
 \sigma
 \end{array}$$

The above is an attempt to restate Selkirk's Monosyllable Rule in the terms of the proposed analysis, as described above. The bracketed alternative realisation of the weak syllable is meant to cover the few cases where the vowel is not reduced entirely to schwa. Although one cannot state this rule in absolute terms, since not all vowels reduce to schwa under these circumstances, it is still possible to make a statement of probability, indicating that [-full] is the most likely, or 'default' value of [ $\pm$ full]. This reanalysis of the rule gives a natural explanation of the process which is not available under Selkirk's system - namely, that it applies only to words with no semantic weight [2] and that it causes reduction in prosodic status leading to segmental vowel reduction in most cases. Vowel reduction is a phenomenon that can be accounted for under this theory, but not under Selkirk's.

Selkirk also observes: 'Note that the bisyllabic about, though a  $\omega_w$  which corresponds to a non-lexical item, does not reduce [i.e. is not subject to this rule].' (Selkirk 1980b:19) Selkirk's system does not explain what it is about being a bisyllable that prevents a word from reducing when it is in a position to do so (i.e. it is a  $\omega_w$  corresponding to a nonlexical item); whereas under the proposed system, 'about', a preposition, would be marked as [1 content]. This means that it would nearly always be a weak prosodic word. Since [1 content] is very close to [0 content], there would be some tendency for the word to be reduced, being of relatively little importance semantically. If this were to happen, only the weak syllable of its bisyllabic foot would reduce to schwa.

---

[2] These are not always nonlexical words, but Selkirk's rule makes it seem so.

In Selkirk's system, words such as 'has', 'been', etc. differ only in number of syllables from words like 'about'; whereas under the proposed system, values of [content] constitute the vital difference between words which always reduce and those which may remain unreduced, when both kinds of word are weak in metrical structure.

### 3.4.5.2 Dewording

In the case of [1 content] words, reduction may be of a different kind. Whereas the 'Monosyllable Rule' refers only to [0 content] words and involves both 'defooting' and 'dewording', words which are [1 content] could simply be 'deworded' but not 'defooted' - a less severe reduction. This process would be optional, dependent on such things as tempo of utterance and speaker's intention. A tentative statement of the rule is as follows:

$$\begin{array}{ccc}
 \text{Dewording (optional)} & & \\
 \omega_w & & \\
 [1 \text{ content}] & \longrightarrow & \Sigma_w \\
 | & & [1 \text{ content}] \\
 \Sigma & & 
 \end{array}$$

It will be noted that this rule applies only to words that are weak and composed of only one foot; in this much it is identical to the Monosyllable Rule. The foot may contain one, two, or three syllables, and the word is [1 content]. The rule is more likely to apply, the fewer the number of syllables in the foot. The rule is applicable to polysyllables since it retains the unit of the foot, within which the prominence relations are usually in the order s-w. Since the rule concerns only words containing a single foot, this means that within the  $\omega_w$  units in question, any prominence relations will mostly be in the order s-w. Above the  $\omega_w$ , they will nearly always be in the order w-s. Therefore the foot cannot be dispensed with as well as the word, in polysyllables, since there is a need for a unit to mark the point where the basis for assigning 'w' and 's' to nodes is changed. The point does not arise for monosyllables, which can therefore be 'defooted' as well as 'deworded'.

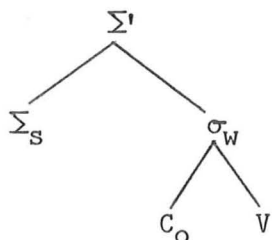
### 3.4.5.3 Dewording needed for [1 content] words

If these rules are viewed as mere formalism, it may be objected that it is equally possible, in [1 content] words, for 'defooting' alone to take place, rather than 'dewording', since there would still be a unit (the single-footed word) to mark the boundary between levels having different principles of 'w' and 's' assignment. But such an objection would ignore the hierarchical nature of the units, whereby a  $\omega$  must contain at least one  $\Sigma$ , and a  $\Sigma$  must contain one or more  $\sigma$  units (underlyingly, at least). Even after the application of either of the rules described above, the newly-reduced unit would not constitute the sole daughter of a unit two 'levels' higher than itself, since it will always constitute the 'w' member of a w-s pair. Thus the hierarchical relationship of the units is preserved. This would not be the case were the rule for [1 content] words one of 'defooting' rather than 'dewording', since one or more syllables could then be directly dominated by a  $\omega$  unit with no intervening  $\Sigma$  unit. Such a position would hardly be logically tenable, since a single-footed word 'is a' foot, by definition, and merely erasing the  $\Sigma$  symbol in the diagrammatic representation will not alter the fact that such a word has the internal structure of a stress foot and therefore is a stress foot, as defined by one of Selkirk's 'basic foot templates'. The same objection does not hold in cases where the  $\omega_w$  unit has been 'defooted' or 'deworded' according to the rules proposed, since then one of two considerations will apply:-

- i) the defooted syllable may end in a single non-tense vowel with no following consonant. Thus it cannot constitute a foot by any of Selkirk's 'foot templates', its original designation as such being due to the consideration of the syllable in its citation form only, when the vowel would be lengthened. The process of reduction of a word to a  $\sigma_w$  unit in an utterance where the word is a  $\omega_w$  unit corresponds to the effect on that word of being in an utterance together with many other words, with at least one of which it must combine in order to be designated a 'weak' unit; i.e. weak in relation to another unit.

- ii) even in the case of syllables capable of constituting a foot on their own account (according to Selkirk's basic foot templates), one finds the same restriction as with the syllables covered by condition i). That is, such syllables (or, in the case of dewording, feet capable of constituting a word on their own account) are never the sole daughters of a higher-level unit. They always combine with a unit of intermediate level (a  $\omega$  or a  $\phi$ , respectively) which itself is capable of being the sole daughter of the higher-level unit that dominates them both. Something like this principle is seen at work in Selkirk's template for the 'stress super foot' (where the  $\Sigma_s$  units may have varying internal structure):

stress super foot



In this template, the final  $\sigma$  unit is unable to form a new  $\Sigma$  on its own, since it contains a non-tense vowel not followed by a consonant. It is accordingly incorporated into a higher-level unit together with a foot, and must always be the weak member. The  $\Sigma_s$  unit, on the other hand, is not only a foot in its own right, but is also capable of constituting the sole daughter of a  $\omega$  unit. This situation is parallel to that described above, though there is no direct equivalence because of the recognition here of a 'super foot' unit, with no analogue at levels above that of the word.

### 3.4.5.4 Disyllabic [0 content] words

A minor problem is posed by [0 content] words (Thompson's -SALIENT) which are disyllabic, eg, 'because', 'become'. These will not be deworded and defooted by the rule for [1 content] words. Yet it is possible, in fast speech, for each of the syllables of the above examples to be reduced to schwa, exactly as if a reduction rule of some sort had applied. Thompson, in his summary of features, lists conjunctions and copula verbs under -SALIENT, but adds the following as a footnote; 'Simple one syllable conjunctions such as and or but are clearly -SALIENT. I am not sure about others like since or because etc.; more data is needed before a sure determination can be made.' (Thompson 1980:138) It is possible that a disyllabic [0 content] word such as 'because' would undergo the dewording rule, or alternatively it could be seen as being in fact [1 content]. The latter view seems preferable for two reasons:

First, 'because' could be said to have a certain degree of inherent semantic weight in that, like prepositions, it refers to relations between real-world entities (in this case, logical relations), although it makes no direct reference to real-world entities. Thompson's uncertainty about the status of 'because' etc. with respect to  $\pm$ SALIENT could stem from a native-speaker perception of this sort of intermediate status, which is akin to that of prepositions and negatives.

Secondly, it is by no means unknown for such conjunctions to be stressed in speech, and they are stressed more often than would be expected if this were merely a case of 'contrastive' or 'emphatic' stress. On the other hand, similarly to prepositions, when stressed they hint at a higher level of emphaticness overall in the utterance, however casually spoken, and never seem to achieve the complete lack of emphaticness found in a stressed lexical word.

For these reasons, it seems best to designate 'because', 'become', etc. as [1 content], which makes them eligible for dewording. However, this is as yet only a tentative suggestion, and it may well be found that there is a better way of dealing with the problem.

### 3.4.6 Examples

On page 149 follows a fairly long example of the workings of the proposed analysis, the example being adapted from Selkirk 1978:17. The circled units indicate the points of application of the reduction rules, as described earlier. This is only one of the possible phrasings of the sentence, and other divisions into higher-level prosodic units are also possible. This phrasing could be realised by the following intonational pattern:-

The 'absent-minded pro,fessor | has been .avidly 'reading

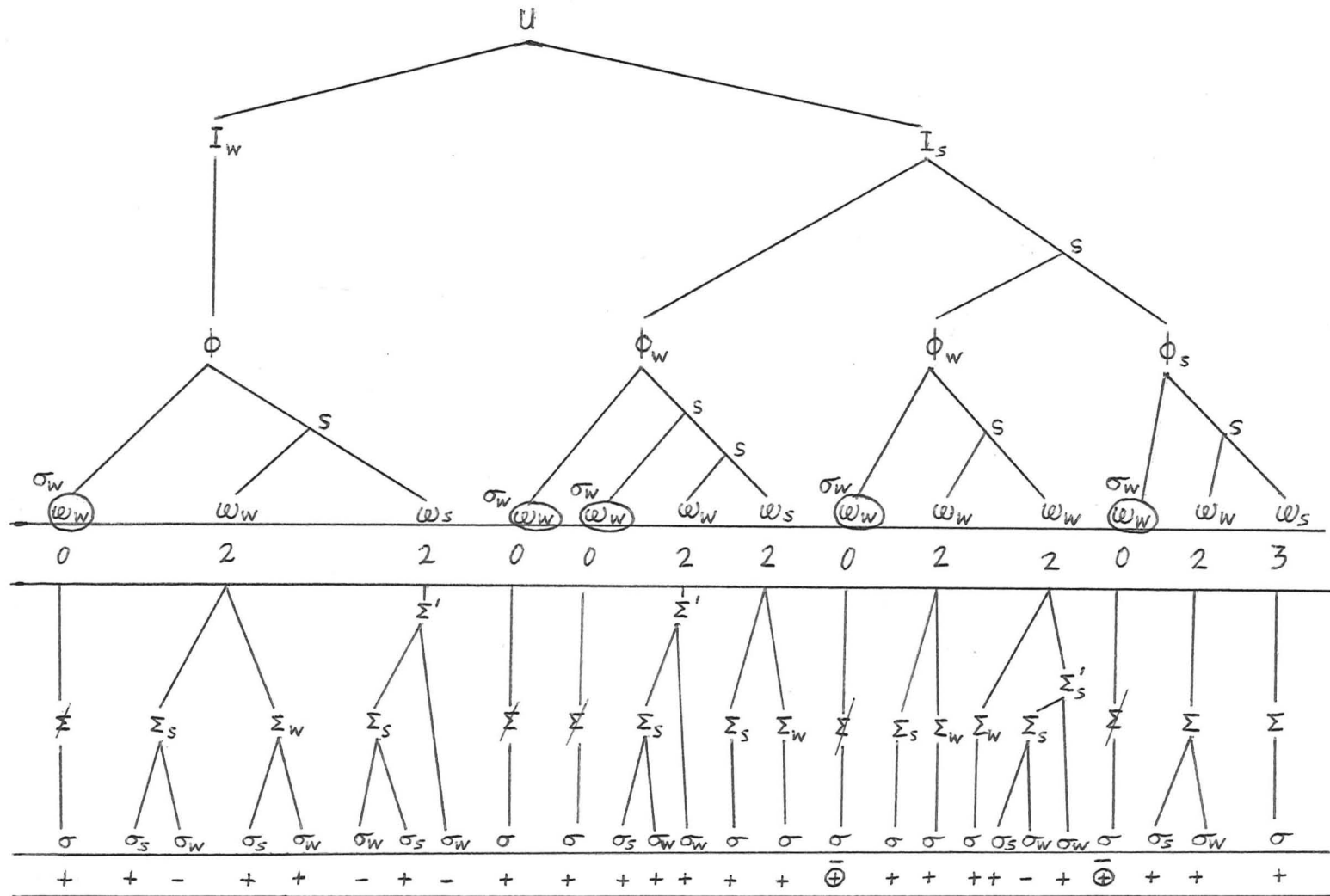


the .latest bi'ography of 'Marcel \Proust||

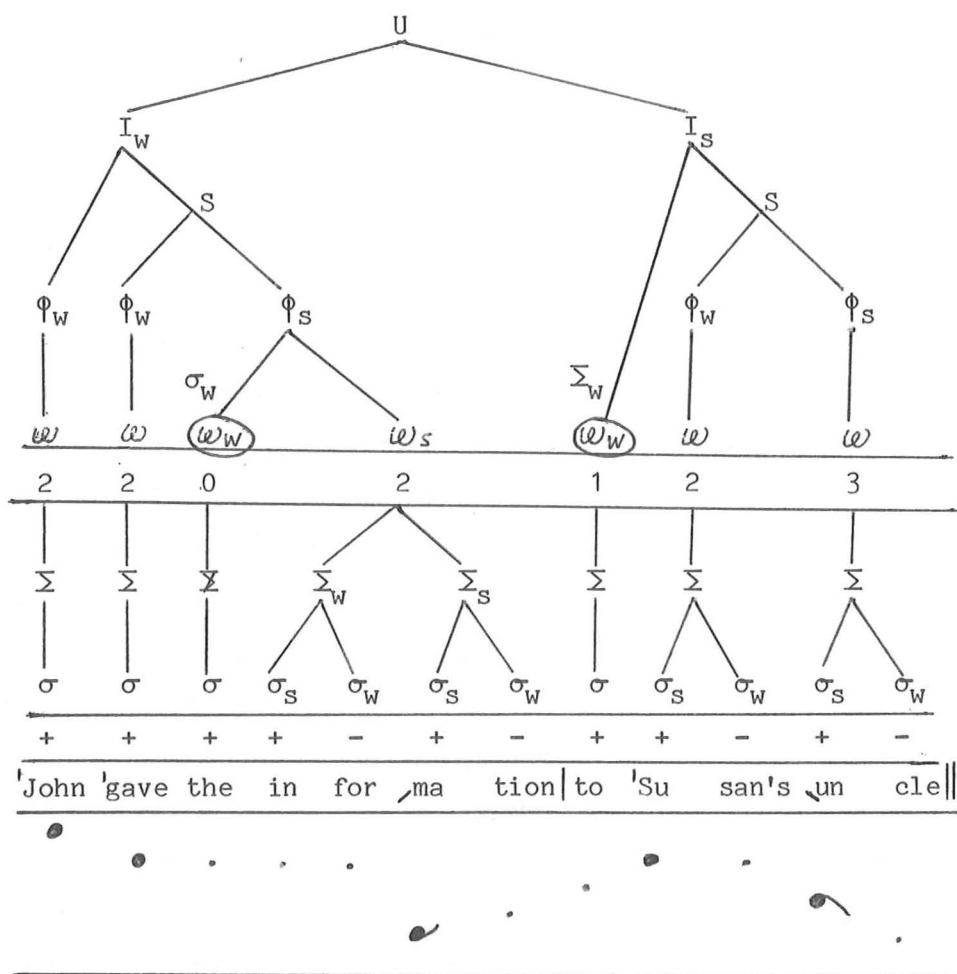


This phrasing corresponds to a rather quick and unemphatic style of utterance; a slower style would involve 'promoting' more  $\phi$  units to the status of intonational phrase, or 'minor tone-unit'. On page 150 is given the metrical structure, and a suggested phrasing, of the example sentence 'John gave the information to Susan's uncle'. Whatever the phrasing actually chosen by the speaker, the basic s-w relations will be the same, so the overall structure does not alter.

As in the previous example, the circled units indicate the points of application of the reduction rules, and the intonation pattern is only one of several patterns that would have been possible given this phrasing (i.e. this precise assignment of 'I' and ' $\phi$ ' labels to nodes higher than the word). The process of dewording and of defooting and dewording will be seen to be important later, in connexion with the 'metrical grid'.



The absent-minded professor has been avidly reading the latest biography of Marcel Proust



### 3.5. THE METRICAL GRID

#### 3.5.1 Introduction

The 'metrical grid' is an entirely abstract concept, first introduced in Liberman (1975), and elaborated upon in Liberman and Prince (1977). It is a method of resolving the purely hierarchical structure of the metrical tree into a time-bound, serial, structure, while at the same time preserving relative prominence of units at various levels. Although Liberman, in his original formulation, used numbers as place-holders in the metrical grid, in order to be able to refer to each position explicitly, it seems more logical to follow Thompson (1980) and use a constant symbol as a place-holder. Thompson uses an asterisk, I will use a circle, but the precise nature of the symbol is of no importance. The reason for using a constant symbol is that it makes explicit the important feature that all these grid positions are identical in kind, and differ only in degree, according to the level where they appear.

As Liberman and Prince write, '...the "degree of stress" of a given syllable is represented by the height of the column of marks that stands over it.' (LP:312) Also: 'Elements are metrically adjacent if they are on the same level and no other elements of that level intervene between them; adjacent elements are metrically alternating if, in the next lower level, the elements corresponding to them (if any) are not adjacent; adjacent elements are metrically clashing if their counterparts one level down are adjacent.' (LP:314)

This framework allows LP to describe the operation of the Rhythm Rule, which resolves clashing stresses by swapping a pair of w/s nodes, so that the stress of the first word is shifted from the final (or penultimate) syllable to a previous syllable in the word. The metrical grid provides a clear formalisation of the prominence relations at various levels, and from it one may pick out the stressed (and secondary-stressed) syllables,

leaving matters of pitch-prominence to the intonational/metrical tree. LP argue for the metrical grid as an entity distinct from the metrical tree, writing;

'Representing such grids as trees, although possible, requires us to define rows and columns derivatively, and also requires the imposition of constituent-structure relations that will have no relevance to our present purposes. Therefore, we choose to formalize such grids as, in a sense, hierarchies of intersecting periodicities.' (LP:313)

The same general principles apply to the metrical grid as were discussed in connexion with the semantic grid in section 3.2. That is, the grid is not simply a scalar feature written in graphical terms, since the prominence relations within each level are of great importance. Also, a word defined as unstressed in the metrical grid of a long sentence may, in a shorter sentence where the word has the same number of 'place holders', become stressed due to the relative lack of stress of the other words.

### 3.5.2 Description

In the first row of the metrical grid (under the analysis proposed here), every syllable containing a [+full] vowel receives a place-holder, while syllables containing a [-full] vowel may be indicated by a dash. This is a way of representing the fact that not only do schwa and syllabic consonants not receive stress in this sentence, but they are never likely to, in this or any other sentence. This is in contrast to syllables which may or may not receive at least a certain amount of stress, depending on context.

In the second row of the metrical grid, a place-holder is assigned to all syllables that are the DTE of a stress foot - that is, to all monosyllabic feet, and syllables of the ' $\sigma_s$ ' kind.

In the third row, another place-holder is assigned to all syllables that are the DTE of a prosodic word.

In the fourth row, a place-holder is assigned to all syllables that are the DTE of a phonological phrase: in the fifth row, to all DTE's of an intonational phrase: in the sixth row, to the DTE of the overall utterance. There is a general constraint on the grid, following from the hierarchical nature of prosodic units, such that a place-holder may not be assigned at any level higher than the first without the syllable also having a place-holder on the previous level.

Liberman and Prince do not give any details as to the formulation of their metrical grid in these terms, and Selkirk does not even use the notion of a metrical grid, dealing only with tree structure. The method presented here has the advantage that it will cause different phrasings of the same sentence to be reflected differently in the metrical grid, to a minor degree. However, the grid will preserve the basic distinction between 'stressed' and 'unstressed' which it is intended to represent. 'Stressed' syllables are those that are associated with a longer column of place-holders than are neighbouring syllables. Particularly important for such comparison is the preceding syllable, rather than the following syllable. If the preceding syllable is associated with a 1-column in the metrical grid, and the syllable in question is associated with a 3-column or sometimes even with a 2-column, then the syllable in question is stressed. However, if the preceding syllable has a 4-column, and the syllable in question a 2-column or often a 3-column, then this syllable is not stressed. What is important is not any 'inherent stress' the syllable may have, but its stress level relative to the neighbouring syllables, so that a column of a given height may be interpreted as stressed or unstressed, according to context.

Also important in connexion with the metrical grid is the notion of probability, as against certainty of occurrence. For example; if a 2-column is flanked by 1-columns, then it is probable that its syllable will be counted as stressed. However, this is not necessarily so, since the preceding 1-column may be itself preceded by a 4-column, which is even more likely to be stressed than the 2-column. Given that this 4-column is stressed, the 2-column is fairly likely to be stressed, but need not be - the question will be decided by outside factors, such as tempo of

utterance, pitch-prominence patterns, and any segmental modifications. [3]

This is the phenomenon known as 'secondary stress'; when a syllable is said to have a certain degree of stress, but less than the full amount possible. There is sometimes disagreement over whether such a syllable is stressed at all, in cases where it is not also pitch-prominent. The examples in the next section may help to dispel some of this confusion.

### 3.5.3 Use of the metrical grid

#### 3.5.3.1 Assigning [±stress]

The question of whether or not a syllable is stressed could be most easily approached by means of an algorithm similar to that proposed for the 'semantic grid' (see section 3.4.2.2.). This algorithm is almost identical to that proposed by Thompson for identifying the first syllable of each 'foot' (Thompson 1980:140), and is as follows;

At the start of the utterance set the threshold to 2. If the first column is of height 2 or greater, that syllable is stressed. If it is not, subtract one from the threshold and move to the next syllable. If its column is greater than (or sometimes even equal to) the threshold in height, that syllable is stressed; otherwise it is unstressed. On designating any syllable as stressed, the threshold is reset to four before moving on to the next syllable. The threshold value will seldom be less than nought, as long stretches of completely unstressed syllables are uncommon in English. If a column is equal to the threshold value in height, it is fairly likely to be stressed, especially when followed by one or more very low columns. However, it is impossible to be certain on this point, since factors such as tempo and speaker idiosyncracies affect it.

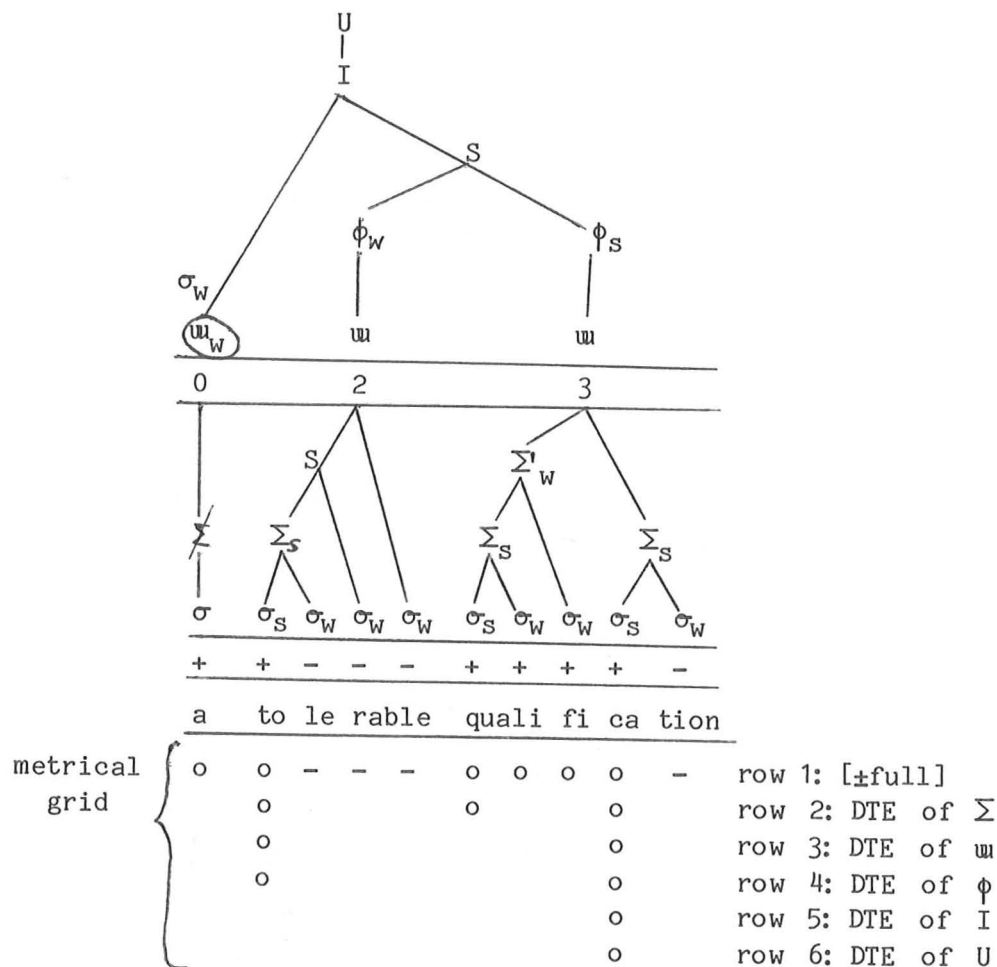
---

[3] This description looks at the process from the point of view of the 'hearer'; one could also take the 'speaker's' point of view and say that if it was intended that this 2-column be stressed, the syllable will show modification of pitch pattern and segmental features, and the overall tempo of the utterance will be slower. The phenomenon is not intended to be viewed as a directional process, but these features are to be seen as co-occurring with a stressed syllable.

An increase in tempo could be represented in the algorithm by an increase in the reset value of the threshold, so that fewer syllables will be counted as stressed. Conversely, a slowing-down in tempo (usually corresponding to greater emphaticness in style of utterance) would be represented by a small decrease in the threshold's reset value, so that it would be possible to count more syllables as stressed.

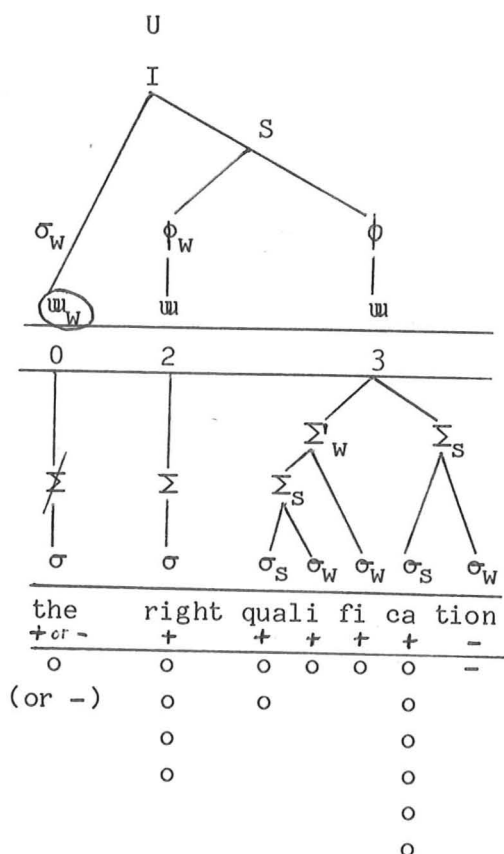
### 3.5.3.2 Examples

i)



In example i), the first syllable of 'qualification' could be said to have 'secondary stress', while the fourth syllable has primary stress. The absence of other 2-columns for some distance on either side of 'qual-' leads to its being counted as stressed. This is for want of a 'better' syllable to take the stress that is needed to break up such a long stretch of comparatively non-prominent syllables. It would be rare - and rather difficult - to produce this sentence with 'qual-' as one of a string of six unstressed syllables.

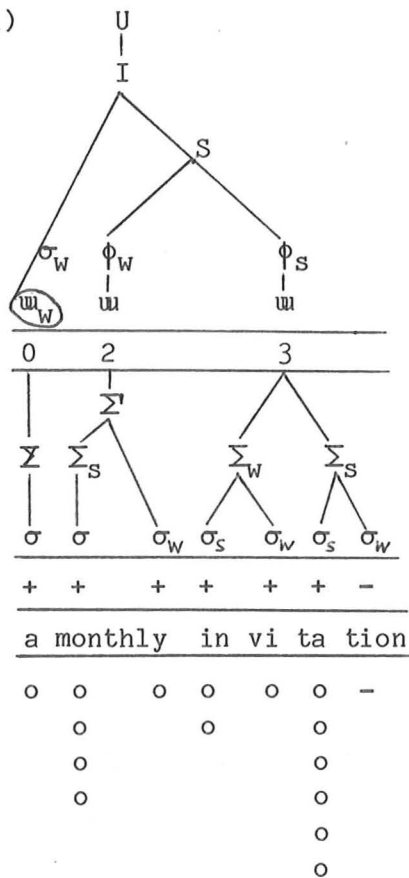
ii)



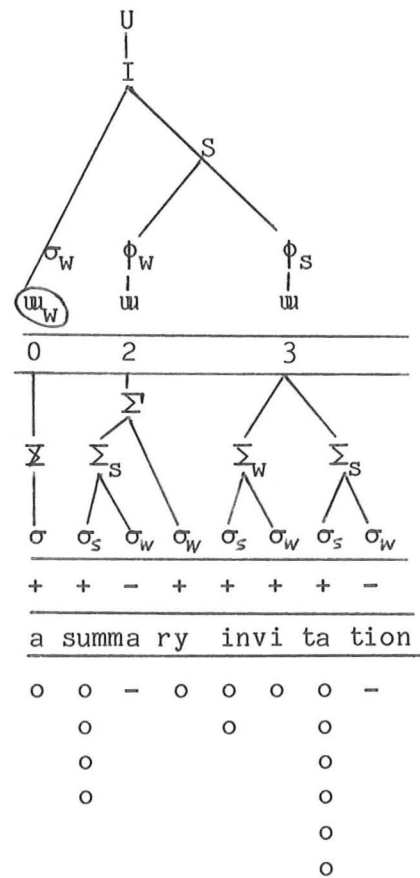
In example ii), the syllable 'qual-' is again associated with a 2-column, but this time it is immediately preceded by a 4-column. A 4-column corresponds to a stressed syllable in all but the most exceptional cases. Thus the stress on 'qual-' is now not nearly as 'prominent' as it was in example i), and it is quite likely that it will count as completely unstressed. This is despite the fact that, as in i), it is a 2-column. If the utterance is spoken slowly, it is possible for 'qual-' to be stressed. In such a case, it will cause a stress clash with 'right' that cannot be resolved by rhythm rule (i.e., backshifting of stress on the first word of the pair), because of the lack of other syllables to which to shift the stress.

In example iii), given a reset value of four for the threshold, the syllable 'in-' would not be counted as stressed. In example iv), the column for 'in-' is equal to the threshold, and so could well be stressed. In example v), the value of the threshold at 'qua-' is equal to that for 'in-' in example iii). In example v), however, 'qua-' is followed by two 1-column syllables (as against one for 'in-'), and so is slightly more likely

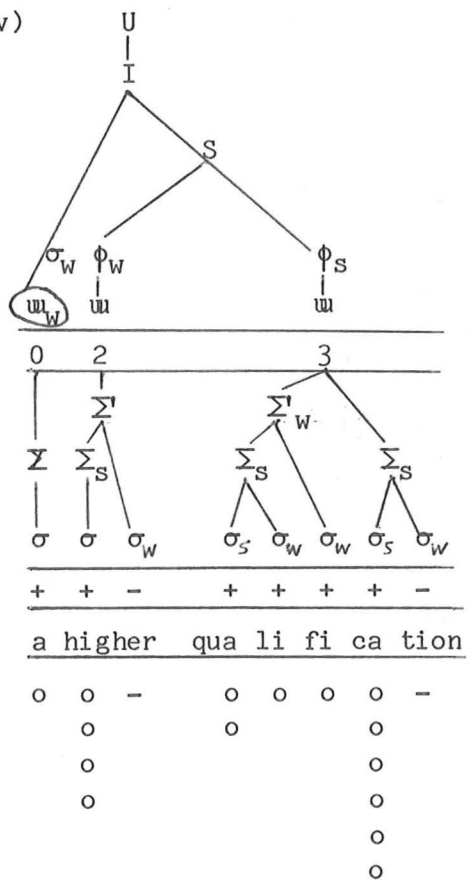
iii)



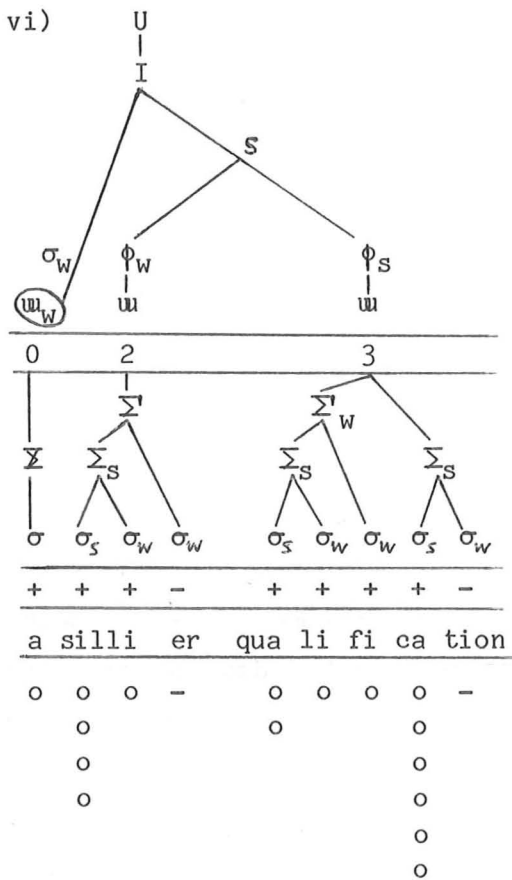
iv)



v)

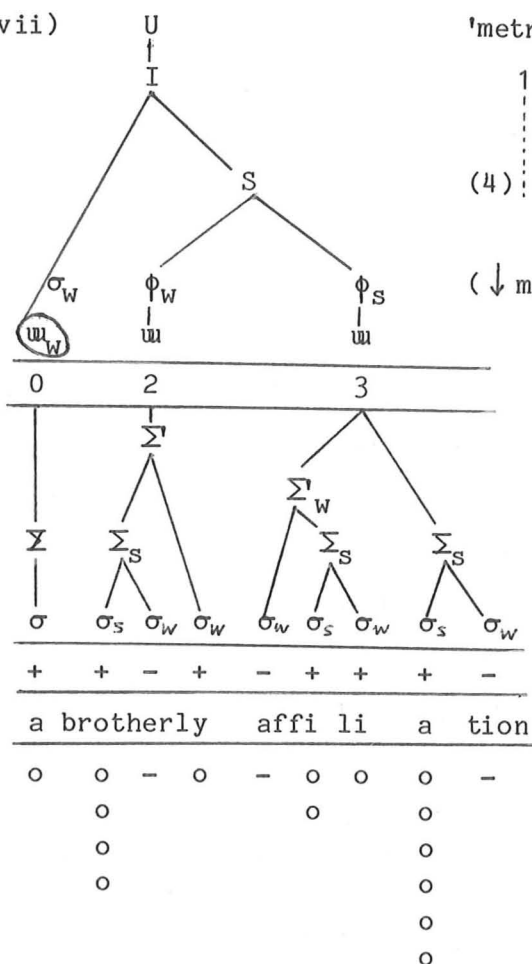


vi)

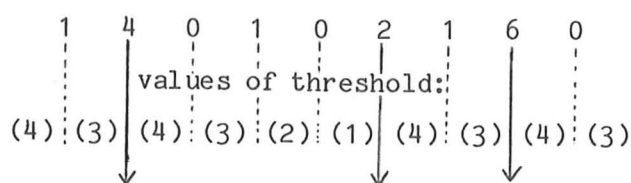


to be counted as stressed than is the syllable 'in-' in example iii). In example vi), the value of the threshold at 'qua-' is equal to that for 'in-' in example iv) - that is, equal to the 2-column on 'qua-' and 'in-'. Since 'qua-' is followed by two 1-column syllables (as against one for 'in-'), it is even more likely that 'qua-' will count as stressed than 'in-', though 'in-' is nevertheless quite likely to count as stressed.

vii)



'metrical grid' in numerical values:



(↓ marks syllables counted as stressed)

In example vii), the 2-column on '-fi-' is greater than the threshold at that point, and so this syllable is likely to be counted as stressed.

It will be noted that the algorithm reduces the variable values of the columns of the metrical grid to plus or minus values of [ $\pm$ stress], represented in interlinear graphs by large and small dots respectively. This is the point at which an essentially multi-valued feature is reduced to a binary feature for the purposes of an actual utterance. While theory may require sophisticated descriptive apparatus, the more pragmatic requirements of practice call for clear distinctions. Since the maximal

form of distinction holds between just two elements, it is perhaps not surprising that the binary distinction between stressed and unstressed is the one used in practice.

### 3.5.4 The need for the metrical grid

Although the columns of the metrical grid are being treated as simple numerical values for the purposes of this algorithm, this is not meant to imply that the process of assigning place-holders at various levels is redundant. Far from it: the metrical grid, formulated in terms of levels as well as in terms of the columns necessary for the algorithm, is vital for the specification of such things as stress clashing and the rhythm rule (which relieves stress clashes where possible). It is also important in that it preserves the prominence relations of units embedded inside higher-level units. Perhaps most tellingly of all, it is not clear how the numerical values would be arrived at independently of the grid. It is true that a fairly crude algorithm has been suggested (without conviction) by Liberman and Prince, but this is unsatisfactory, as will be seen. The algorithm suggested by LP is as follows, where the number 1 would correspond to the highest degree of stress:

'If we wished to mimic closely the numerology of previous theories, we could make use of the following definition:

- (12) If a terminal node t is labelled w, its stress number is equal to the number of nodes that dominate it, plus one. If a terminal node t is labelled s, its stress number is equal to the number of nodes that dominate the lowest w dominating t, plus one.' (LP:259)

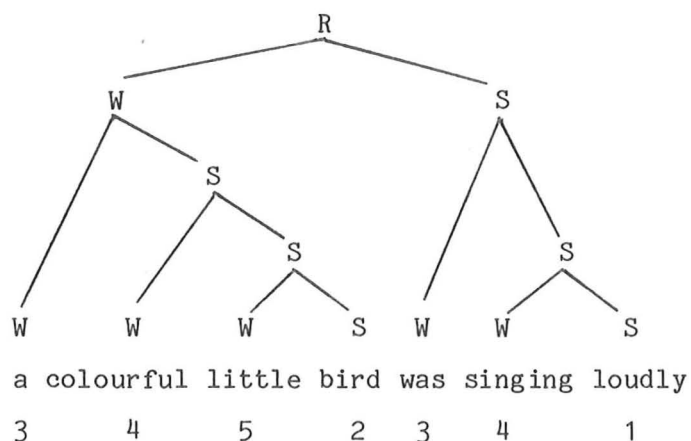
Liberman and Prince point out that this algorithm

'simply expresses the fact that the information that goes into the results of the numerological cycle is also present in the annotated tree of the relational theory. There is nothing inherent in the relational method of representation that would lead one to the particular rank-ordering of terminals implied by (12), as expressing the notion "degree of stress". The relational representation says, simply, that at a given level one subconstituent is stronger than the other.' (loc. cit.)

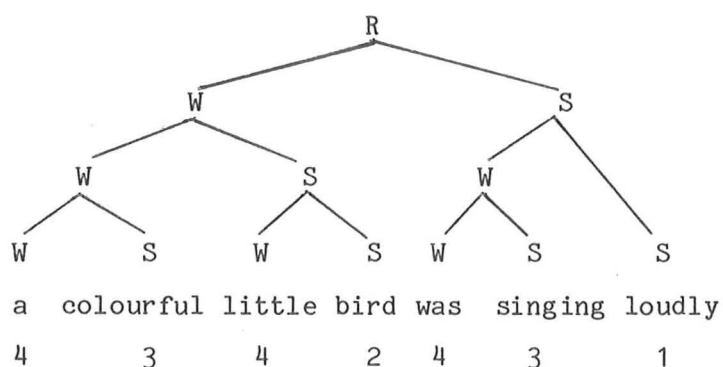
However, they do not mention that not only is this algorithm unnecessary, but it can also lead to incorrect results, due to 'stressed'

words being embedded more deeply in the metrical structure than are unstressed words, as is seen in the following example;-

i)

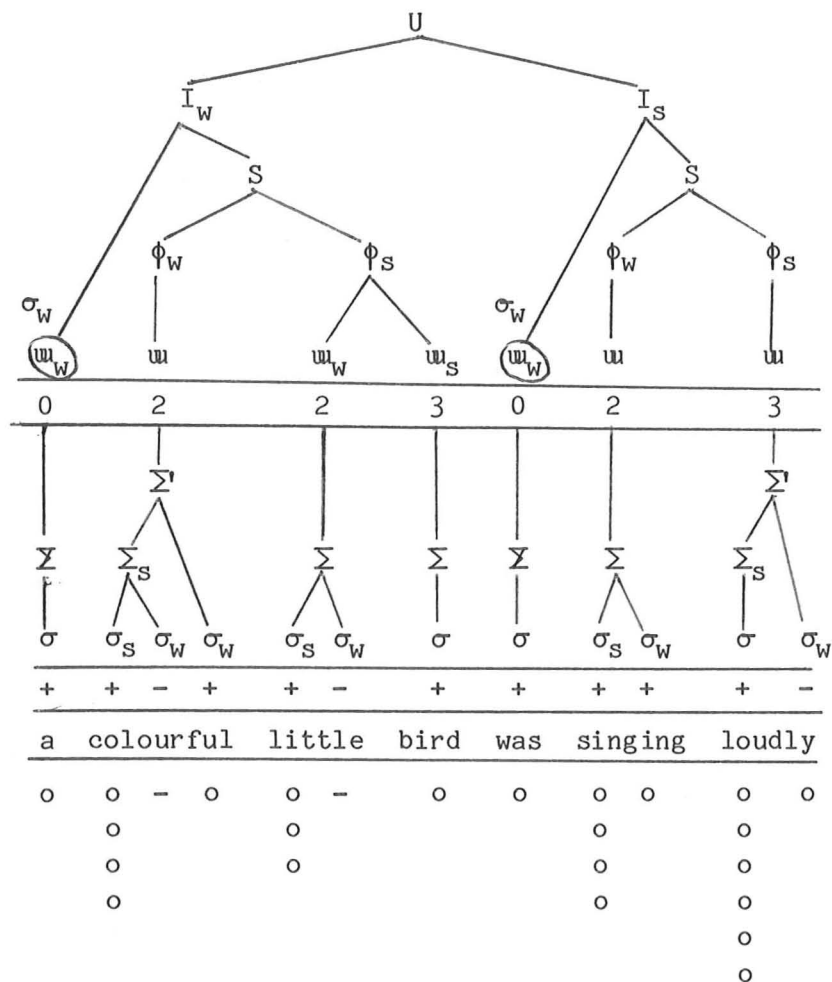


ii)



These examples are drawn out in the style of LP, though without marking values for their feature [ $\pm$ stress]. The numbers are derived by the algorithm quoted earlier. Example ii) has undergone rebracketing by a process which L sees as akin to 'cliticization' (though no word boundaries are deleted). This is the process by which 'certain non-lexical words become attached to lexical words in their vicinity' (L:110). Even when this adjustment has been made, however, the result is not totally satisfactory. Admittedly, example ii) is more satisfactory than example i) (the same sentence before rebracketing), where nonlexical, normally unstressed words 'a' and 'was' seem to have more stress than lexical words 'colourful', 'little', and 'singing'. However, the adjusted structure of ii) still assigns the same stress level to the two nonlexical words as to the lexical word 'little', which usually has at least a certain degree of stress. No explanation is offered for this effect, and so it seems LP are wise to reject this type of algorithm. The following example shows the same sentence analysed according to the system proposed here:

iii)



a possible style of utterance would be:

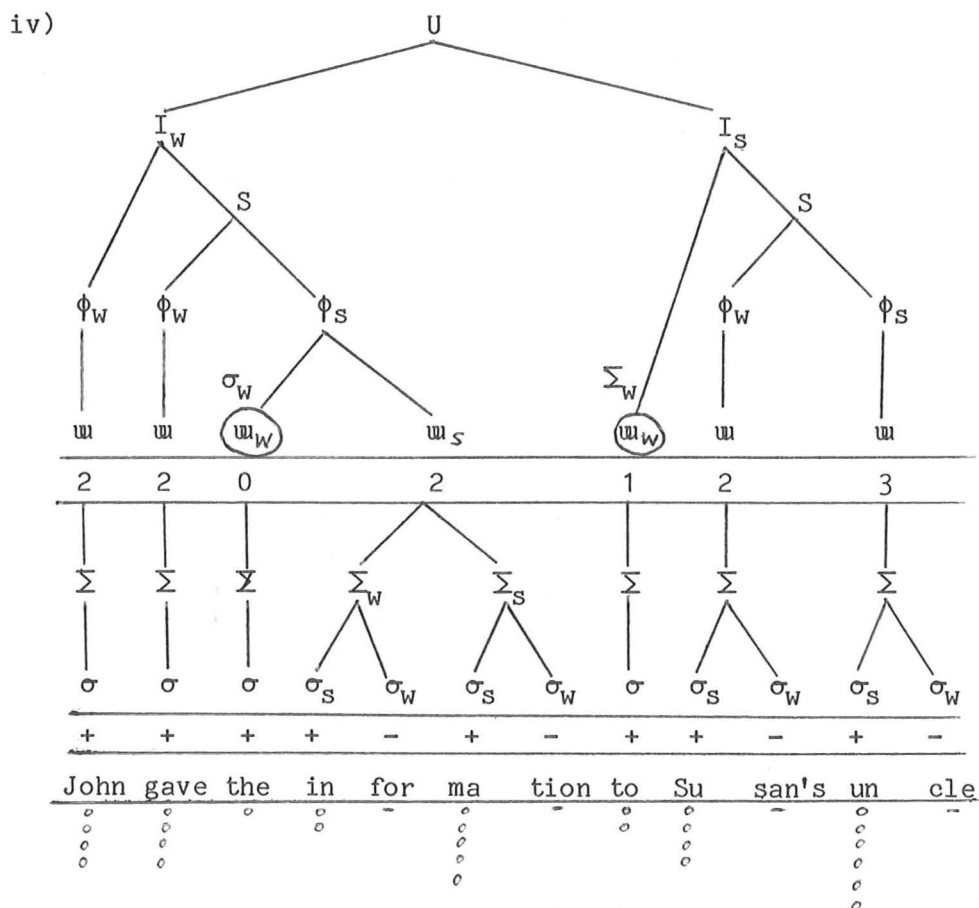
a 'colourful ·little ,bird | was 'singing \loudly ||

The metrical grid is therefore a vital part of the total representation. It provides the basis for defining stressed syllables and stress clashes. The collapsing of the many different column heights into plus and minus values of [ $\pm$ stress] is carried out by a simple algorithm that is, however, flexible enough to be able to account for the results of changes in overall tempo. For all practical purposes, in the actual use of language, syllables are either stressed or unstressed, and so such a 'collapsing' process is necessary. However, there are cases when it is unclear whether or not a particular syllable is stressed; such cases can be characterised as occasions when either:

i) the column height is merely equal to the threshold value at that point, or

ii) if the column height is (a little) greater than the threshold value at that point, there is a maximum of one column of height one or nil after the syllable and immediately before a column higher than that of the given syllable (see end of section 3.5.3.).

The complete representation is a composite of semantic, syntactic, metrical, and segmental phonological considerations. As such, it is an attempt to give a fuller account of stress in English than is possible using only one or two of these aspects: semantic (Thompson), syntactic (Chomsky and Halle, and many others), metrical (L, LP, and Selkirk), segmental phonological (Chomsky and Halle: [ $\pm$ stress] assigned to vowels). The following is an example of a complete representation.



a possible intonation pattern for this phrasing is:

'John 'gave the ·infor,mation | to 'Susan's ,uncle ||

The metrical grid is then converted to values of [ $\pm$ stress] for each syllable, using the algorithm described previously. This algorithm is almost identical to Thompson's, which runs as follows:-

'Set the threshold at 2. Proceeding left to right through the constituent grid, at each syllable

1) if the associated column is a 4-column, start a foot with this syllable, set the threshold at 4, and carry on.

2) otherwise if the height of the column is greater than or equal to the threshold, and there is a next column whose height is less than the height of this one, either start a foot with this syllable, set the threshold at 4 and carry on, or else decrease the threshold by 1 and carry on. Starting a foot is preferred if the column height is greater than the threshold.

3) otherwise decrease the threshold by one and carry on.'  
(Thompson 1980:140)

Thompson, however, does not mention the possibility of adjusting the reset value to correspond to changes in tempo. Neither does he mention the situation where the smaller column after the given syllable precedes a column larger than that of the given syllable. In such a case, it is very likely that the given syllable will not count as stressed, especially if the overall tempo is fairly fast, corresponding to a higher reset value of the threshold.

### 3.5.5 Stress shifting

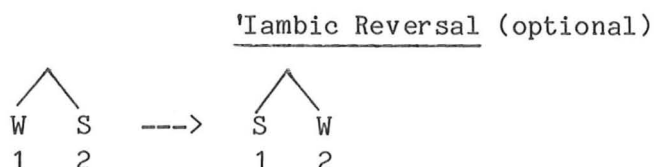
Stress shifting, or the 'thirteen men' phenomenon as it is sometimes called, is a very common feature of English. Many examples of different kinds are given in Bolinger (1965). A few examples follow:

clear-cút	a cleár-cut problem
Constitútíon	the Cónstitution forbids it
cast asíde	then cást aside your fears
he struck out	he strúck out twice
(he was outráged!	an outräge)
thirteeén	thírteen men

Bolinger points out that this backshifting is general among all 'two-word verbs' as he calls them - i.e. verb with post-position. But it is seen also in nouns and adjectives of various kinds, and so does not seem to

be tied directly to syntactic considerations.

According to Bolinger, these shifts '....apparently can only occur when the syllable to which the shift takes place is already long.' (Bolinger 1965:160) This agrees with condition number 2 of LP's rule of 'Iambic Reversal', formulated to avoid clashing stresses (LP:319):



- Conditions:
1. Constituent 2 does not contain the designated terminal element of an intonational phrase.
  2. Constituent 1 is not an unstressed syllable.'

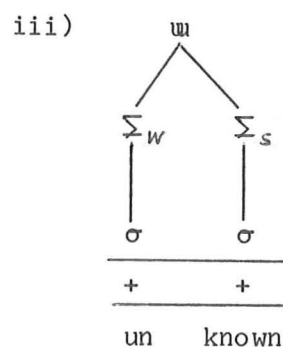
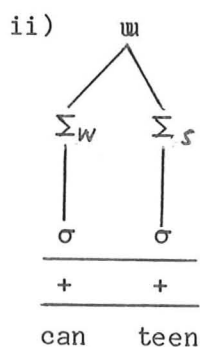
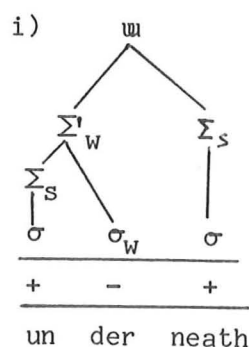
This condition reflects the general constraint that no 'S' node may directly dominate a [-stress] element in Liberman's formulation. Under the proposed system, the constraint would be that no [-full] vowel may be the DTE of a stress foot: i.e. it will only be a ' $\sigma_w$ ' unit.

Bolinger's conclusion has significant implications for the initial assignment of metrical structure. He writes;-

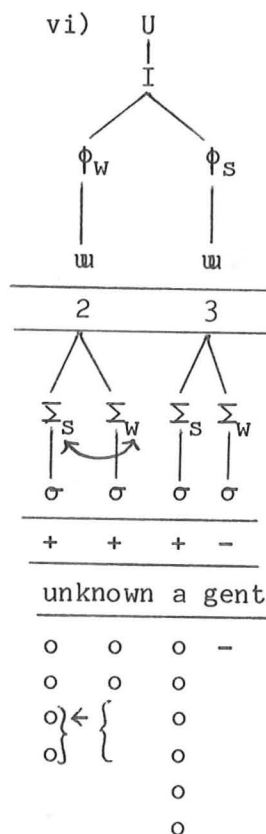
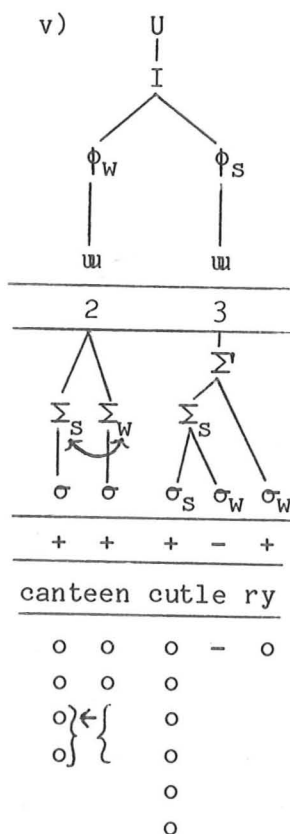
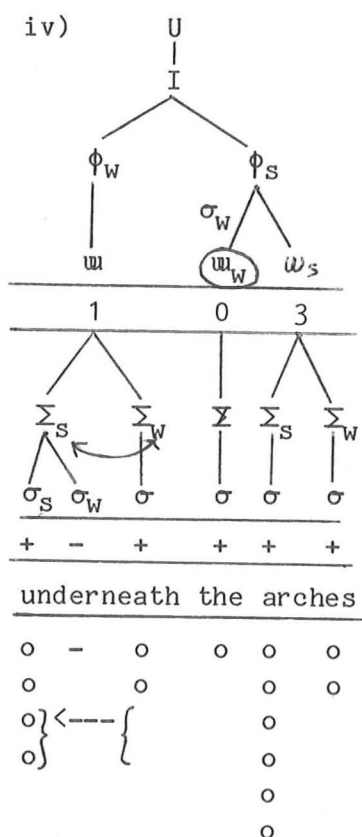
'stress patterns are apt to congeal in the way position in the sentence may predispose them, which means that nouns and adjectives will tend to be forestressed, verbs to be endstressed, and ultimately, in many cases, to carry this stress with them regardless of position. The influence of the major sentence accents in crystallizing the stress in the first place has been overlooked mainly because of our reliance on citation forms, which has fixed in our minds the conviction that we ought to be able to find 'the' stress pattern of every lexeme. Since citation forms coincide, in their stresses, with what we find at the END of an utterance, the almost equally potent initial accent has not been given its due.' (Bolinger 1965:164)

For some words at least, it could be a fallacy to assume that there is only one stress pattern possible. The citation form of a word, as Bolinger observes, provides the motivation for the convenient fiction that the word is stressed in a particular way, which subsequently must be altered by rhythm rule in very many cases. The citation forms of some words, in fact, would contravene Selkirk's rules of prominence relations within the word (see end of section 3.2.2.). According to her rules, given a pair of sister

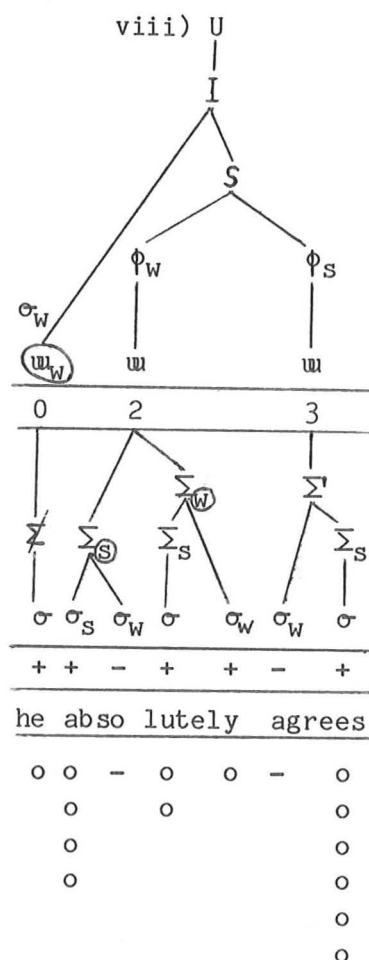
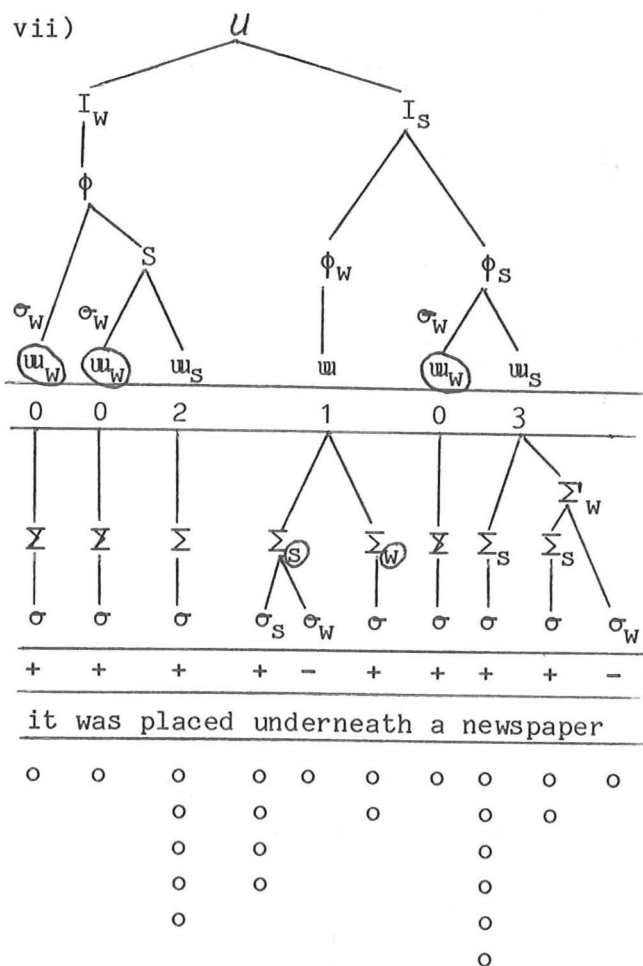
nodes, the second is strong if and only if it branches. This rule is broken in the citation forms of words such as the following;



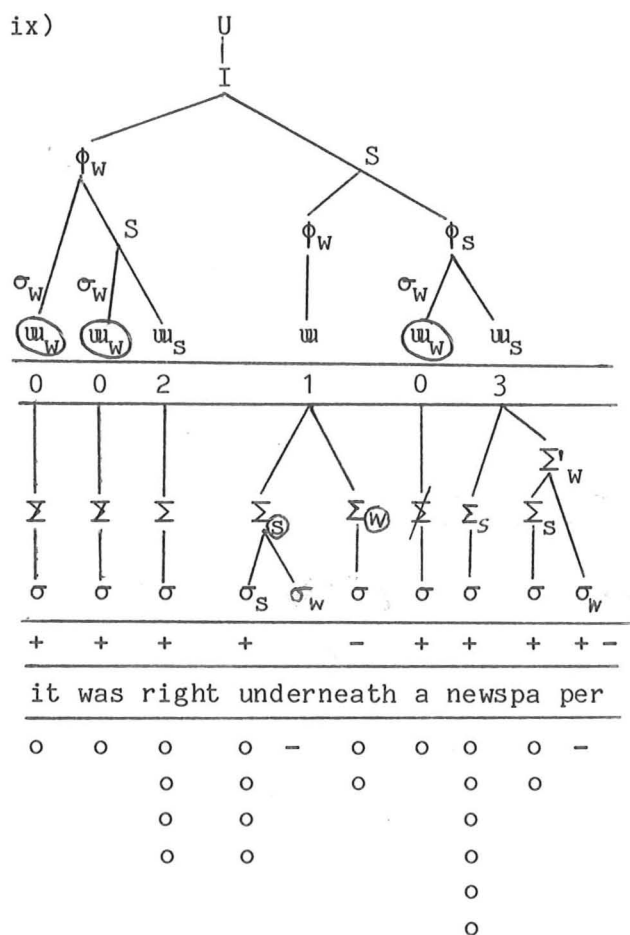
However, in context, the rhythm rule operates to backshift the stress in words such as this, where a preceding syllable is capable of taking stress (i.e. it is [+full]). In such cases, the shifted form conforms to Selkirk's rule, which is essentially the 'Lexical Category Prominence Rule' of Liberman and Prince. In the following examples, the effects of backshifting are shown by arrows.



It appears that backshifting of stress can be optional in some cases, such as those where one or more syllables intervene between the (unshifted) stressed syllable of a word and a following stressed syllable. It is possible for such cases to be produced without backshifting, but they then sound somewhat awkward and unnatural, particularly if there are only a few intervening syllables. In nearly every case the stress undergoes backshifting. Further examples follow.

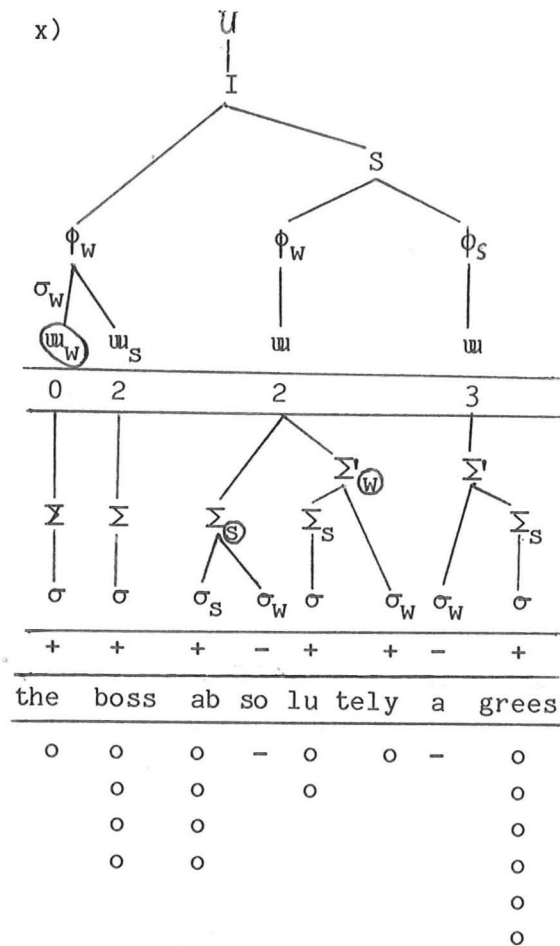


Examples vii) - x) have had the stress backshifted, as indicated by the circled 's' and 'w' symbols. In vii) and ix) it would be possible to produce a version without backshifting. But this would sound most unnatural, probably necessitating an intonational break immediately after stressed '-neath' (which does not usually happen with unstressed '-neath'). In viii) and x) a non-backshifted version would be more acceptable, since the greater number of syllables between the two stresses in such a case makes for greater evenness in the rhythm. Thus there is more scope for



variation in cases such as viii) and x), while in the other cases backshifting is almost compulsory. This points to a gradual scale of vulnerability to stress shifting among words, depending on context but also on the make-up of each word. Words such as 'absolutely' have a final unstressed syllable already present as a 'buffer', while words such as 'underneath' do not. When the following word also contains the 'buffer' of an initial unstressed syllable, e.g. 'agree', then it is even more likely that the rhythm rule will fail to apply, since a more evenly-distributed effect can be obtained by retaining the citation form. As Bolinger writes: 'Stability of stress pattern seems to vary all along the scale.' (Bolinger, 1965:162, footnote)

In examples ix) and x), the stress shift actually causes a new clash between the shifted stress and the preceding stressed monosyllable. In such cases, it is very likely that the stress shift backwards will not occur. If it does occur, there will probably be a pause between the clashing stresses, or else the first of the two clashing stressed



syllables will be lengthened slightly. Liberman discusses a similar situation, using the example phrase 'the anaphoric reference'. He proposes two alternative solutions to the problem of the stress clash in this phrase. One solution is to backshift the stress, as in the examples given here, so that it shifts from '-pho-' to 'an-'. The other solution is to add 'additional grid elements', as in example xii). Example xi) gives the original form, xii) contains two extra lower-level elements, and xiii) shows backshifting (L:196).

xi)

	3	3	
	2	2	2
1	1	1	1
the	anaphoric	reference	

xii)

	3		3
	2	2	2
1	1	1	1
the	anaphoric(∅)(∅)	reference	

xiii)

	3		3
	2	2	2
1	1	1	1
the	anaphoric	reference	

Lieberman describes the solution exemplified in example xii) as representing one way of saying the phrase. The phonetic nulls are realised as a pause. He observes that the pragmatic effect of this rendition seems to be to emphasise 'anaphoric', though this is not contrastively stressed (as deduced from basic stress pattern, tune-text association, and the nature of the requirements for 'pause').

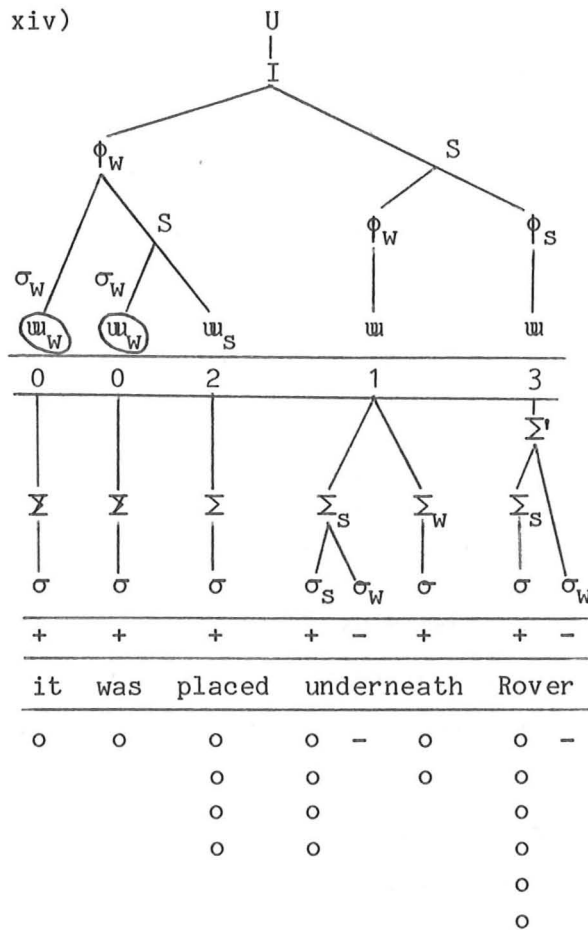
Lieberman observes of this pause:

'The phonetic nulls....(perceived as a kind of pause) seem to be different in source from the sort of intonation-break we perceive between clauses, or setting off parentheticals, or (in some cases) between subject and predicate. This difference can be observed in the fact that some such "pauses" constitute barriers which semantic operators like negation cannot penetrate - the "pause" in ["the anaphoric reference"] is certainly not of this character.' (L:196)

From Lieberman's description, he would seem to be saying that most pauses occur between two 'intonational phrases', in Selkirk's terms, while those pauses brought about by adding elements to the metrical grid are different in kind from the majority. The latter kind of 'pause' is completely independent of the structure of the utterance in terms of prosodic units, and so conveys no information about syntactic or semantic structure. Thus it cannot affect 'semantic operators', as Lieberman terms them, being of a different nature. This latter type of 'pause' is a phenomenon of rhythm and timing, and has the effect of adjusting the timing of the utterance in the direction of isochrony (though no claim for perfect isochrony can be made). This is because it lengthens the interval between two stressed syllables so that the interval approximates to the longer intervals between the other stressed syllables. An intonational contour can be continued over this type of 'pause' without sounding forced. The same cannot be said of intonation contours with the first type of pause, that arises from breaks between prosodic units.

Another feature of stress-shifting is seen in example xiv). Example xv) shows a non-backshifted version.

xiv)



xv)

it	was	placed	underneath	Rover
0	0	0	0 -	0 -
		0	0	0
		0	0	0
			0	0
				0
				0

In this example, there is a stress clash whether or not the stress is backshifted. However, the more natural of the two possible renditions seems to be the one where the stress has been backshifted so that it occurs immediately after the 4-column on 'placed'. In such a context, the 4-column on 'un-' would probably not count as stressed at all, coming immediately after a syllable which unambiguously counts as stressed. If the stress had remained on '-neath', it might have appeared to gain a certain degree of stress due to the preceding 'low' columns; but because of the 6-column immediately following, it would probably not have counted as stressed in this context. The simpler situation of the two, therefore, is that where backshifting has applied, since the syllable in question here, 'un-', would be less likely to appear stressed than would '-neath' in

the non-backshifted case. Naturally, these comments apply only to cases (the more usual ones) where the sentence is spoken reasonably quickly. In cases where the whole utterance is spoken slowly, 'underneath' may well contain a syllable that counts as stressed. Even here, however, it is the backshifted version that sounds slightly more natural and unforced.

If one examines backshifting of stress from the point of view of the metrical grid, one may easily conclude, with Bolinger, that

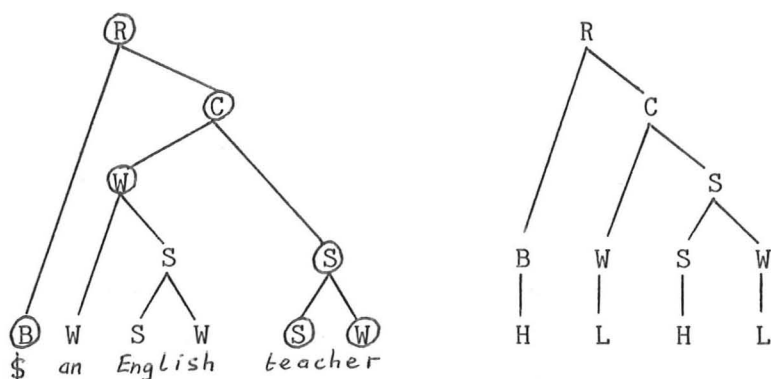
'....an example like a clear-cut problem can more accurately be described as a sacrifice of the accent on cut than as a shift of the accent "from" that syllable "to" clear, since clear normally has an accent anyway....' (Bolinger, 1965:172)

Although the process may involve a swapping of 's' and 'w' nodes and a consequent reallocation of place-holders in the metrical grid, the net effect is to deprive a syllable of the prominence it had earlier, even if this prominence was merely of a secondary nature. In the case of examples such as 'It was placed underneath Rover', both stressable syllables of 'underneath' lose their prominence, given the normal rate of utterance. Thus backshifting of stress could be seen as merely part of the larger phenomenon of stress downgrading due to context.

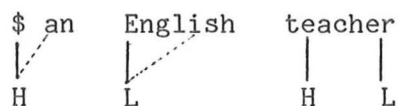
### 3.6. SOME INTONATIONAL CONSIDERATIONS

### 3.6.1 Congruence

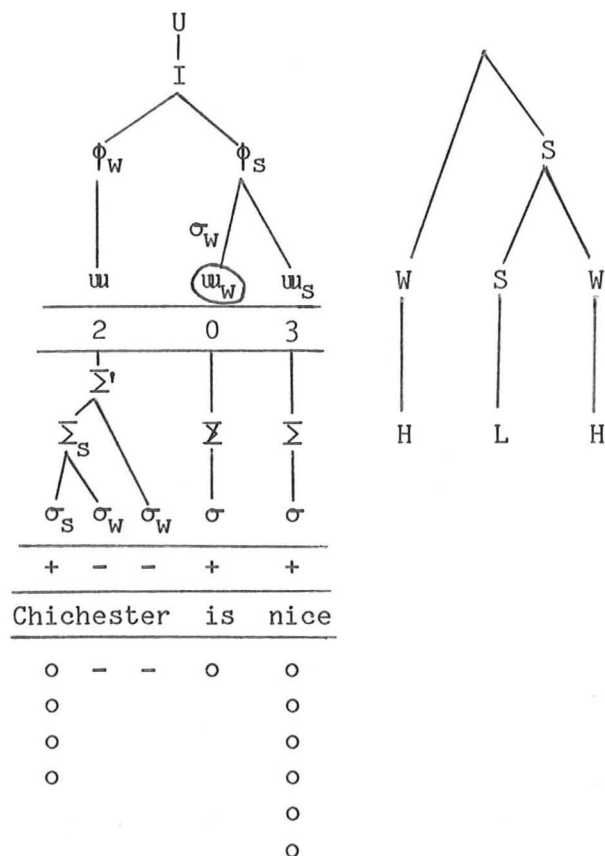
Lieberman accounts for intonational patterns by matching the metrical structure of the 'text' to the 'congruent' metrical structure of the 'tune', one of a certain number of specific and pragmatically significant such 'tunes'. The structure of each 'tune' is made up of distinct 'tones': 'L' (low), 'H' (high), 'L-M' (low-mid), 'H-M' (high-mid). These are arranged in a metrical structure with binary 'w' and 's' nodes. These nodes must match with the 'w' and 's' nodes of the text's 'metrical structure' for this intonation pattern to apply to the text (though often an initial 'w' tone is optional). By 'text', Lieberman means 'sentence' or 'utterance'. The example below shows congruence between tune and text (L:78):



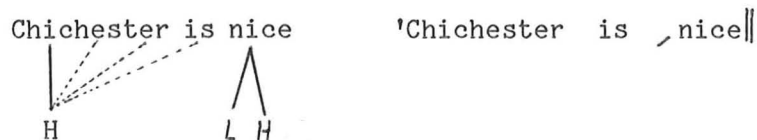
The circled nodes are those which are 'congruent' to the nodes of the metrical structure of the tune. The 'B' is a 'boundary tone', i.e. a tone which is metrically weak and occurs at the beginning (and end) of shorter utterances. The 'C' refers to 'content', that is, non-boundary material, which is metrically strong. By means of the boundary tone, one can account for such phenomena as the 'high pre-head'. The 'tune' given above is referred to by Liberman as the 'surprise/redundancy tune with "high pre-head"' (L:77), and associates with the text as follows, where the dotted line represents an association derived by rule. Later rules will convert to a glide any sequence of pitch levels assigned to the same syllable.



### 3.6.2 'Neutral' pattern

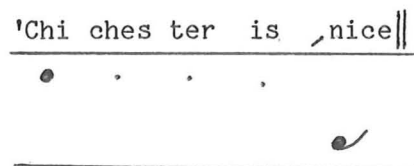


In this example, there is perfect congruence between the metrical structure of the text and that of the suggested tune, giving rise to the following association. The representation using tonetic stress marks is not meant to imply that this nucleus type is the only one possible.

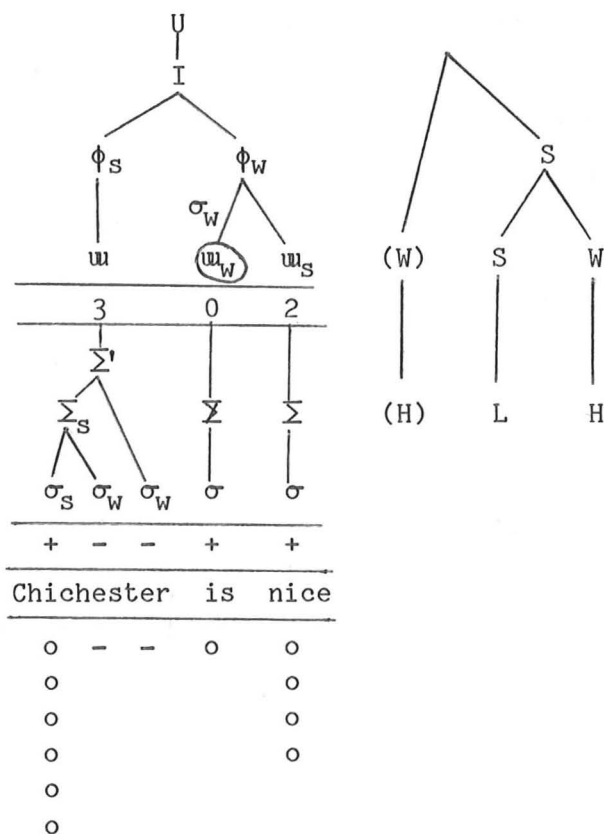


The focussed element is 'nice'; this pattern, the statistically unmarked rendition, draws attention to the 'niceness' of Chichester. The element 'Chichester', while not actually completely predictable, is not unexpected - as in a walk around the town, or when reading a description

of it. Thus 'Chichester' is metrically weak. In British school terms, 'Chichester' begins a 'high head'; this begins at a high pitch level, but most often gradually sinks in pitch towards the low rise on 'nice'. The high head could be replaced by a falling head (symbolised 'v') without altering the connotations as described above. The pattern of the intonation as described above would be approximately as follows:

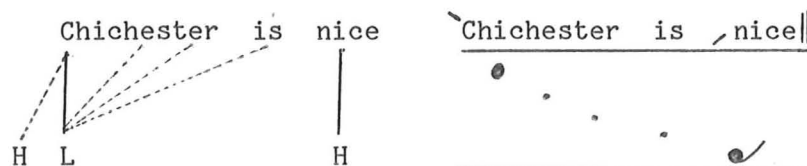


### 3.6.3 Compound nucleus



In this example, 'Chichester' is focussed, having the highest value of [content]. This entails reallocation of 's' and 'w' nodes, so that there is no longer perfect congruence between the metrical structure of the text and that of the suggested tune. In such a case, the 'stray' initial 'H' tone is realised as the starting-point of a glide on the first syllable of



'Chichester'. The final pitch glide is brought about by a rule triggered by 'low' pitch and a following 'high' pitch specification. This results in a rise from a low pitch.



The metrical bracketing of the text follows the syntactic structure where possible. The shift of the focussed constituent to 'Chichester' causes the least possible disturbance to the overall structure. In terms of information conveyed, attention is drawn to both 'Chichester' and 'nice' to an equal degree, and there are no hidden implications.


'Boundary tones' are ignored in these examples, since the initial syllable of the sentence is stressed. In such a case, 'the distinction between H and L initial boundary tones is neutralized' (L:81) - in effect, no boundary tone is realised. Liberman writes that one can 'think' the difference between boundary tones, but that the phonetic effect is negligible or non-existent in such cases: thus I have ignored them.

Example 3.6.2, if the head were of the falling type, would have exactly the same realisation as example 3.6.3. in terms of intonation contours. Thus the sequence (falling) head plus (low rise) nucleus would be indistinguishable from the sequence (high falling) nucleus plus (low rise) nucleus. However, it is possible to justify the difference in analysis: in example 3.6.3. the speaker can pause briefly (before 'is') and still retain the connotations of the version without pause (i.e. attention is drawn to both 'Chichester' and 'nice' to an equal extent). This is not true of example 3.6.2. D. Crystal has observed that a compound tone-unit has a 'bowl' or 'trough' of unstressed syllables between the two nuclei. These intervening syllables rise or fall gently, level off, then fall or rise (respectively) very gently towards the next nucleus. Crystal describes the pattern thus:

'The syllables between the two kinetic elements must display an evenness of pitch pattern, continuing the pitch movement in a 'trough' or sustained arc from one to the other, for example  or .' (Crystal 1969:218)

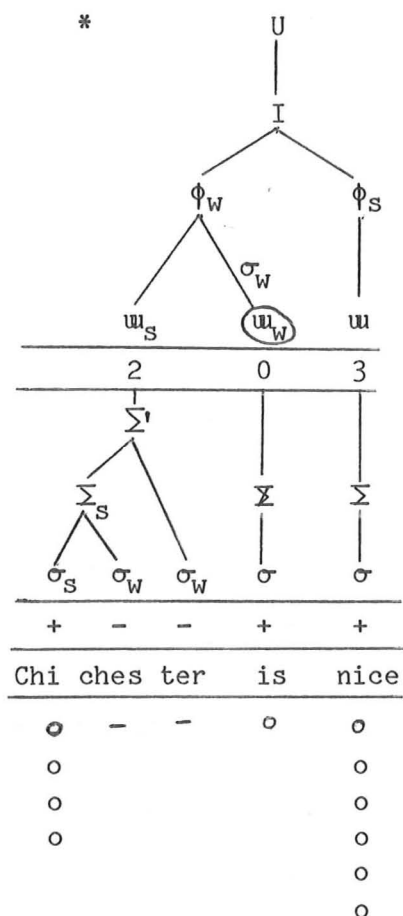
The sentence in 3.6.3. could be produced in this way, but that of 6.2. could not be so produced without losing its connotations. The 'gradual' contour would be roughly as follows (for 3.6.3):

Chichester is nice||



If there is a pause in 3.6.3., the two separate tone-units (intonational phrases) could have the nuclear tones \+ \ or /+/. Without such a pause, no sudden 'change in direction' of the intonation contour is allowed, so these nucleus combinations would not occur. Crystal distinguishes between 'exocentric' nucleus sequences, necessitating a sharp change in direction of intonation, and 'endocentric' nucleus sequences, showing the 'trough' effect - e.g. /+ \, \+ / (Crystal 1969). In 3.6.3. however, there can be no pause and thus no such possibilities of different types of contour sequence, since it is held by some scholars that head type is not completely independent of nucleus type (see O'Connor and Arnold, 1961).

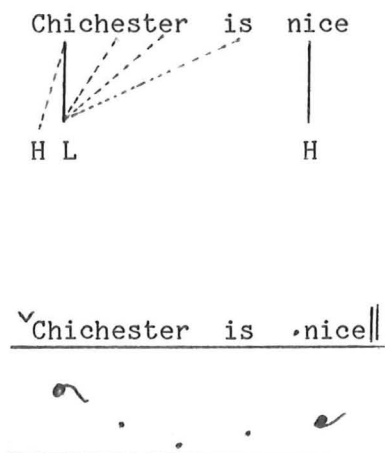
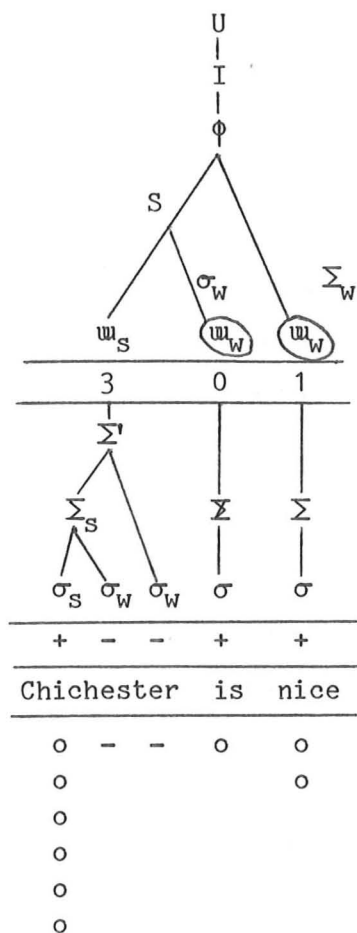
The metrical bracketing of 3.6.2. must be as shown. The only alternative bracketing (given below) is incorrect:-



It is not possible to pause after 'is' without seeming to hesitate. So, although this bracketing does not contravene the metrical bracketing rules, it must be ruled out as incorrect. As a general principle, one could state that in simple sentences such as this, the (surface) syntactic structure must be followed as far as it is possible to do so.

Lieberman's system does not readily allow for compound tone-units, as it deals only in simple tone-units, or 'tunes', with recognised intonation patterns that perform a particular function. Thus his system would be unable to distinguish between a compound tone-unit and a sequence of head and nucleus, in cases where the intonation pattern is identical. Since the connotations are quite different, and the potentialities for pause and choice of contour are also different, the proposed system seems to give a fuller account of intonation than does Liberman's system.

### 3.6.4 Complex nucleus (i)



In British school terminology, this pattern is an example of a 'complex nucleus', in this case a fall-rise. Attention is drawn to 'Chichester', while 'nice' is not unexpected, (as in the situation of looking through brochures of 'nice' places to visit, etc.), and so is metrically weak.

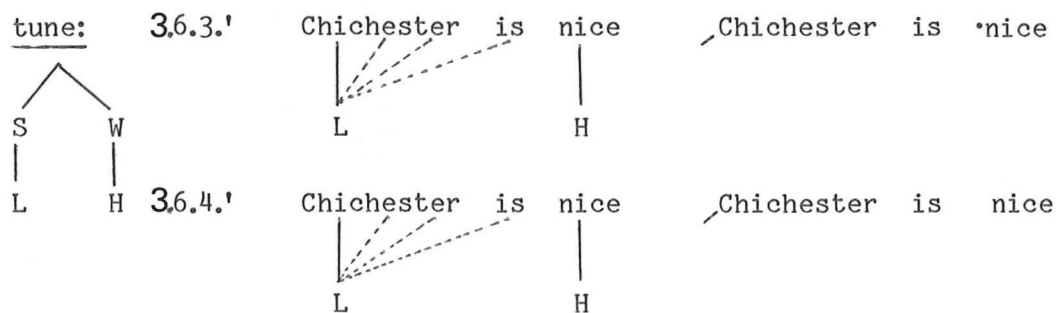
In this example, the semantic weight of 'nice' is reduced, and so it becomes merely [1 content]. This leads to its being characterised as metrically weak, and thus subject to 'dewording', so that it is reduced to the status of a weak (stress) foot. Because of this, it receives only two place-holders in the metrical grid, and so may be either stressed or unstressed, depending on the reset value of the threshold of the algorithm (i.e. depending on tempo of utterance). The word 'nice' cannot, however, be pitch-prominent, even if it is stressed, since it cannot be the DTE of a  $\phi$ , remaining a mere  $\Sigma_W$  unit. Since 'nice' has been reduced to the status of stress foot, it cannot form the sole realisation of the 'H' tone. Thus the

rise is realised on some preceding syllables as well. If 'nice' is counted as stressed (which depends on tempo, etc.), it may continue the rise with a small, non-pitch-prominent glide.

The metrical bracketing shown is the only one possible, given the [1-content] value of 'nice' and the [0 content] value of 'is', which dictates that they be metrically weak. This bracketing does not follow the syntactic structure, and if a pause were made after 'is', as suggested by this bracketing, the result would be unnatural. However, this pattern must be produced without any pause, in order to retain the connotations; this is not so in 6.3., where a pause could be made after 'Chichester', and where the metrical structure follows the syntactic structure.

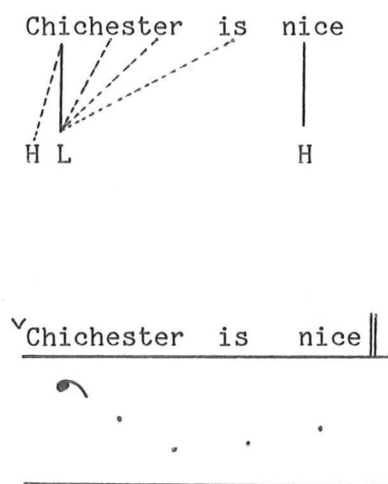
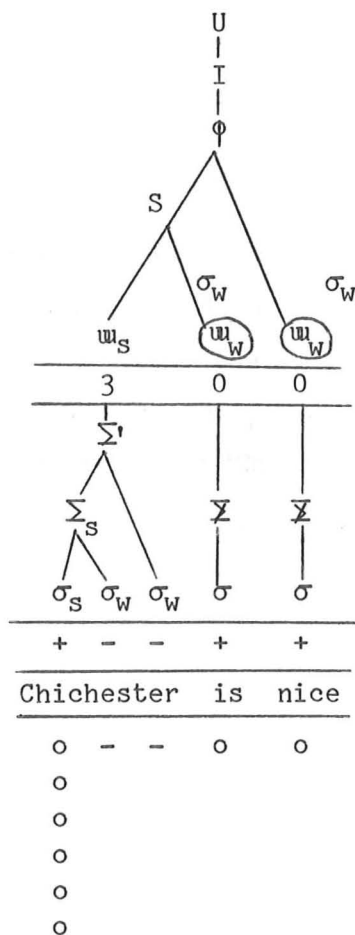
Another possible connotation of 3.6.4. is the hidden implication: '(Chichester is nice....) but (another town) is even nicer.' This implication could be conveyed at least partly by the fact that 'Chichester' is the focus of the sentence, 'nice' is not accented, and the intonation contour rises, indicating non-finality. Here again, there can be no pause, either after 'Chichester' or after 'is', without giving undue emphasis to either 'Chichester' or 'nice' respectively, and thus losing the original connotations. The nuclear tone type may be changed - for example, to a low rise - without changing the relative semantic weight given to each word, or the 'hidden implication'. This potential for changing the nucleus type is what distinguishes this sentence from one such as 3.6.3., where different results would be obtained.

In example 3.6.4. 'nice' is deaccented and so cannot take the 'H' tone on its own account. If the tune were the 'L-H' portion of the suggested tune, the effects on 3.6.2. and 3.6.3. would be different:



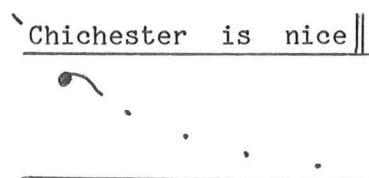
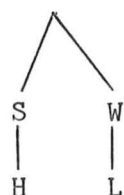
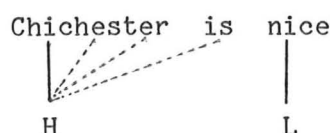
The effect of the new tune would be to assign a low rise nucleus in each case. A fundamental difference between the two sentences, however, is revealed in the fact that the structure of 3.6.4' is still nucleus plus tail, as it was in 3.6.4., while the structure of 3.6.3', also nucleus plus tail, differs from the compound nucleus structure of 3.6.3., and has completely different connotations. Thus, although in one particular realisation 3.6.3. and 3.6.4. may seem identical on the surface, their structural difference becomes evident when one considers such things as potential for pause, connotations, and the effects of varying tunes and nuclei. This kind of phenomenon is not considered at all by Liberman, and thus indicates another possible area where the proposed analysis improves on that of Liberman. This is because the latter has no 'semantic level' affecting other levels.

### 3.6.5 Complex nucleus (ii)



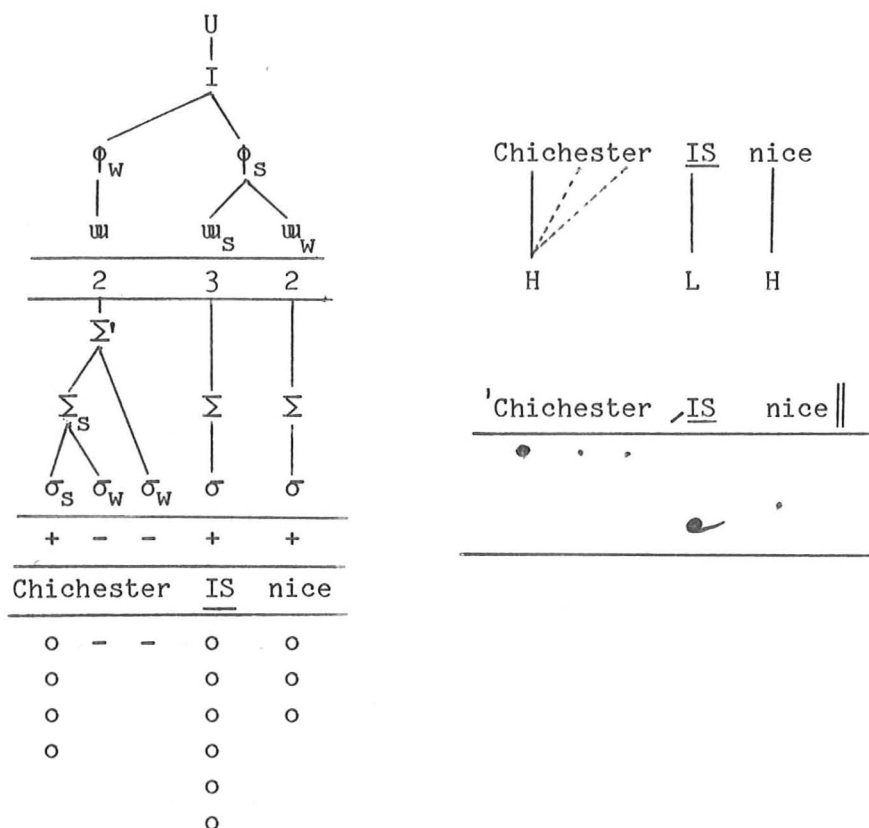
The bracketing shown in this diagram is the only possible one; again, no natural pause may be made within the utterance. The word 'nice' has been completely deaccented to [0 content], leading to defooting. Since the only word with a positive value of [content] is 'Chichester', this must become [3 content], leading to its bracketing as the DTE of the utterance, and assignment of the nucleus on this word.

This example differs from 3.6.4. in that, in 3.6.5., 'nice' cannot be stressed without coming to seem too important for its context. The context is one in which 'nice' is anaphorically destressed; e.g. when the utterance is a response to a question such as 'Which town did you say was nice?'. A compound nuclear tone is not possible in 3.6.5., since 'nice' cannot be the DTE of a phonological phrase, being reduced in status to a mere syllable. Thus it can never be either pitch-prominent or stressed. A more usual tune for this type of utterance would be as follows (i.e. a high fall nucleus):



Anaphoric destressing is the change to [0 content] from a positive value of [content]. 'Contrastive stress' is the extreme case of anaphoric destressing, when all the words after (or before) the focussed constituent have been destressed. This is what has happened in 3.6.5., which has contrastive stress on 'Chichester'. In this case, as in 3.6.4. 'nice' cannot be the DTE of a  $\phi$ , being defooted, and so it cannot take a 'tone' on its own account.

### 3.6.6 Stressed nonlexical word



In this fairly unusual situation, stress falls on a nonlexical word, in this case the copula. The nonlexical word then takes [3 content] and becomes the DTE of the utterance. The word 'nice' may or may not be counted as stressed, depending on the reset value of the threshold (corresponding to tempo of utterance). The nuclear tone type is a low rise, in this instance. The metrical bracketing follows syntactic structure, and there may be a slight pause after 'Chichester'. This utterance could be spoken in a situation where the copula is to receive emphatic stress, as if protesting the truth of what is asserted. More usual would be a falling nucleus on 'is', giving the utterance more of an air of 'emphatic finality', which is more appropriate to the function of asserting the truth of what is said. If 'nice' were destressed to [1-content] or [0 content], it could never count as stressed; if it is [2-content], there remains the possibility of its being stressed. It cannot be a  $\phi_W$  unit, as the DTE of the U is the last  $\phi$  unit. To become pitch-prominent it must thus be an I unit, which would be very rare.

### 3.7. LOCATION OF [3 CONTENT]

#### 3.7.1 Neutral pattern

Within each I unit, only one word is usually [3 content]. This word is usually the last lexical word in the I unit. However, in some cases more than one word may be [3 content], i.e. the focussed constituent may be a prosodic unit of a level higher than the prosodic word. For example, an answer to the question; 'To whom did John give the information?' could perhaps be; 'John gave the information to Susan's uncle', with destressing of 'John gave the information to' and consequent contrastive stress on 'Susan's uncle'. [4] In a neutral rendition of the sentence, only 'uncle', not 'Susan's', would be prominent, as in the example on the next page, which is in a less emphatic style than the rendition given in sections 3.4.6, and 3.5.4. A suggested intonation pattern is also given.

#### 3.7.2 Contrastive stress

With contrastive stress on 'Susan's uncle', both 'Susan's' and 'uncle' would be [3 content], while all preceding words would be [0 content], having been destressed. Defooting applies to all monosyllables that are [0 content], giving the example on page 185 where the syllable 'in-' may perhaps also count as stressed. A possible intonation pattern is given.

It is difficult to produce a long stretch of completely unstressed syllables, and the stress on '-ma-' (also the possible stress on 'in-') will break up an otherwise intolerably long stretch. This could be seen as further support for the fact that the defooting rule as formulated in 3.4.5.1 applies only to monosyllables. Since 'Susan's' is [3 content], it is

---

[4] See section 3.3.7.2. for remarks on contrastive stress and emphatic stress.

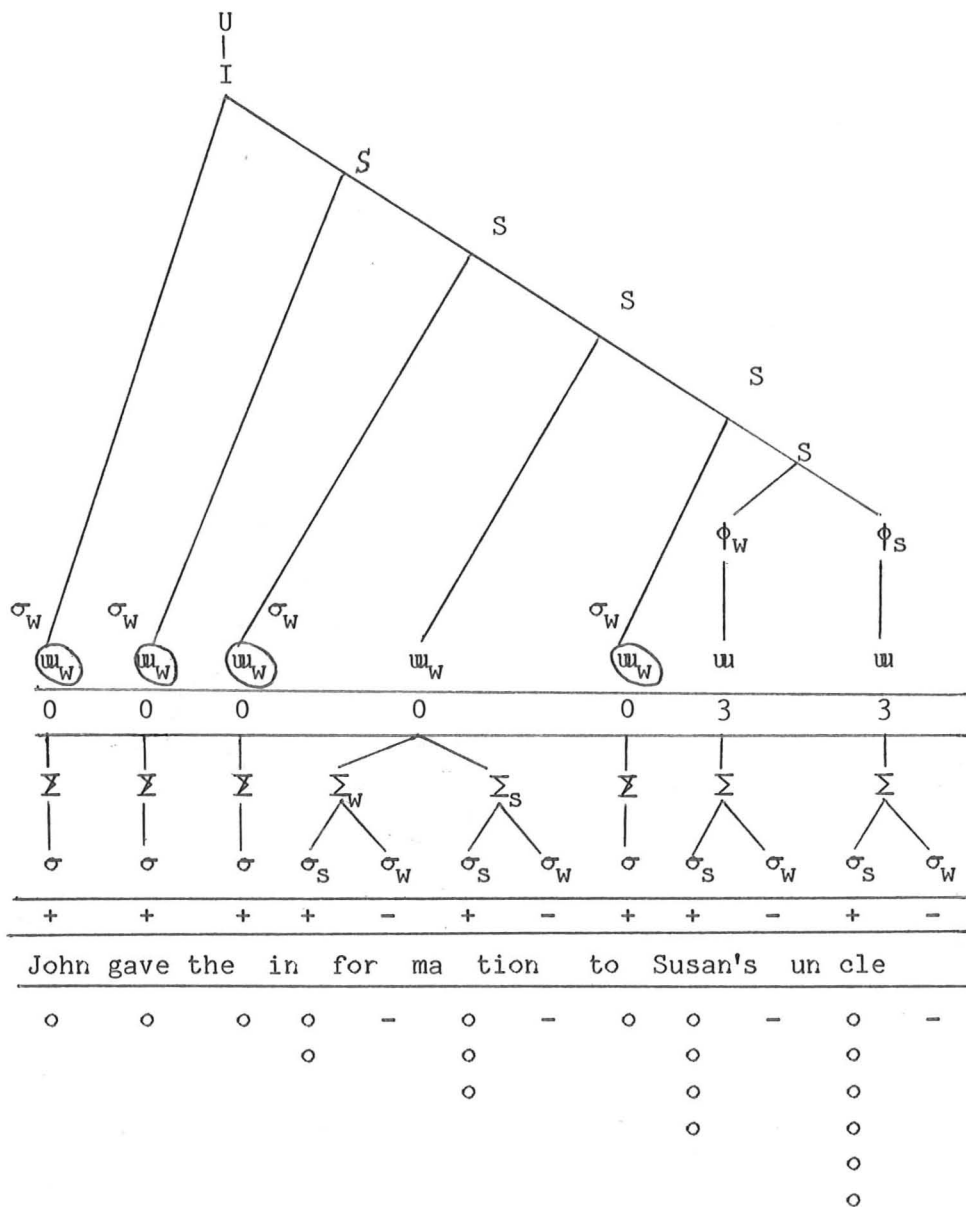
for 3.71.



### 3.7.3 Emphatic stress

184

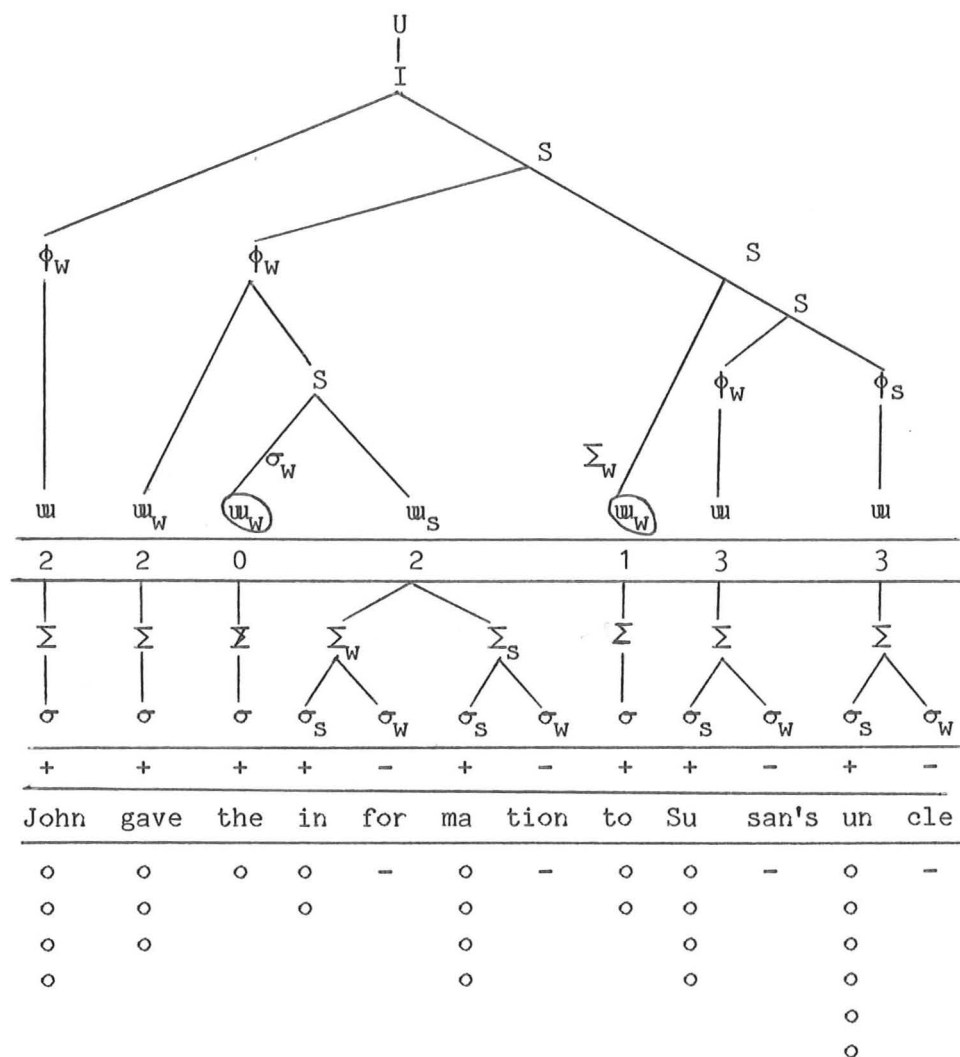
for 3.7.2.



John gave the information to 'Susan's uncle||

The only difference between the example in 3.73, and the 'neutral' pattern of 3.71, is that 'Susan's' is pitch-prominent in the former, owing to its greater semantic weight. In the suggested intonation pattern it is also highlighted by a pitch jump up. While a step down in pitch would have sufficed to make it pitch-prominent, the fact that this word is [3 content] calls for even greater prominence. This can be achieved by means of the step up in pitch, indicated here by an upwards arrow.

for 3.7.3.



'John gave the ·infor'mation to ↑'Susan's uncle||

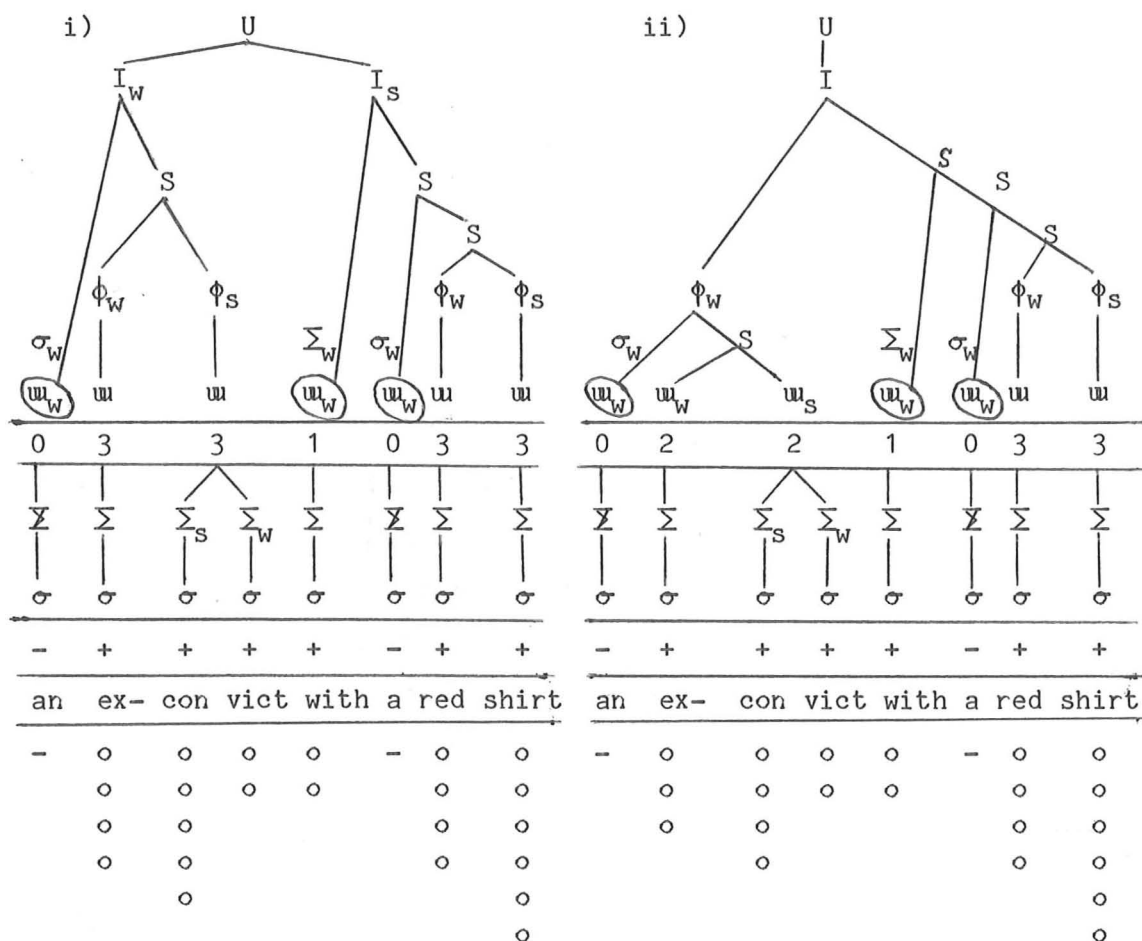
### 3.7.4 Discussion

As noted, the preceding examples would have the nucleus on 'uncle' only. In a relatively short sentence such as this, it would be unusual to find that both 'Susan's' and 'uncle' were the DTE of an intonational phrase, and thus the nuclei of separate tone units. Because of this, the different phrasings do not affect nuclear tone placement, and relying on the latter alone will give no clue as to the phrasing. This point is made by Chomsky, taking as his example the phrase 'an ex-convict with a red SHIRT' (Chomsky

1970). The capitalised word, in Chomsky's notation, carries the nuclear tone. He points out that this phrase could be interpreted as focussing on any of the following constituents;-

- i) an ex-convict with a red shirt
- ii) with a red shirt
- iii) a red shirt
- iv) shirt

These different interpretations correspond to different phrasings of the same sentence, where the focussed constituent may be either the last lexical word (as in iv) or else the last prosodic unit of some higher level (as in i, ii, and iii). The different interpretations may be represented as in the next diagram. (Interpretation iii would have exactly the same representation as interpretation ii, and so is not shown separately.)



i) an 'ex-convict | with a 'red shirt||      ii) an ex-'convict with a 'red shirt||

iv)

0	2	2	1	0	2	3
$\Sigma$	$\Sigma$	$\Sigma_S$	$\Sigma$	$\Sigma$	$\Sigma$	$\Sigma$
$\sigma$	$\sigma$	$\sigma$	$\sigma$	$\sigma$	$\sigma$	$\sigma$
-	+	+	+	-	+	+
an ex- con vict with a red shirt						
-	o	o	o	o	-	o
	o	o	o	o		o
	o	o				o
		o				o
						o
						o
						o

In each of the three examples, the suggested intonation pattern makes prominent the appropriate words. In cases where the intonation pattern is such as to make explicit the difference between various phrasings, it can always be traced back to these different structures. Where the intonation pattern does not make structural distinctions explicit, as in 3.6.2 and 3.6.3, the differences are seen in varying potentialities for such things as pause position and tune type.

In the discussion of intonational considerations in sections 3.6 and 3.7, it can be seen that there are various degrees of divergence from the normal, more frequent pattern, which is represented by 3.6.1. The smallest degree of divergence is seen in 3.6.2, where a [2 content] and a [3 content] value are exchanged: this necessitates no rebracketing, merely a change

in higher-level node labelling. A little more divergent than this is the example in 3.6.3, where, as well as a relocated [3 content], a [2 content] word is destressed to [1 content]. This pattern on the semantic level necessitates rebracketing on the phrasal level, but need not affect the distribution of stressed and unstressed syllables as defined on the metrical grid. More divergent still is example 3.6.4, where a normally [2 content] word has been completely destressed to [0 content]. As in 6.3., this necessitates rebracketing at the metrical level, but there is the difference here that the destressed word cannot possibly count as stressed, in terms of the metrical grid. Thus the distribution of stressed and unstressed syllables in the sentence has changed. Most divergent is 3.6.5, where a normally [0 content] word becomes [3 content], and so is not affected by the defooting rule. This leads to changes in the node-labelling at a level higher than that of the word, also to the development of a 'new' stressed syllable where there was none in the earlier interpretations. If the nonlexical word involved had contained a normally [-full] vowel, this would have been changed to a [+full] vowel, when stressed, and only in this (very divergent) situation.

Likewise, in section 3.7 it is the case where lexical words take a value of [0 content] that corresponds to the greater divergence from the norm - that is, contrastive stress is more disruptive of the normal pattern than is emphatic stress, in terms of phenomena at the semantic and phrasal levels. Some tentative support for this grading of contrastive and emphatic stress may be seen in the fact that the latter (as observed in 3.3.7.2) is essentially paradigmatic in operation, and so cannot affect the syntagmatic metrical structure as profoundly as can the more syntagmatic contrastive stress.

### 3.8. APPLICATION OF THE SYSTEM TO WELSH

#### 3.8.1 Vowel level

Welsh does not reduce vowels in the same way as English, since schwa can be stressed in Welsh. Welsh vowel mutation, though similar to vowel reduction in that it involves a change to schwa, is in fact independent of stress (which nearly always falls on the penult) and depends instead on the location of the given vowel with respect to the final syllable. Thus there is little justification for a feature [ $\pm$ full] (or any other such segmental feature), and there are no distinctions to be made at the 'vowel level' in Welsh.

#### 3.8.2 Word level

##### **3.8.2.1 Vowel length rules**

Selkirk's system may be adapted for Welsh, with one important modification. In English, the status of a syllable in terms of 'stress feet' can be said to reflect the 'syllable weight', which involves e.g. tenseness of the vowel and number of following consonants. In Welsh, on the other hand, although vowel length is phonemically distinctive, it can nearly always be predicted from the quality of any following consonants within the same syllable, subject to the overriding principle that vowels in unstressed syllables are short whatever the following segments. The general principles behind Welsh vowel length may be stated informally as follows. [5]

---

[5] Rules and exceptions adapted from T. Rhys Jones (1977:21-22).

i) In unstressed syllables (non-penults and unstressed monosyllables)

All vowels are short. (This class includes the 'proclitics').

ii) In stressed syllables (penults and stressed monosyllables)

Vowels are short in the following environments:-

α) When followed by one of /p,t,k,m,ŋ/

β) When followed by two or more consonants.

γ) When all vowels except orthographic 'i' and 'u' (i.e. /i/ and some cases of /ɪ/ respectively) are followed by one of /l,n,r/.

Vowels are long in the following environments:-

α) Syllable-finally.

β) When followed by any consonant other than those mentioned above (i.e. /b,d,g,f,θ,x,v,ð,s/).

γ) Orthographic 'i' and 'u' are always long before /l,n,r/.

iii) Exceptions

α) Exceptions to these rules are few; some vowels are marked with a circumflex orthographically to indicate a long vowel in an environment where a short vowel would have been predicted by the usual rules.

e.g.      ôl    behind              pâr    pair              pêl    ball  
            /o:l/                      /pa:r/                      /pe:l/

β) A very few words contain vowels that are long where a short vowel would have been predicted, and yet are not marked with a circumflex orthographically.

e.g.      hen    old    /he:n/                      dyn    man    /dɪ:n/

γ) In polysyllables, a stressed long vowel becomes somewhat shortened to 'half-long', but is still recognisably distinct from a stressed short vowel, the difference lying partly in the quality of the vowel.

e.g.      ceg    mouth                      cegin    kitchen              ceginau    kitchens  
            [ke:g]                              [ke:grɪn]                      [kɛgi.nɛ]

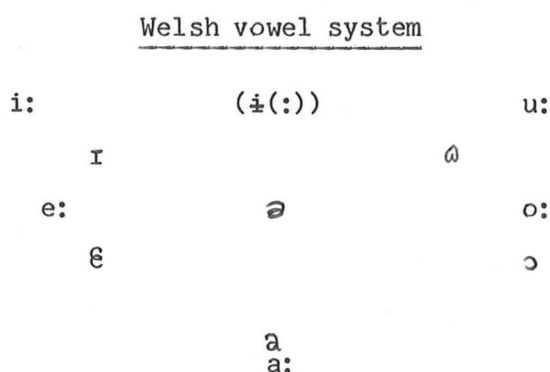
R. Jones (1967), in a study of three Welsh dialects, states that this 'half-length' occurs in only one of the dialects, namely that of Tŷ Ddewi (St. David's), the only Southern dialect of the three. His two North Welsh dialects have a short vowel instead. When pre-

consonantal, a stressed short vowel in a polysyllable provides the environment for 'extra duration' of the consonant, which is markedly lengthened (R. Jones 1967:38).

δ) Some English loan-words contain vowels that are short where by Welsh rules they would have been long; such words are few in number.

e.g. bag bag /bag/, pin pin /pin/, bil bill /bɪl/

The vowel schwa, while a separate phoneme in Welsh and capable of being stressed, has no long counterpart. Even under the conditions appropriate to a long vowel, it remains short. The other vowels can be seen as long/short pairs, as in the following diagram. [6]



The quality differences arising from the length variation are seen by Jones as merely allophonic, determined solely by the independent length 'prosodeme' (R. Jones 1967:211). However, there are a few minimal pairs distinguished by the occurrence of a long or short vowel, for example:-

sê	<u>seat</u> , <u>pew</u>	/se:t/	set	<u>set</u>	/sê/
tôn	<u>tune</u> , <u>tone</u>	/to:n/	ton	<u>wave</u>	/tɔn/
dî	<u>team</u> (lenited)	/di:m/	dim	<u>nothing</u>	/dim/
mân	<u>tiny</u>	/ma:n/	man	<u>place</u>	/man/
gŵn	<u>gown</u>	/gu:n/	gwn	<u>I know</u>	/gɔn/
bû	<u>I was</u>	/bi:m/	bum	<u>five</u> (lenited)	/bim/

Such examples support a view of the corresponding long and short vowels

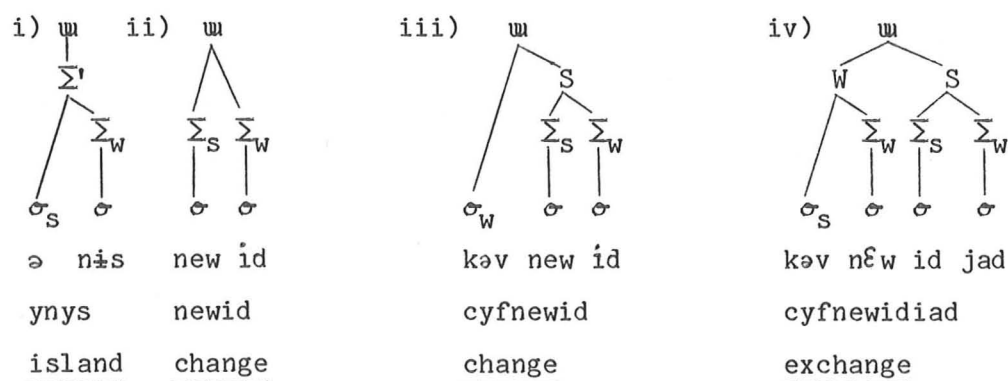
---

[6] Only North Welsh dialects possess /ɨ(:)/; in South Wales this has become [i:] or [ɪ], as appropriate.

as separate phonemes. [7] These (descriptively adequate) rules may be taken as the basis for an analysis of Welsh on the same lines as the proposed analysis of English. It will also be sufficient to consider only 'normally-stressed' words - that is, stressed monosyllables and polysyllables stressed on the penult. There are some exceptions, but these are due either to direct borrowing from the English, or to historical syllable contraction, or to the presence of an unstressable proclitic as the penult.

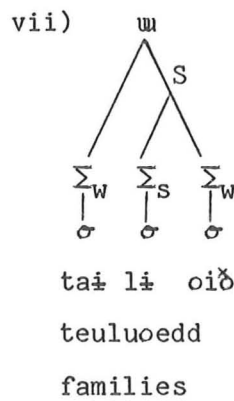
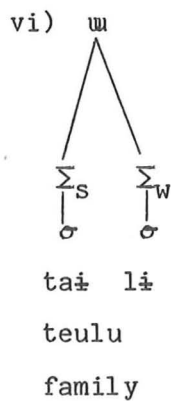
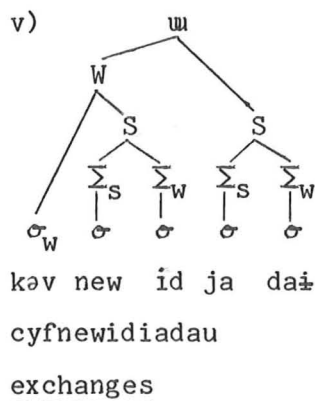
### 3.8.2.2 Underlying word structure

Selkirk's system of 'stress feet' may be adapted for Welsh so that a vowel which is 'long' according to the above rules constitutes a monosyllabic foot, while two neighbouring 'short' vowels join to make up one stress foot. The stress superfoot also appears, where the 'stray' syllable may be either to the left or to the right of the first-level foot, and, in underlying structure at least, must contain a 'short' vowel. The feet are joined in a right-branching structure according to Selkirk's rule for English; i.e., of two sister nodes, the second is strong if and only if it branches (see end of section 3.2.2). Since schwa can never be long, it can never occur in a syllable constituting a monosyllabic foot. Examples of underlying structures follow.

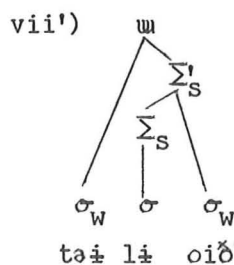
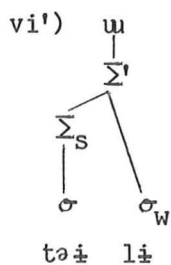
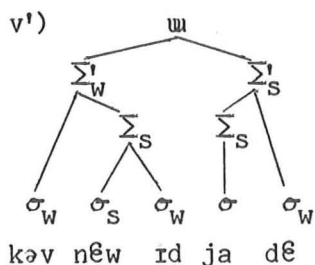
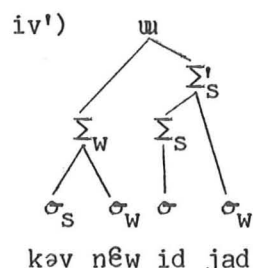
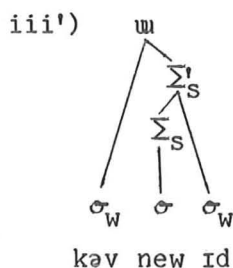
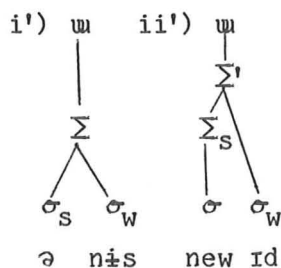



---

[7] Williams (in press) presents an analysis of the Welsh vowel system in which certain consonants are viewed as underlyingly geminate, and the preceding vowel therefore short. Vowel length is seen as partially dependent on the number, rather than the nature, of following consonants; thus vowels are not marked for length in underlying structure.



This type of underlying representation makes explicit the identity of sequences such as 'newid' under different stress patterns which may change the distribution of short and long syllables within the sequence. Thus, in the examples above, each syllable of 'newid' (whether this is a single word or part of a longer word) constitutes a monosyllabic foot, even where one or both syllables are unstressed (i.e. not the DTE of the word). In the derived representation, however, no syllable that is not the DTE of the word may constitute a monosyllabic foot - i.e. it may not be long - and the DTE may only constitute a monosyllabic foot if the environment is appropriate. Thus there is a rule 'defooting' all syllables that are not the DTE of a word, giving such surface representations as the following.



### 3.8.2.3 Surface word structure

In the above examples, 'defooting' has applied to non-DTE syllables, which are now represented as containing short vowels. The final syllable of v'), the plural morpheme '-au', becomes [-ɛ] in colloquial Welsh, and is represented here in this form, while the first diphthong in vi') and vii'), the diphthong '-eu-', becomes [əɪ] in non-final syllables, and is represented here in this form.

The rules of defooting for monosyllabic [0 content] words, and dewording for monosyllabic [1 content] words, apply in Welsh as in English (see 3.4.5.1. and 3.4.5.2). A more detailed consideration of the segmental differences between underlying and derived representations will be found in Williams (in press). It is worth pointing out that this analysis allows the following structure (i.e. a word consisting of a single syllable that does not count as a foot).

ω

|

σ

Such a pattern would not be found in English, where such 'stray' syllables could not constitute a prosodic word. In Welsh, however, it is possible for any vowel, short or long, to be stressed, and this fact may be easily reflected by allowing both 'footed' and 'non-footed' syllables to constitute monosyllabic words. Such a pattern is seen in the case of unstressed monosyllables, i.e. nonlexical words such as the possessive adjectives 'fy', 'dy', 'ein', etc., the particle 'yn', the definite article 'y', 'yr', etc. (see also section 3.8.2.5.2. in the case of prefixes). Also, very many lexical monosyllables take this pattern, such as 'plant', children; 'cot', coat; and 'twp', stupid.

### 3.8.2.4 The prosodic word

In Welsh as in English, there seems to be some difficulty in defining what constitutes a prosodic word and what is an element on the phrasal level. Welsh also makes use of 'compound words', formed from two or more simple (i.e. noncompound) words. These compounds are often stressed on the penult only, as in monomorphemic words. Other compounds, however, have more than one stressed syllable. This difference in stress pattern is discussed by <sup>M.R.</sup>Allen (1975) as one of the reasons for distinguishing 'strict' and 'loose' compounds; the latter have an internal word boundary, the former have merely an internal morpheme boundary.

Allen deals also with vowel mutation and inflectional phenomena in support of her analysis. Vowel mutation is a morphological alternation whereby certain vowels appear as schwa in non-final syllables; for example, when the addition of a suffix deprives them of final syllable status. The changes in vowel quality are as follows.

ai --> i    ai̯ --> i̯    u/a --> ə (some cases of i̯(:) --> ə )  
i.e. ai --> ei    au --> eu    w --> y (y --> y) orthography

In some compounds, vowel mutation occurs regularly as in simple words, e.g. 'gwraig', wife --> 'gwreig-dda', goodwife. In other compounds, however, this does not occur: e.g.-

cyn-faer	/kɪn vair/	<u>ex-mayor</u>
gau-broffwyd	/gaɪ brofwɪd/	<u>false prophet</u>

Another distinction between compounds, according to Allen, is seen in the case of inflection. Some compounds inflect only the second element, thus 'behaving as if' they were simple, monomorphemic words - e.g.-

dryg-ddyn	<u>wicked person</u>	<u>plural</u> dryg-ddynion
/drág ðɪn/		/dræg ðɛnjɔn/

Other compounds inflect both elements, as though each were a separate word, as in the following:-

nefolyon-wybodeu	<u>heavenly sciences</u>	/nɛvɔljɔn wɪbódɛ/
------------------	--------------------------	-------------------

It is in this type of compound only that suspension of the second element can occur, as in the following example:-

nefolion- a'r daerolion- a thandaerolion bethau  
of heavenly and earthly and under-the-earthly things  
/nɛvɔljɔn ar dairɔljɔn a θandairɔljɔn beθɛ/

Allen also uses evidence from the productivity of compounds to argue for her distinction between strict compounds (with internal morpheme boundary) and loose compounds (with internal word boundary). In the preceding examples, these were respectively the first and second type of compound word. Allen states that loose compounds may be freely formed, their meaning easily derivable from the meaning of their parts, while strict compounds are not freely formed, are fewer in number, and have a meaning not transparently derivable from the meanings of their parts, being fixed into a more idiomatic meaning.

Stress patterns also support Allen's analysis. She finds that her 'strict compounds' are stressed regularly, on the penult, while 'loose compounds' have two stresses, supporting an analysis that considers them to be separate words.

Thus an analysis distinguishing between compounds which are a single word and compounds which are made up of more than one word is supported by evidence from vowel mutation, inflection, productivity and semantics, and stress patterns. Strict compounds consist of a single prosodic word made up of two or more feet, while loose compounds consist of two or more prosodic words joined to form a single higher-level constituent.

### 3.8.2.5 Irregularly-stressed words

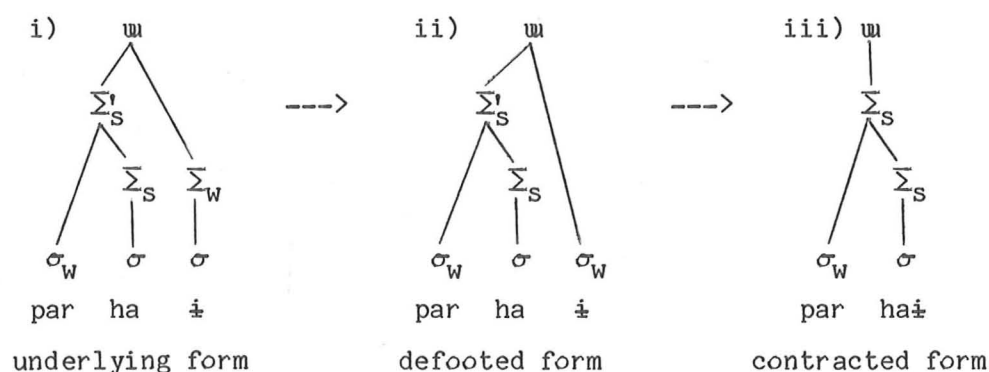
Some words in Welsh are not stressed on the penult. Most of these exceptions are stressed on the final syllable. Those that are not are usually loans from English that have retained the English stress pattern, as in 'Méthodist', Methodist, stressed on the first syllable. Such irregularly-stressed loans are few in number, and will not be considered further. Among the native words the finally-stressed words may be analysed into two main types.

### 3.8.2.5.1. Syllable contraction

Some words that originally contained an open penultimate syllable and an ultima beginning with a vowel later contracted the last two syllables to one syllable, due to the lack of an intervening consonant, leaving the stress on the new final syllable.

Such words are 'parhau', to continue, /parhai/, and any of the verbs ending in '-hau' or '-au'; also isolated instances of contraction such as 'Cymraeg', Welsh (language), /kəmraig/. Some monosyllables, such as 'maes', field, /mais/, were once disyllables, contracting later to a single stressed monosyllable. This process of contraction took place mostly before the Middle Welsh period, in the spoken language at least – that is, before the twelfth century A.D. (see D. Evans, 1976:4–5, for details of contraction).

These words are considered by Griffen (1979) to have an underlying uncontracted form, in which the stress is regular. A later rule contracts the two final syllables at a stage after the assignment of stress, so that the new final syllable remains stressed. Thus 'parhau' would be analysed as follows.



When a final syllable is stressed, naturally its vowel remains long (i.e. 'footed' if the environment is appropriate. Thus the diphthong in 'parhau' remains long.

### 3.8.2.5.2. Prefix as penult

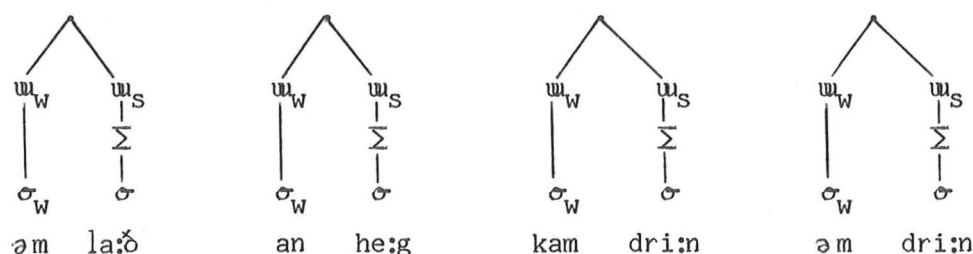
Other finally-stressed words are disyllables in which the penult is a prefix which can never be stressed. These include the prefixes 'ym-'

(reflexive and reciprocal), 'an-' (negative), 'cam-' (wrong, mis-), etc., as in 'ymdrin' to deal with /əmdri:n/; 'annheg' unfair /anhe:g/; 'camdrin' to mismanage /kamdri:n/ (last two examples taken from Jones, 1967:21). This may give rise to (quasi-) minimal pairs, such as:-

ymladd to fight /əmlað/  
 ymlâdd to wear oneself out /əmlá:ð/

There are no examples of monomorphemic words constituting a minimal pair distinguished solely by stress placement; thus there is no equivalent to the English distinction dígest/digést.

These irregularly-stressed disyllables have the following underlying form:



The prefixes, being marked [0 content], will subsequently be deworded and thus reduced to the status of a single stray syllable attached to the following word. Since the rules for node-labelling at the phrasal level nearly always assign the order w-s, this analysis agrees with previously formulated rules. The 'dewording' is a natural result of the 'defooting' rule seen earlier; as the syllables in question do not constitute monosyllabic feet, it is only the dewording effect that is seen. This analysis makes the prefixes formally identical to the stressless proclitics (e.g. possessive pronouns) and other stressless nonlexical monosyllables such as 'y(r)' the, and 'yn' (particle). These three groups of monosyllables have the common property of causing mutation (if possible) of the initial consonant of the following word. The prepositions also cause mutation, and these, being [1 content], will be likewise subject to dewording if monosyllabic. Other monosyllables causing mutation are 'a' and, 'un' one, the personal pronouns, interrogative pronouns, some intensifiers, and 'neu' or, 'pan' (noninterrogative) when, 'mor' 'cyn' as (in comparisons), and 'sy' which is. ('Y(r)' and 'un' cause

soft mutation to a following noun only if it is feminine.) All these other mutating monosyllables would be [0 content] and thus subject to dewording. Thus there seems to be a case for a unitary analysis of all these types of monosyllable (proclitic, prefix, and nonlexical word) in terms of a separate [0 content] word joined to the following word, which will always have a positive value of [content] and thus be the 'strong' partner of the two. This analysis reflects the close syntactic and semantic relationship between one such monosyllable and the following lexical word, as well as the reduced status of this monosyllable (dewording) which leads to its being seen as a subordinate element in a larger unit dominated by the lexical word. R. Jones (1967) treats this type of larger unit as his 'Class C' type of 'working unit', since within this unit only one 'primary stress' occurs. He writes (op. cit.:28):

'The units within this class could be called "grammatically marked forms". The immediate constituents of such units are - a discontinuous morph, an exponent of the possessive adjective, plus a Class A or Class B unit. They all exhibit phonetic and phonological features which are grammatically relevant, e.g.

[ən....i] 1st person singular possessive morpheme plus  
[ka:θ] cat = [əŋ 'ha:θi] my cat

In each example, one and only one primary stress occurs. Emphatic stress may occur in connected speech, replacing either the primary stress or the final secondary stress, e.g.

[əŋ 'ha:θi]	[əŋ 'ha:θi]	[əŋ 'ha:θi]
T + P + S	T + E + S	T + P + E

[In Jones' notation, T = tertiary stress, S = secondary stress,  
P = primary stress, E = emphatic stress]

Jones gives little consideration to the nature of his four degrees of stress, but quotes Gimson on stress: 'When a sound or syllable is stressed, it is being uttered with more muscular effort, increased air pressure, and greater amplitude of vibration.' (Gimson 1962:24) This suggests that he views stress in terms of intensity (and duration) rather than in terms of pitch-related phenomena. However, he does not consider how far, if at all, such a definition may be applied to Welsh. He writes:

'Attention to suprasegmental features, however, was limited to an impressionistic survey of stress and the duration of speech sounds relative to each other. It was found adequate to establish four degrees of stress although phonologically within these units [his 'working units'], only one is relevant. T stress and S stress are in complementary distribution, i.e. T = pre-P, S = post P, except that it

may occur pre-P when there is a potential emphatic stress placement on that syllable. Since their distribution is determined by the obligatory P-stress, they may all be regarded as positional variants of a phonological stress which need not be marked orthographically except for those forms which depart from the stressed penultimate patterning.' (op. cit.:29-30)

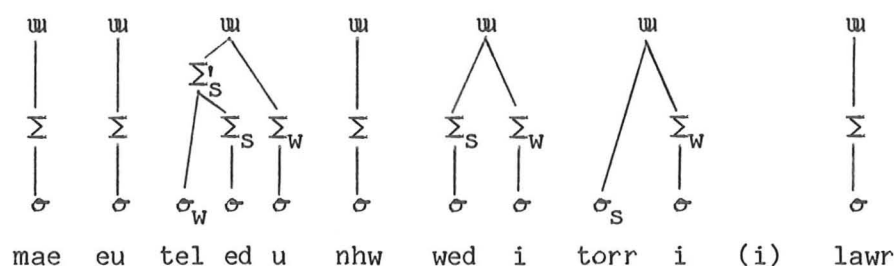
What this appears to mean is that although four degrees of stress may be distinguished (and Jones gives a reference on page 30 to a study recognising no less than six degrees of stress), nevertheless the only important distinction is between primary-stressed and non-primary-stressed. This distinction is precisely what the 'metrical grid' of an LP-type analysis sets out to make plain, in that a theoretically unbounded number of stress levels are resolved into the simple distinction between stressed and unstressed. 'Nuclear stress' and 'emphatic stress' are then as it were overlaid on this basic distinction, since considerations from metrical structure (e.g. locating the DTE of a particular type of prosodic unit) are then brought to bear on the structure defined by the metrical grid and by its algorithm for locating stressed syllables. Such an analysis resolves the difficult question of how many 'levels of stress' there are (on which it seems even native Welsh speakers cannot agree), in that it leaves the answer unspecified. If one wishes to think in terms of 'stress levels', the number present in a given sentence is dependent on such things as word class membership, semantic weight, syntactic and metrical structure, and morphological considerations. The same sentence may be uttered in different fashions, corresponding to different phrasings in terms of prosodic units, and this will also affect the number of 'stress levels' present in each utterance of the sentence. Perhaps a common failing in past considerations of stress and intonation has been the almost universal tendency to take one phrasing of a sentence - the 'unmarked' and medium-tempo phrasing - as the only possible way of uttering that sentence.

An example of the word-level structure of a short Welsh sentence is given on the next page.

'Mae eu teledu nhw wedi torri i lawr'

Their television has broken down

/mai ÷ tɛle:di̯ nɔ̯ we:di̯ tɔ̯ri̯ (i) laɔ̯r/



### 3.8.3 Semantic level

The 'semantic level' for Welsh will be very similar to that for English. The only differences are necessitated by differences in Welsh syntax; added to the list of inherently [0 content] words are the following.

particles: e.g. 'yn' (preverbal and pre-adjectival particle)  
'a' (interrogative particle in literary language)  
'nid', 'na' (negative particles)  
'fe', 'mi' (preverbal particles)

possessive adjectives (both parts) - in Welsh, these quite often consist of two separated morphs:-

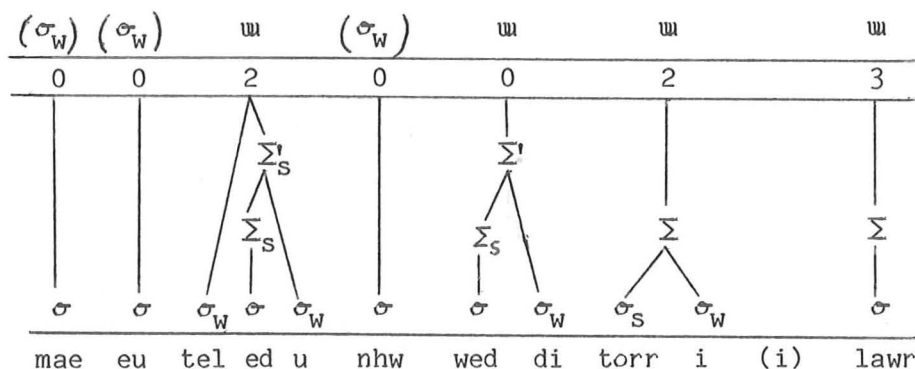
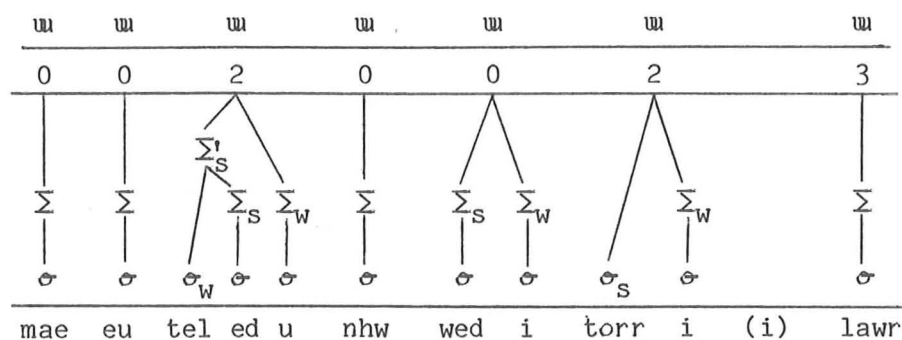
fy....i <u>my</u>	ein....ni <u>our</u>
dy....di <u>thy</u>	eich....chi <u>your</u>
ei....e <u>his</u>	eu....nhw <u>their</u>
ei....hi <u>her</u>	

These morphs flank the noun possessed, and may also be used as object pronouns.

relative pronouns/particles: 'sy', 'mai', 'a', etc.

As observed before, very many [0 content] words, especially monosyllables, cause mutation to the following lexical word. The mutation is usually lenition, i.e. the soft mutation, but the nasal mutation and spirant mutation also occur. These mutations, once phonetic in origin, have been grammaticalised and have lost their phonetic basis.

Here follows an example of a sentence, showing the semantic level and its effects, leading to the derived structure.



#### 3.8.4 Phrasal level

There has been very little work on Welsh syntax within the framework of transformational-generative grammar. Only two widely-available works are published in English. Of these, perhaps Jones and Thomas (1977) is the clearest and simplest. I have followed their proposals for syntactic structure, adapting them where necessary to form metrical structure along the lines suggested by Liberman for English - that is, allowing a kind of 'cliticization' to take place (see 3.54) to prevent too deep a nesting of the DTE and to split up longer stretches into more manageable sub-units.

For simple sentences, Jones and Thomas recognise the following syntactic constituents:-

S, NP, N, VP, V, AUX (tense inflection), PT (pre-sentential particle),

ADJUNCT (prepositional phrase, etc.), PP (prepositional phrase within the VP), ASP (aspect marker), PRED.PH (predicate phrase), COP (copula), FEL ('fel' as /vɛl/), ADJ, P (preposition).

### 3.8.4.1 Some PS rules for Welsh

Jones and Thomas set up PS rules for simple sentences, as follows.

- i)  $S \rightarrow PT \quad AUX \quad NP \quad VP \quad ADJUNCT$
- ii) 
$$VP \rightarrow (ASP) \left\{ \begin{array}{l} V \left\{ \begin{array}{l} (NP) \quad (PP) \quad (PP) \\ (PRED.PH) \end{array} \right\} \\ COP \left\{ \begin{array}{l} (PRED.PH) \\ NP \\ PP \end{array} \right\} \end{array} \right\}$$
- iii)  $PP \rightarrow P \quad NP$
- iv) 
$$AUX \rightarrow \left\{ \begin{array}{l} -ITH \\ -AI \\ -ODD \end{array} \right\} \begin{array}{l} \text{(future tense marker)} \\ \text{(imperfect tense marker)} \\ \text{(past tense marker)} \end{array}$$
- v) 
$$PRED.PH \rightarrow \left\{ \begin{array}{l} YN \\ FEL/COP\_NP \end{array} \right\} \left\{ \begin{array}{l} NP \\ ADJ \end{array} \right\} \quad ('YN' = \text{the particle 'yn'})$$

### 3.8.4.2 Some transformational rules for Welsh

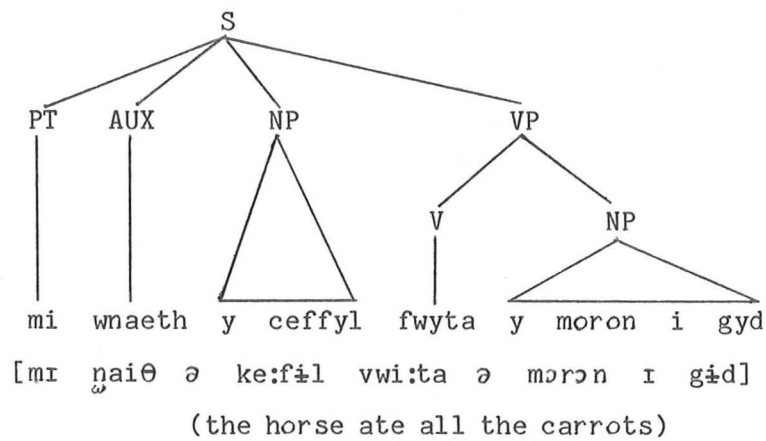
Jones and Thomas also set up four transformational rules for Welsh:-

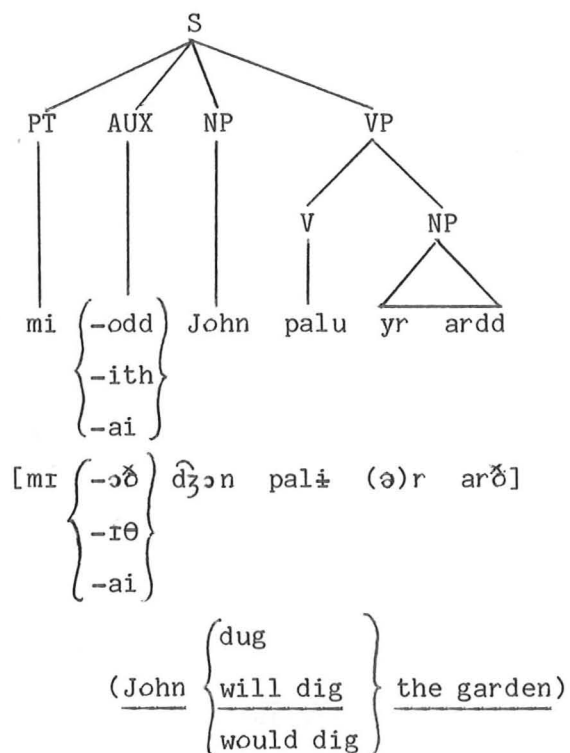
- i) The auxiliary carrier transformation moves aspectual BOD ('bod' to be /bo:d/) if selected, or the verb or copula BOD if there is no aspectual BOD, to the place of auxiliary constituent, to carry the auxiliary features in various inflected forms. The ASP constituent rewrites as one of the following: BOD YN (progressive), BOD WEDI (completive), BOD AR (about to do something), BOD NEWYDD (to have just done something).
- ii) The identificatory noun phrase inversion transformation (an obligatory rule) moves the object NP of identificatory copula sentences to initial position.

- iii) The YNA-insertion transformation applies optionally to insert YNA ('yna', there /əna/) immediately after the auxiliary carrier, in the case of locative sentences involving an indefinite subject noun phrase.
- iv) The GAN-movement transformation applies to locative sentences involving GAN (or GYDA) i.e. 'gan', 'gyda' with /gan, gəda/; the rule interchanges the GAN phrase (or GYDA phrase) with a subject noun phrase.

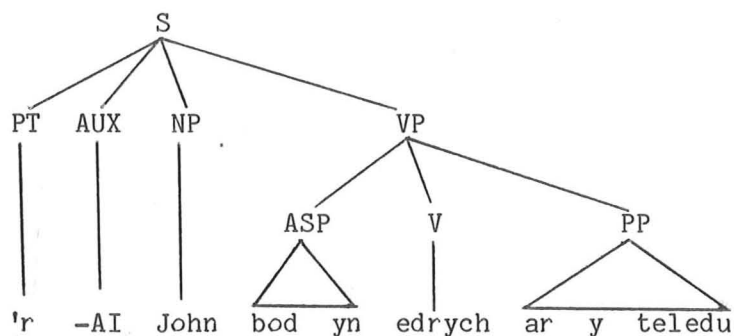
### 3.8.4.3 Examples

Jones and Thomas give some examples of the structure of a simple sentence, such as the following. ('Mi', the presentential particle, would be 'fe' in South Wales.)





The second structure has more generality, since the AUX there rewrites solely as the tense markers. The auxiliary carrier transformation then moves either the verb or the auxiliary verb 'gwneud', to do, to the AUX position.

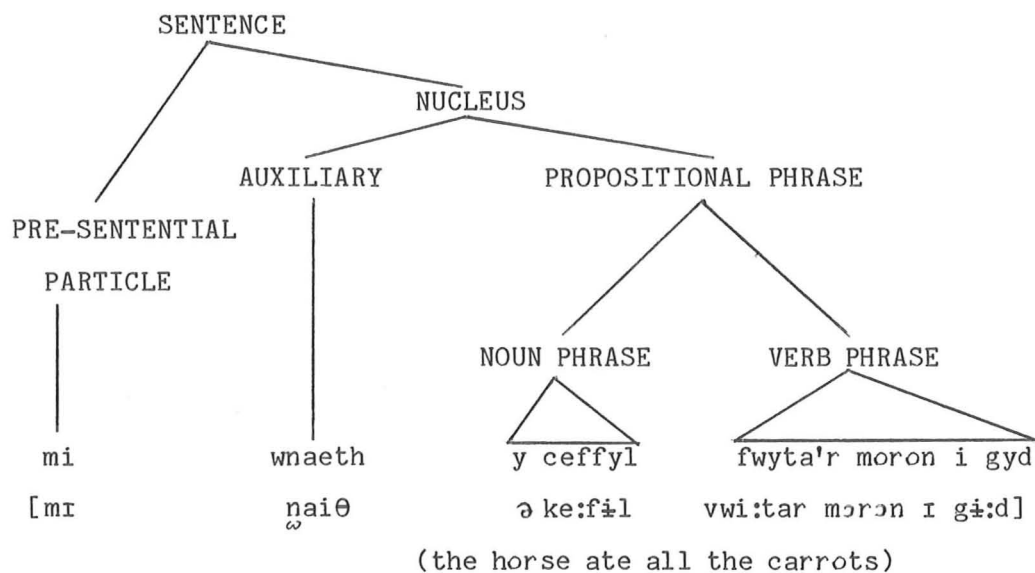


The auxiliary carrier transformation then moves BOD to the AUX position. As carrier of the AUX item -AI, BOD becomes OEDD; this process compares with the moving of the verb to the front. The derived structure is therefore as follows;-

roedd John yn edrych ar y teledu  
 [roið dʒɔn ən ɛdrix ar ə tele:di]  
 (John was watching the television)

### 3.8.4.4 An alternative syntactic structure

Jones and Thomas suggest a 'more sophisticated diagramming of these segments which reveals their relationship to the sentence and to each other....' (op. cit.:13) This is as follows:

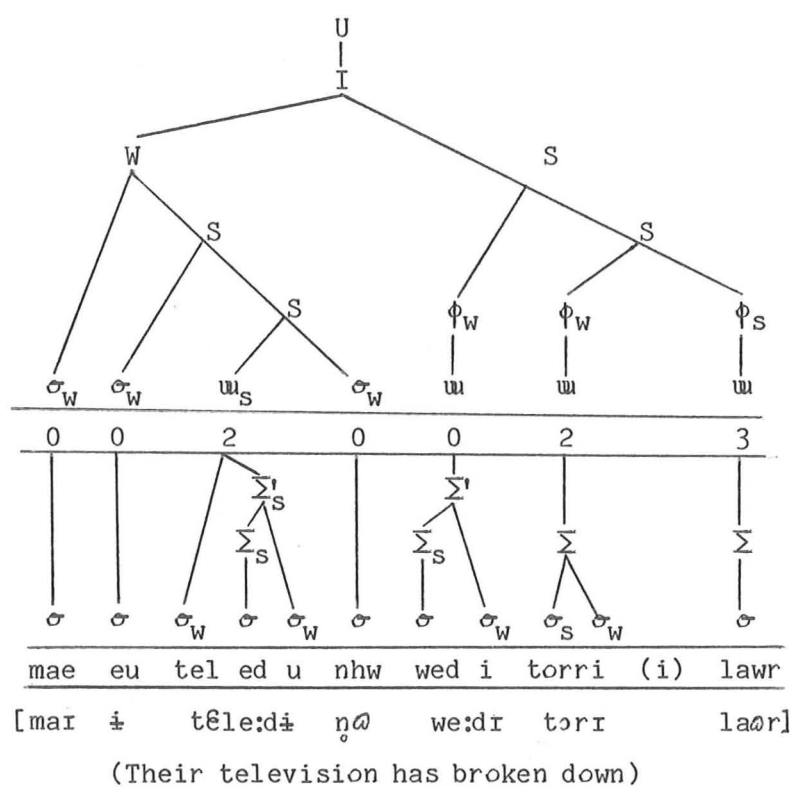


Jones and Thomas write, of this structuring: 'Although the use of a nucleus segment and a segment equivalent to a pre-sentential particle is attested in Katz and Postal (1964), the use of such a tree-diagram would involve terminological and conceptual innovations which would be outside the scope of this book.' (op. cit.:14) [8] Since their book is mainly concerned with a simpler and more general approach to the theory, Jones and Thomas retain the simpler structure defined by the PS rules cited. However, the binary-branching structure described by them as a 'more sophisticated diagramming' is more similar to a metrical tree, and it is this type of structure that I have used as the basis for the tree diagrams of Welsh metrical structure, where possible - with some modifications similar to 'cliticization', along the lines suggested by Liberman (see 354 and 3840).

[8] The reference is to Katz, J.J. and Postal, P. (1964) in 'An Integrated Theory of Linguistic Descriptions', (Cambridge, Mass.:MIT Press)

### 3.8.4.5 Prosodic structure

Selkirk's system of prosodic units may also be used for Welsh. They have the advantage of allowing a clear description of variable phrasing, and integrating with the intonational pattern. Thus an example may be given as follows, where the word-level structure is derived, rather than underlying, structure.



The prosodic units at the phrasal level are joined in the order w-s, as in English. If an attempt is made to join them in the order s-w (as suggested by the [content] values), it will be found that i) the result has very little relation to syntactic divisions such as NP, VP, and also ii) there are far more 'exceptions' to the rule than there would be under the w-s ordering, in sentences that do not have contrastive or emphatic stress.

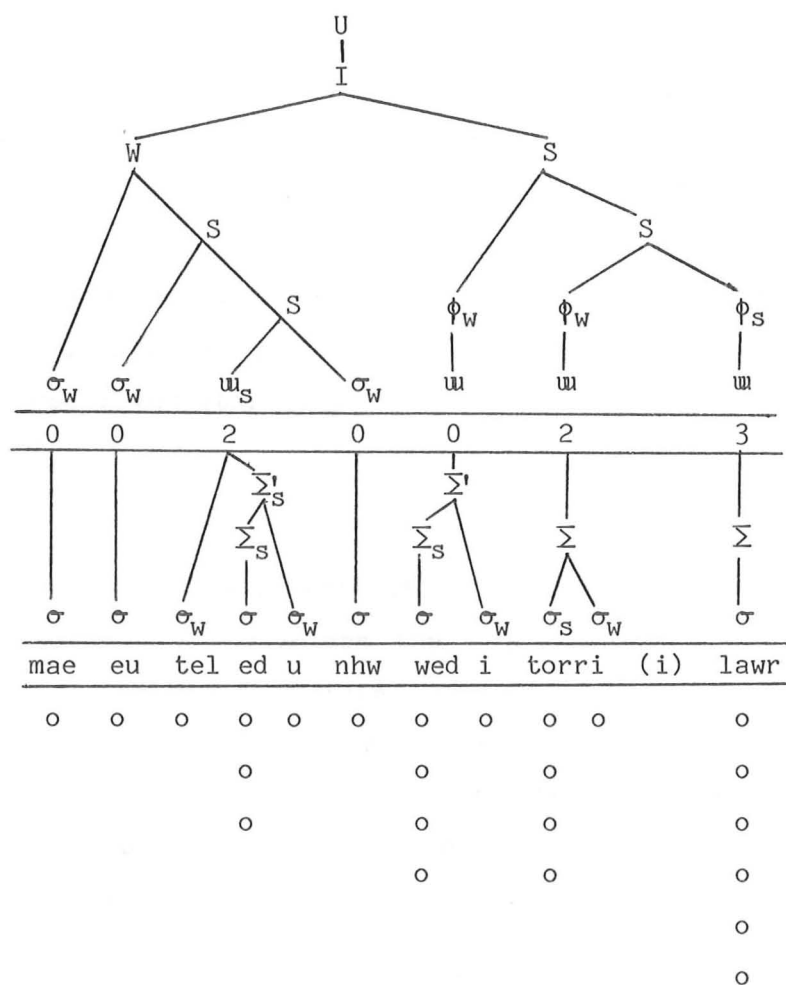
Thus the ordering w-s seems most satisfactory, both from the point of view of correspondence with (surface) syntactic structure and from the point of view of greater consistency and the beginnings of a reason for any lack of consistency. Moreover, under the s-w ordering, the first stressed syllable of the utterance would be assigned the greatest number of place-holders in the metrical grid. If the w-s ordering is used, it is the last stressed syllable of the sentence which is assigned the greatest number of place-holders in the metrical grid, and this can be seen as corresponding to the 'nucleus' of the sentence. Intonational phenomena thus support the ordering w-s, as in English.

### 3.8.5 Metrical grid

The metrical grid is defined exactly as for English, on the basis of metrical structure. The only difference is that Welsh has no [ $\pm$ full] distinction. Thus every syllable will receive a place-holder at the first level of the metrical grid, and there will be no need to use the dash, since there are no [-full] vowels.

Because of the fixed nature of Welsh stress, backshifting of stress is not possible. Clashes do occur, naturally, but the impression is that they occur far less often than in English; to the English ear, Welsh sounds more 'even' and 'flowing'. The fact that, in polysyllables, the stressed syllable is always followed by an unstressed syllable, and the fact that Welsh has a great profusion of unstressable prefixes and proclitics, may help to explain the more evenly-distributed pattern of stressed syllables in Welsh. Where stress clashes do occur, the most common pattern seems to be a slight pause, and/or a break in the intonational contour. This method of dealing with stress clashes is identical to that suggested by Liberman as an alternative method to backshifting of the stress (see 3.5.5, example xii, on 'additional grid elements'). Since Welsh has fixed stress, this is the only possible way of dealing with stress clashes in Welsh.

On the next page follows an example of the metrical grid. The prosodic structure at the word level is the derived, rather than the underlying, structure.



### 3.8.6 Intonational considerations

The same general principles may be applied to Welsh as to English, as regards fitting metrical structure to intonation pattern. The details of Welsh intonation, however, are somewhat different from those of English. The following comments are drawn from a study of the intonation patterns in Welsh language learning tapes (Williams 1980). These, it was hoped, would provide clear intonational cues and maximum exploitation of the intonational possibilities. Different sentence-types were considered: declaratives, yes-no questions, wh-questions, and imperatives. The analysis was conducted in the terms of the British school of intonation analysis, using concepts such as 'nucleus', 'tone-unit', 'head', etc. The following is a summary of the main findings, which disagreed in minor points with C. Thomas (1967), the only such study published on Welsh intonation.

### 3.8.6.1 Nucleus

The nucleus in Welsh is a similar entity to that in English, although the high rising nucleus does not seem to be systematically present in Welsh. The rise-fall is perhaps more common in Welsh than in English. Unstressed syllables in the tail may exhibit a slight glide which is, however, not pitch-prominent. This is made possible by the fact that the 'unstressed' final syllables of polysyllables are actually longer and louder than the 'stressed' penults.

### 3.8.6.2 Head

The head in Welsh intonation is very different from that in English. To begin with, while Crystal describes the 'unmarked pitch level' in English as a slight drop in pitch (Crystal 1969:144), the 'unmarked pitch level' for Welsh seems to be a slight rise in pitch from one syllable to the next. In pre-nuclear position, this is by far the most frequent type of pitch relationship between successive syllables. Thus a slight step-up in pitch, in Welsh, is seen as being not pitch-prominent - any other change in 'pitch level' will be pitch-prominent. Whereas Crystal postulates different norms of pitch movement for stressed syllables and unstressed syllables, Welsh seems to require no such division in the definition of pitch-prominence. It is only when there is no preceding contour to refer to that such 'absolute' measures of pitch-prominence are needed.

The head in Welsh has a basic pattern referred to by Ceinwen Thomas (1967) as the 'saw-toothed' head. This pattern consists of a sequence of stressed syllables each slightly lower in pitch than the preceding stressed syllable. The intervening unstressed syllables rise slightly from the level of the preceding stressed syllable, giving an effect as in the following:

'ers 'pryd 'rydych chi 'wedi 'bod yn 'dysgu Cymvraeg?//



[ɛrs prɪ:d rɛdɪx xɪ we:di bo:d ən dəsɡɪ kəmraɪɡ]

(How long have you been learning Welsh?)

A slightly less common type of head is one in which the stressed syllables rise gently, including the intervening unstressed syllables, to produce a steady rise in pitch over the whole sequence. This means that only the first stressed syllable is pitch-prominent, given the definition of pitch-prominence in Welsh adopted above. An example of such a 'rising' head is the following, where there is a pitch glide on the unstressed final syllable of the last disyllable.

'Beth ydy ei ^henw hi?//

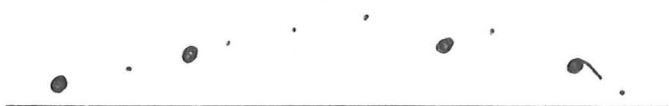


[be:θ ədɪ i hɛnə hr]

(What is her name?)

The 'unmarked' and 'rising' heads can occur in the same tone-unit, in which case the rising sequence usually precedes the drop which begins the 'unmarked' head contour:

'Rydw i wedi bod yn 'prynu hadau//

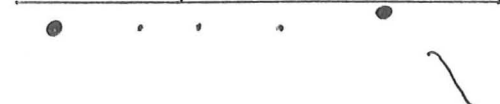


[rədwr we:dr bo:d ən prɛni ha:də]

(I've been buying seeds)

In a few cases, the head contour remains at the same pitch level throughout. Any following head contour may take either the ascending or the level form. The level head tends to be fairly high-pitched, and is usually seen in questions, such as the following:

'Beth rydych chi'n feddwl?//



[be:θ rədɪx xɪn ve:ðəl]

(What do you think?)

It will be seen that the 'head' in Welsh shows great variety. As Ceinwen Thomas observes: '...there seems to be far less predictability about the pattern of this part of the tune in Welsh intonation than is the case in English intonation.' (op. cit.:19) Because of this, it is more convenient to deal with each head contour separately (where 'head contour' refers to a stretch beginning with a stressed and pitch-prominent syllable, and including any following non-pitch-prominent syllables, whether stressed or unstressed, up to but not including the next stressed and pitch-prominent syllable). A head contour may be seen as corresponding to a 'phonological phrase' in metrical structure.

### 3.8.6.3 Prehead

The prehead in Welsh comprises the syllables at the start of the utterance up to, but not including, the first stressed and pitch-prominent syllable, as in English. The most usual type of prehead in Welsh begins at a medium pitch level and rises gently; i.e. the following syllables are not pitch-prominent. In some cases, the prehead may be a level sequence at mid-voice range, and therefore also non-pitch-prominent. Other types of prehead are: the low level prehead, where the level sequence is at a low pitch level, and the high rising prehead, rising gently from a fairly high pitch level.

### 3.8.6.4 Tonetic stress marks

Tonetic stress marks for Welsh are summarised below.

<u>nucleus</u>				<u>head</u>	<u>prehead</u>		
/s	rise	^s	rise-fall	's	unmarked	s	unmarked
/s	low fall	vs	fall-rise	's.s	rising	→s	(mid) level
\s	high fall			-s	level	_s	low level
(no high rise nucleus)						°s	high rising

s = syllable;    ·s = stressed but non-pitch-prominent syllable

s = sudden unsystematic step-up in pitch (rare)

// = boundary of major tone-unit (/ = of minor tone-unit)

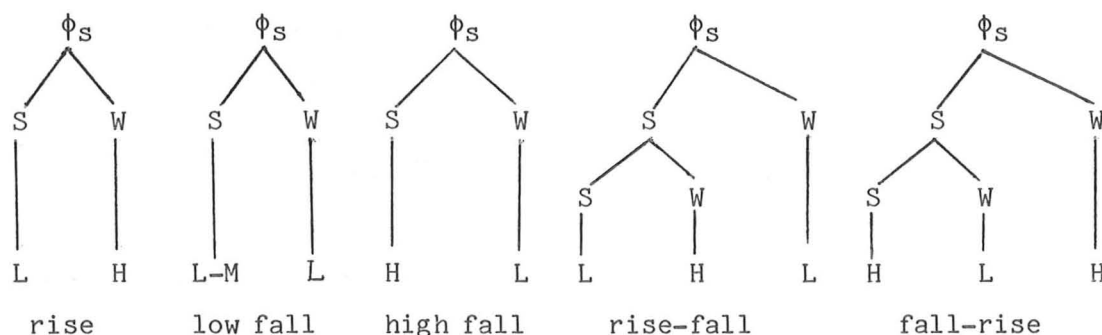
### 3.8.6.5 A Liberman-style analysis

Liberman's approach to intonation analysis may be extended to Welsh, with appropriate modifications. The most important difference is that since head contours in Welsh are less predictable, some 'tunes' will apply not to the utterance as a whole, but to some lower-level prosodic unit - that is, the phonological phrase. The relevant distinction is between  $\phi$  units that are the DTE of an intonational phrase, and those that are not, since different 'tunes' apply to each kind. The intonation pattern in the head seems to bear little, if any, relationship to the type of nuclear tone chosen; unlike English, which has certain restrictions on the co-occurrence of head types and nuclear tone types, according to O'Connor and Arnold (1961). Thus different 'tunes' are appropriate for the different structures: head contour (or  $\phi_w$  unit) and nucleus (or  $\phi_s$  unit that is the DTE of an I unit). The different types of prehead correspond to different types of 'boundary tone' - as Liberman writes:

'Notationally, such boundary tones will be represented as  $\frac{B}{T}$ . For the purpose of tune-text association, it will be assumed that corresponding boundaries exist in the text, and that CONGRUENCE requires a matching of boundary to boundary. In the configuration  $\frac{B}{T}$ , the tone T is a terminal element, while the symbol B is a non-terminal. For purposes of grid alignment, the B positions are (redundantly) weak. In order to include these boundary elements in our metrical trees, non-boundary material will be called 'content', symbolized by the node label C.' (Liberman 1975:77)

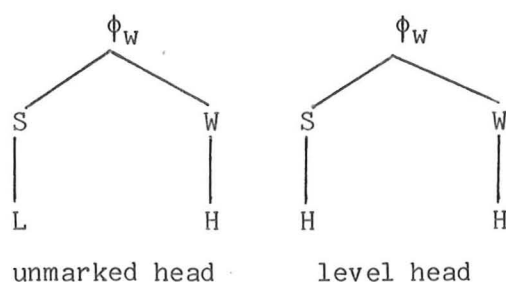
The appropriate 'tunes' for Welsh are presented below.

#### DTE of I unit (i.e. nuclear tones)

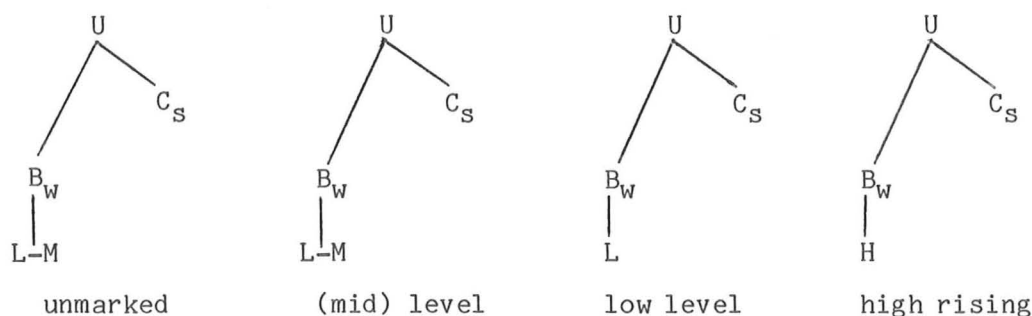


L = low; H = high; L-M = low-mid, also H-M = high-mid

### Non-DTE of I unit (i.e. head contours)



### Boundary tones (i.e. preheads)



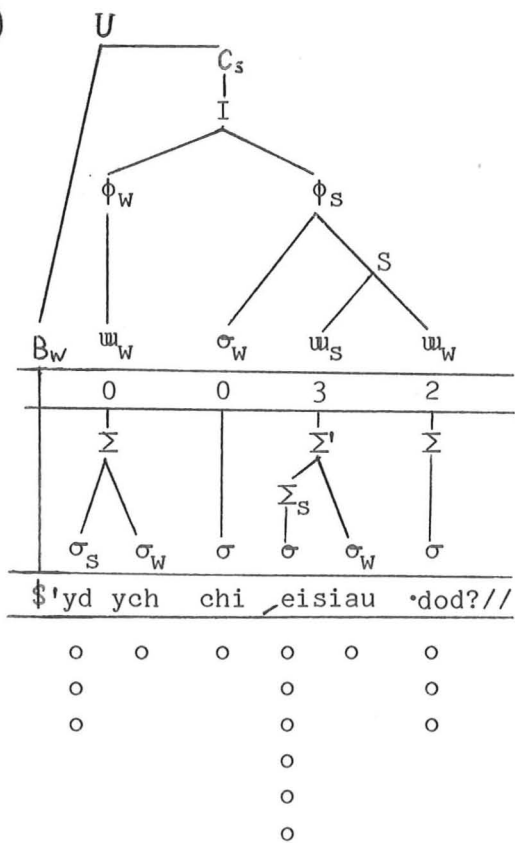
### 3.8.6.6 Examples

Some examples of this system are given on the following pages. All the examples are taken from the materials studied for Williams (1980).

In example ii), the stressed syllable of 'oren' is not the DTE of a  $\phi$  unit, and so cannot be pitch-prominent. Thus it becomes a stressed syllable in the tail of the nucleus. This is an example of a Welsh sentence with contrastive stress. The symbol '\$', following Liberman, is taken to represent the boundary in the text, and is necessary for establishing congruence between tune and text. The 'boundary tone' chosen here is the unmarked type, rising gently from a medium pitch. The nucleus is a rise-fall, and the head is the unmarked type.

Example iv) is spoken in a very emphatic style, represented in the phrasing as a proliferation of I units. This leads to the assignment of more than one tone to certain stressed syllables, and corresponding glides on nuclear syllables. Such a style of utterance was rare among the utterances collected from language learning tapes; most had a maximum of two I units.

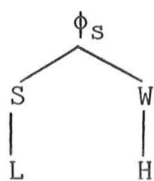
i)



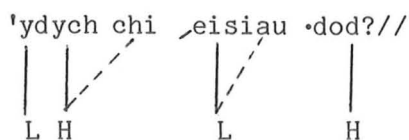
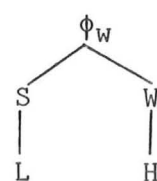
[ədix xi əi{ɛ do:d]

(Do you want to come?)

DTE of I:



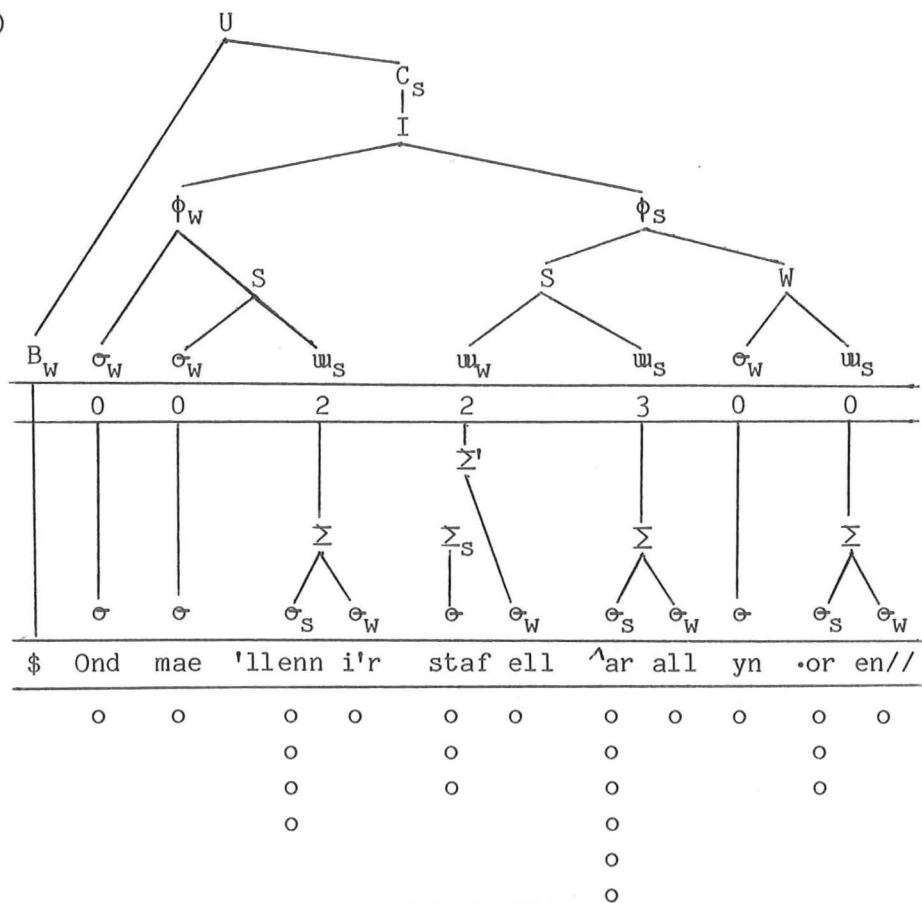
non-DTE of I:



'ydych chi eisiau dod?//



ii)



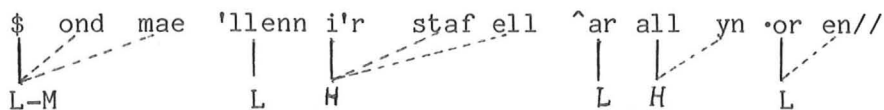
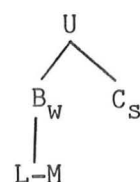
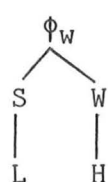
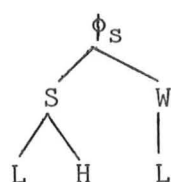
[ənd maɪ ˈtʃɛnrɪ stɑːvɛɪ ərəʃ ən ɔrɛŋ]

(But the curtains of the OTHER room are orange)

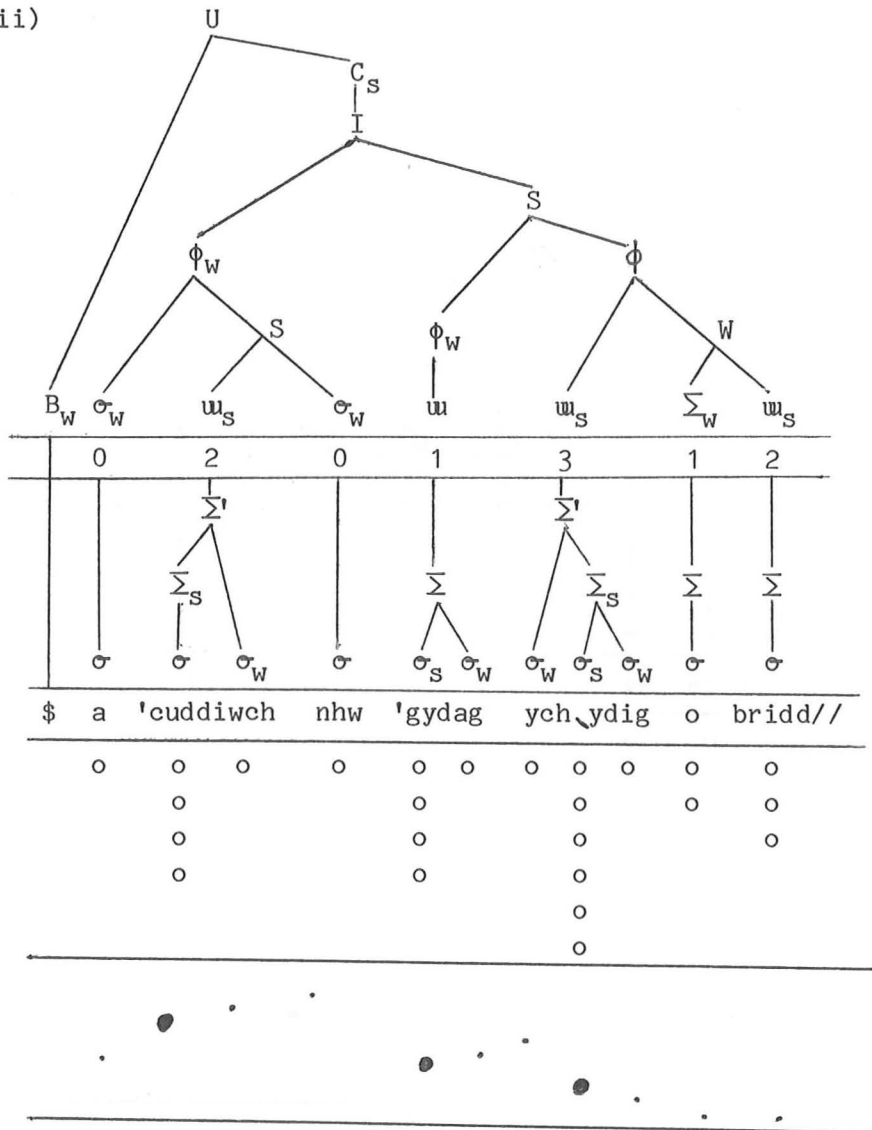
DTE of I:

non-DTE of I:

boundary tone:



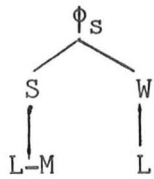
iii)



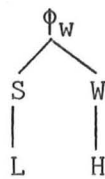
[a ki:ʔjəx n̩ ɡədag əxədɪg ɔ bri:ʔ]

(....and cover them with a little soil)

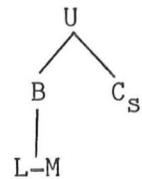
DTE of I:



non-DTE of I:

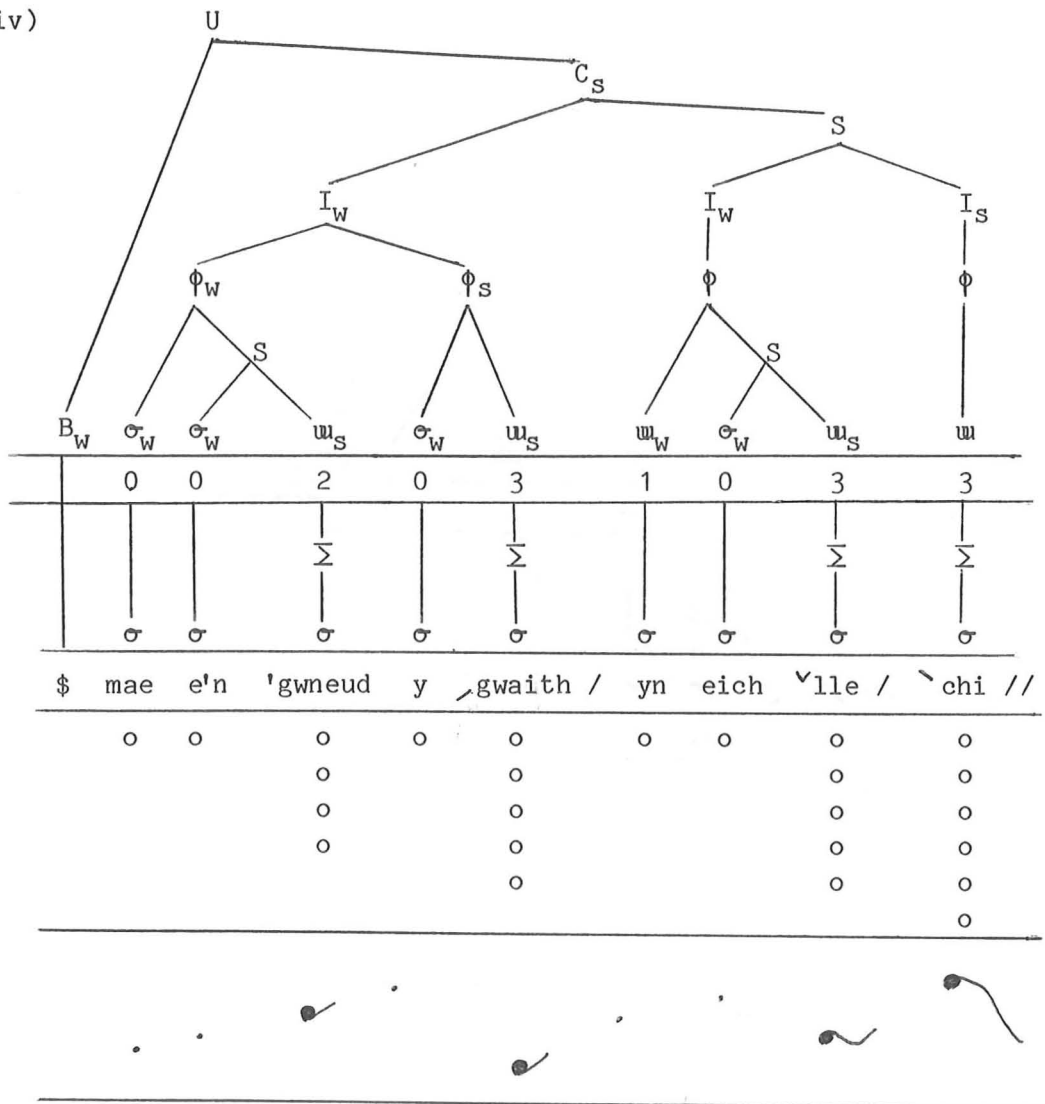


boundary tone:



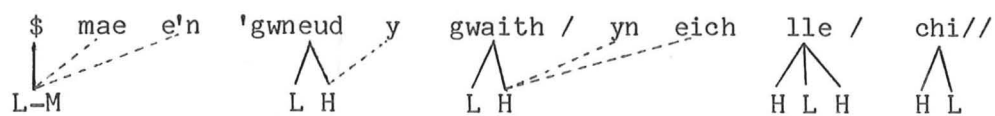
\$ a 'cuddiwch nhw gydag ych ydig o bridd//  
L-M L H L H L-M L

iv)



[mai ɛn gneɪd ə gwaɪθ ən ɪx ʔe: xi:]

(He's doing the work instead of you)





## 4. DISCUSSION

### 4.1 Penultimate stress in Welsh

The modern Welsh pattern of stress on the penult can be realised by two types of mechanism. Both phonetic and phonological phenomena add prominence to the penult, as can be seen from the following.

#### **4.1.1 Phonetic mechanisms**

##### 4.1.1.1. Post-stress consonants

The 'extra duration' of the post-stress consonant (in the case of plosives) was noted by R.O. Jones (1967) in a purely auditory study. That this extra duration is also present on the acoustic level, however, is seen by the results of the measurements in section 2.2.4.1., while the results of the perceptual experiments in section 2.3. show that this cue is valid in the perception of stress.

The post-stress consonant's greater length favours the penultimate syllable in that it may partially make good the syllable length lost when the vowel of the penult is short. It may also indicate greater muscular effort, in that, in a continually-varying stream of speech sounds, it would require a little more attention and effort to hold an obstruent steady rather than to submit to the default timing and articulate it rapidly. This, however, is only speculation. The post-stress consonant may also contribute to the prominence of the penult in the way outlined in section 4.2.1. - i.e. 'interval hearing'.

##### 4.1.1.2. Tendency towards isochrony

Another phonetic phenomenon adding prominence to the penult is rhythm, or rather the tendency towards it. It has been seen (in section 2.2.2.) that the acoustically weak penult plays a crucial part in rhythmic

structure, unlike the acoustically strong ultima. If this were the main cue to stress in Welsh, the lack of inherent acoustic cues in the penult would then not matter, since the strategy of stress realisation would rest on a different basis entirely. This indeed seems to be the case, for no other strategy for stress realisation has been found. Also, it is hoped, the statistics provided in the course of this investigation go some way towards answering Buxton's complaint that descriptions of temporal structure in speech 'have been limited to approximate claims about a "tendency towards isochrony"' (Buxton 1983:119).

#### 4.1.1.3. Lack of inherent cues

It will be noted that the above two phonetic phenomena are both 'relational' in nature rather than 'inherent', or 'syntagmatic' rather than 'paradigmatic'. They can only exist in conditions where the utterance as a whole builds up a norm of timing (for consonants) or a rhythmic structure (for vowels) which allows the relevant units to be characterised as marked and therefore stressed. They could not exist 'in citation form'; that is, divorced from the context of other units of a similar type in connected discourse.

This is in contrast to the situation in English, where the stressed syllable is most often (but not always) pitch-prominent also. Pitch-prominence, as defined in terms of a pitch glide within one syllable, or very high or low pitch, can be regarded as paradigmatic. This is because such features are readily recognisable in themselves, without the need for reference to be made to neighbouring units of a similar type in order to estimate the 'norm'. Pitch-prominence as defined in terms of obtrusion from an intonational contour may be more readily seen as syntagmatic, but even here there is a difference between English and Welsh: while pitch obtrusion may be equally well a step up or down in pitch, the stressed penult of Welsh differs in only one direction, in that its vowel is shorter (when phonologically short), and, unlike that of the ultima, is often reduced to schwa or even lost altogether. Thus penultimate stress in Welsh seems to be linked to 'relational' or 'syntagmatic' phonetic strategies for realising stress, while stress in English, though making use of such strategies (as is seen from the tendency to subjective

isochrony found by past workers - see section 1.2.5.2.) also makes use of 'inherent' or 'paradigmatic' cues of pitch, intensity, etc., which are not reliable stress cues for Welsh.

#### 4.1.2 Phonological mechanisms

##### 4.1.2.1. Most recent sound changes

Phonological phenomena can also add prominence to the penult. One such phenomenon is that of modern sound changes. For example, in rapid colloquial contemporary Welsh, final /aɨ/ has become shortened to /ɛ/, while final /v/ has been lost altogether in many cases. Final /aɨ/ is a common plural morpheme, seen in such words as llyfrau (now pronounced /l̪ɛvrɛ/ rather than /l̪ɛvraɨ/), and so this development is widespread. Final /v/ was often lost historically, but is now dropping not only in nouns but also in both verbs (where it originally formed part of the inflection /av/ for the first person singular future form), and in the superlative forms of adjectives (denoted by the ending /av/, now pronounced /a/). These are examples of recent developments that reduce the ultima and thus lend prominence to the penult.

##### 4.1.2.2. Vowel length rules

Phonological vowel length is distinctive only in stressed syllables (penults and stressed monosyllables). In other contexts the distinction between long and short vowels is neutralised and all vowels are seen as phonologically short, whatever their phonetic duration (which may be considerable in the case of ultimas) - see Rhys Jones 1977 and Williams (in press) for summaries of Welsh vowel length rules. In general, vowels are short (in stressed syllables) when followed by two or more consonants or by one voiceless stop, nasal, or liquid, within the same syllable. Vowels are long when syllable-final or before one voiced stop or one fricative of any kind. Schwa is always short, but never occurs in the ultima. It will be noted that the rules of phonological vowel length in Welsh give prominence to the penult, since this is the only syllable of polysyllables in which vowel length is distinctive.

#### 4.1.2.3. Vowel mutation

Vowel mutation is the process in Welsh by which certain non-final vowels undergo reduction to schwa by morphological alternation. This process applies to original /u/ (orthographically 'w') and to those cases of original /i/ that are orthographically 'y'. The first element of the diphthong /ai/ may also be reduced to schwa in fast speech. (For a detailed consideration of vowel mutation, see Allen, M.R. 1975, Thomas, A. 1979, and Williams, in press). For the present purposes it will be enough to note that this process, unlike the two mentioned above, gives prominence to the ultima by reducing non-final vowels in certain contexts - even the supposedly stressed penult may be reduced, while the supposedly unstressed ultima is immune to reduction. This situation seems to go against all that we know about the workings of stress from other languages; an attempted explanation is outlined below in section 4.4.3. It is enough to note for now that the penult is not (yet?) completely dominant in the pattern of stress in Welsh.

#### 4.1.3 Type of descriptive theory needed

##### 4.1.3.1. Welsh

There is a choice of three main types of theory about stress in the attempt to characterise Welsh stress. The first, and more traditional, could be termed the 'inherent cues' type. This comprises both the 'British' and 'American' approaches, and views stress as a complex of pitch-prominence, greater intensity, longer duration, and more peripheral vowel quality. Such a view has been dominant for many years. The second type is Bolinger's theory of stress as 'accent', defined solely on the basis of pitch-prominence. This type also has 'inherent' or 'paradigmatic' elements, in that pitch-prominence can in some cases exist by virtue of the intrinsic cue of a pitch glide on one syllable. On the other hand, it also has 'syntagmatic' elements, in that pitch-prominence may exist by virtue of the relational cue of a difference in pitch between two syllables in a particular direction - the pitch-prominence would not exist were only the one syllable in question to be present. The third type of theory, like the first, has no such mixture of elements: the 'metrical'

theory of stress views stress as purely a syntagmatic phenomenon, defined in terms of the relational cue of rhythm. Under this view, the precise manner in which the pulse of the rhythm is marked is a purely secondary consideration; it is the rhythm that defines the stress rather than vice-versa.

In the case of Welsh, it is obvious that the first type of theory will not do. There are no consistent stress cues of pitch, intensity, duration, and vowel quality - even the 'negative' cue of shorter vowel duration is not wholly consistent. Even restricting the cue to pitch-prominence, as in the second type of theory, will not give satisfactory results; two minimal-pair members with the same pitch-pattern can be interpreted as having stress on different syllables according to segmental cues of duration. Even where pitch-prominence has an influence on stress judgements, this is done not by direct mapping of pitch-prominent syllables with stressed syllables, but rather by indirect mapping via the possible intonational categories of Welsh. Even here, however, there is no very high degree of consistency, since the shift of judgements is by no means total and other factors exert an influence. Therefore it seems best to adopt the third type of theory for stress in Welsh, viewing the latter as based on rhythm alone and only contingently coinciding with pitch-prominence under certain conditions.

In order for rhythm to occur, a stretch of time must be marked off into roughly equal intervals. The nature of this marking-off is not a crucial issue: in some languages, such as English, the marking-off may be done by intensity, etc. In Welsh, there are no obvious intrinsic cues to provide the 'pulse' of the rhythm. Instead, what seems to happen is that the shorter vowel of the stressed penult, and the longer post-stress consonant, together form a reference point for the demarcation of rhythmic feet: great intensity is not necessary for such a purpose. Once set up thus, the (subjective) rhythmic pattern is continued by its own momentum, despite many objective irregularities, and is reinforced by perceived modifications in segmental duration in the direction of isochrony.

A further complication is introduced in the case of stressed monosyllables. Though not separately examined in this study, these show a clear pattern of long duration (comparable to that of final syllables), high intensity (as in final syllables), lack of schwa (as in final syllables), the possibility of a length contrast (as in penults), and the possibility of pitch glides (as mostly occurs on final syllables) and pitch obtrusion (as in both penults and final syllables). Stressed monosyllables were counted as stressed in all measurements of feet in the experiments described in chapter 2. These measurements showed a tendency to isochrony underlying the 'syllable-timed' tendency, even when stressed monosyllables had been thus included. The answer to the puzzle may be that stressed monosyllables have prominence in terms both of rhythmic prominence and of intrinsic acoustic prominence, while stressed penults have only rhythmic prominence and ultimas, though unstressed, have intrinsic acoustic prominence (see section 4.4.3. for further consideration of this point).

#### 4.1.3.2. English

While there is experimental evidence supporting the use of the 'metrical' theory to describe stress in Welsh, as has been seen, this is not the case for English. Isochrony (mostly of a subjective kind) has been noted as a tendency in English for many years, and thus there is some basis for applying the metrical theory to English also. The main difference between English and Welsh in this respect is that the rhythmic 'pulse' in English (marking off the rhythmic units) is provided by vowels with pitch-prominence, greater intensity and duration, and unreduced vowel quality. In Welsh, the pulse is provided by a combination of shorter duration of the vowel and longer duration of the following consonant. The situation in English seems similar to that of stressed monosyllables in Welsh, where both rhythmic and intrinsic cues are the exponents of stress.

However, this is not to bring back inherent cues by the back door for English. While it may be true that these cues can often be mapped directly to stressed syllables, it is also true that such a process is not completely accurate; there is always a certain amount of inconsistency. This is where subjective isochrony comes into its own; stress can be

linked, not with the actual occurrences of an intrinsic cue, but with the rhythmic pattern set up by these cues and continuing over those stretches where intrinsic cues fail to give consistent results. A syllable heard as stressed in English may not show a clear pattern of intrinsic stress cues, but it may well fit into the rhythmic pattern set up by previous syllables that did show greater intensity, duration, etc. This is the reason why one cannot stop at intrinsic cues as the primary means of realising stress. It is the construct they set up - the sentence rhythm - which is to be seen as the basis for stress, however few the cases may be where this construct is the sole exponent of stress.

It may be objected by some that the construct is essentially a psychological entity, undiscoverable in the acoustic form of speech. Language, however, is used by human beings and thus is filtered through a human mind; to refuse the psychological the status of 'real' is to deny any validity to what happens in the human mind and to reduce all language to the level of involuntary nervous reflex. Since here we are dealing with a psychological entity, it is necessary to take account of appropriate research in the psychology of perception. The next section gives details of one piece of experimental work in this area.

#### 4.2 Some work on the perception of stress

##### **4.2.1 Interval hearing**

A. Bell (1977) carried out an experiment with speakers of different languages which had differing stress rules. The subjects were asked to make judgements of stress placement when presented with pure tones varied along certain dimensions. It was found that a language's system of stress placement had no significant influence on the speaker's perception of stress patterns. The most important factor was found to be the 'sensory dimension of prominence'; i.e. 'prominence' in terms of  $F_0$  or increased amplitude or duration. There was a distinct preference, for all subjects, for marking their perceptions of the tones as being in groups of three with initial stress - every third tone had been made more 'prominent' along

one acoustic dimension in the experimental tone sequences.

A major exception, however, occurred in the case of the 'length sequence', where every third tone was longer in duration. In 48% of cases, there was a preference for groups with medial accent in the transcriptions of the rhythmic groups perceived by the subjects (as against 36% for initial stress and 16% for final stress). Bell attempts an explanation of this as follows.

'What is surprising is that almost half of the responses indicated a rhythmic group with the longer sound in medial position (-/-). This is probably related to the fact that the longer sound was followed by a correspondingly short interval, since the repetition rate was held constant. The longer sound and its successor would then tend to be perceived within a rhythmic group. In Handel's terms the sequence is composed simultaneously of a length pattern ... \_ \_ \_ ... and an interval pattern ...x xx x xx x... The results can then be explained thus: Some listeners perceived the length pattern, preferring the rhythmic group -vv to vv-. Others perceived the interval pattern, preferring the rhythmic group x xx to xx x. This assumes that listeners were able to identify the interval organization x xx with the notation -/-'

What this means is that the combination 'short tone/long gap followed by long tone/short gap' was perceived by half the subjects as marking the beginning of a rhythmic foot, while the 'accent' was placed in the middle of this foot. From this it is but a short step to the pattern seen in Welsh polysyllables of 'shorter vowel/longer consonant followed by longer vowel/shorter consonant', with the most acoustically prominent vowel placed in the middle of the foot thus defined. If it is this 'interval hearing' that is involved in the case of Welsh stressed penults, it becomes clear that our view of stress must be revised somewhat.

In Bell's experiment, there was no suggestion that any subject attributed any kind of prominence to the short tone beginning each unit. In Welsh, the prominence of the penult resides in the phenomena of post-stress consonant length, isochrony, recent sound changes, and vowel length rules described in sections 4.1.1. and 4.1.2. above. The penult also enjoys a certain amount of importance from beginning the rhythmic unit and forming the basis of the tendency towards isochrony. The acoustically prominent ultima is situated after the start of the foot, as was the

lengthened tone in Bell's experiment. A difference between the two situations is seen in the fact that the longer tone was marked as stressed by Bell's subjects, while the Welsh ultima is not stressed. However, in terms of such phenomena as intrinsic acoustic prominence, lack of vowel reduction, and vowel mutation, the ultima may clearly be said to have some degree of prominence, while not being stressed. In fact, at an earlier stage in the development of Welsh, the ultima was the stressed syllable in polysyllables (see section 4.3.1. for details). For the present, it is sufficient to note that the 'length pattern' mode of perception would characterise the ultima as the stressed syllable beginning a foot, while the 'interval pattern' mode would characterise the ultima as a prominent syllable within a foot begun by the non-acoustically-prominent penult.

#### **4.2.2 Subjective versus objective**

##### **4.2.2.1. For all sensory input**

Most studies of stress, for any language, have concentrated on identifying those features of speech production present in the acoustic signal that can be related directly to the mental perception of stress in the mind. This has led to the tendency to seek inherent acoustic cues of particular types which are then described as cues to stress that can be mapped one-to-one with stress perception.

Such a view is both mechanical and unnecessary - mechanical, because it ignores the large part played by the human mind in the actual use of language; and unnecessary, because such a disregard for psychology is not needed in order to arrive at a watertight description of the phenomenon. It is not obvious why we should need to prove a one-to-one relation between the subjective and the objective aspects of stress; and indeed, such an insistence may even be harmful to our purposes, in that it could be leading the focus of research away from the true situation.

In the case of hearing, seeing, etc. in general, what the brain perceives is very often dissimilar to the physical facts as measured by a machine. In the case of seeing, this is demonstrated by cleverly-planned 'optical illusions'. In the case of heat perception, this is demonstrated by the old game of placing one arm in a bowl of hot water and the other in a bowl

of cold water; if both are then placed together in a bowl of lukewarm water, the first arm will feel cold and the second one warm. In the case of hearing, this is demonstrated by the way in which the ear 'telescopes' frequencies so that, at very high frequencies, a far greater change in frequency is needed to produce a given perceived pitch change than is needed at lower frequencies. It is also well known that our perception of pitch, loudness and length are not directly determined by  $F_0$ , intensity and duration respectively, but by a combination of the main cue with the others. The phenomenon of 'subjective rhythm', when the mind imposes a rhythm where there is none objectively, is also seen in sensory input of all kinds; for example, the chugging of a train or the ticking of a clock (see Allen, G.D. 1975 for a discussion of work on subjective rhythm).

Thus it is clear that, even without narrowing the field to the more specialised case of speech input, what the brain perceives need not be directly related to the physical facts of the situation. In fact, where there is a discrepancy, this is usually in the interests of greater efficiency of perception. For example, most sounds that are important to us (such as speech sounds) lie in a restricted range of frequencies, and it is here that our frequency discrimination needs to be at its sharpest. For very high and very low frequencies, the blunting of discrimination ensures that no unnecessary effort is expended in making fine discriminations where they are not needed. Similarly, sounds at very high and very low frequencies need to have a much greater degree of intensity to be heard at the same loudness level as sounds of medium frequency, which are those most often encountered.

#### 4.2.2.2. For speech input

Having established the lack of a consistent direct relationship between acoustic facts and perceived sounds in the case of sound input in general, it now remains to consider this relationship in the subset of sound input that is speech input. Here again, the brain seems to act as a 'filter', imposing a slightly different structure on the raw acoustic input.

Research on the perception of phonemes carried out by A. Liberman et al. (1963) showed that what is acoustically a series of equal step changes in locus frequency of the second formant transition is perceived in category (phoneme) terms by native speakers. Also, they found that discrimination is greatest where it is most needed, at phoneme boundaries, and that it is least acute where there is least need for it, in the middle of the phoneme category.

Similarly, the perception of rhythm also shows a bias in the case of speech stimuli as against non-speech stimuli. When listeners are presented with two sequences that differ in objective duration, whether the difference is perceived depends first of all on whether or not the stimulus is speech, since the difference limen for speech in this case is higher than that for non-speech stimuli. Thus it is more difficult to perceive a durational difference in the case of speech. Where the stimulus is speech and the objective difference is greater even than this higher threshold, a further safeguard is built in; listeners tend to impose subjective rhythm in the sense, here, of equalising dissimilar durations. Even if this fails, the duration difference may still go unperceived if a rhythmic momentum has been built up previously (see Lehiste 1977 for details of the perception of duration). Thus there is a bias towards the perception of speech as isochronous which ensures that quite large discrepancies on the objective level are simply not perceived.

The conclusion is that a simple equating of acoustic cues with auditory percepts is inadequate, especially in the case of rhythm. The acoustic tendency towards isochrony is perhaps better seen as a cue to a psychological entity of subjective rhythm that is only partially realised at the objective level. This is not to decry the importance of the physical facts, but simply to deny them primacy in the definition of stress and to deny any simple relation between the subjective and the objective aspects of stress.

#### 4.2.3 Function of cues

Given such a situation, one may well be forgiven for wondering what part the inherent acoustic cues present in English actually play in the realisation of stress. One answer is given by Ladd (1980), whose formulation (cited in section 1.3.3.) is repeated here for convenience. He states that such acoustic measurements as  $F_0$ , duration, intensity, etc.,

'...should not be interpreted as demonstrating the acoustic correlates of a particular stress peak, but rather as indicating the sorts of acoustic features which induce us to hear an overall rhythmic organization which includes the particular stress peak.'  
(Ladd 1980:26)

In other words, these cues trigger the realisation of a rhythmic pattern, in terms of which the stresses are then fitted in. In a sense, this strategy saves effort, in that the listener need not constantly be 'on the alert' for an inherently prominent syllable, but has his or her expectations primed by the perceived rhythm so that less attention need be paid at those points where the rhythm does not predict the occurrence of stress. An analogy may be drawn between this situation and that of pushing a child on a swing, where the pusher does not need to direct the motion of the swing continuously, but simply gives a short push at one particular point. Similarly, when ringing a tower bell in the English manner, it is not necessary to use great strength to control the swinging of the bell in a vertical circle; all that is needed is a short pull at the exact moment when the bell is almost at rest at the end of one circle, and gravity and momentum take care of the rest of that swing. In the same way, the speaker of a stress-timed language needs only to insert into the signal certain judiciously-timed cues, and the listener's rhythmic perception will supply the remainder, even making good the lack when a cue has not been supplied. Thus rhythm could be seen as functioning as a strategy for improving the efficiency of communication, by reducing effort for both speaker and hearer.

### 4.3 Historical development of stress in Welsh

#### 4.3.1 The Old Welsh Accent Shift

##### 4.3.1.1. Sound changes before the accent shift

At some point in the early stages of its development, Welsh underwent the Old Welsh Accent Shift. This involved the shifting of word stress from the ultima to the penult in polysyllables, while stressed monosyllables remained stressed. The Welsh ultima had been the (stressed) penult in the parent language, British, but with the loss of word-endings and inflections that characterised the change from British to Primitive Welsh (in the late fifth to early sixth centuries A.D.), this syllable became the new ultima and retained its stress.

The date of the later accent shift from ultima to penult is in some dispute; Kenneth Jackson (1953) places it in the mid-eleventh century, i.e. just before the change from Old Welsh to Middle Welsh, while Arwyn Watkins (1972a, 1972b) dates it much earlier, in the mid-ninth century. However, all scholars are agreed that the shift occurred, using as evidence various sound changes attested in manuscripts and inscriptions. The argument hinges on the fact that the earliest sound changes distinguish between final syllables and stressed monosyllables on the one hand, and non-final syllables and unstressed monosyllables on the other. Later sound changes, however, distinguish between final syllables and unstressed monosyllables on the one hand, and non-final syllables and stressed monosyllables on the other. The following sound changes date from before the accent shift.

- a) In ultimas and stressed monosyllables, long /ɔ:/ diphthongised to /aə/, in the fifth to sixth centuries A.D.- e.g. British \*cāseus, 'cheese' > Primitive Welsh /kɔ:s/ > Middle Welsh caws, /kaəs/.
- b) In ultimas and stressed monosyllables, certain diphthongs became more open in their first element - i.e. ei>ai, (ou>) eū > aū (precise date unknown). For example, Old Welsh gueith, /gweiθ/ 'work' > Modern Welsh gwaith, /gwaɪθ/; O.W. esceir, /eskeir/ 'leg, ridge' > Modern W. esgair

/esgair/ on the one hand, and O.W. \*gueithiou /gweiθjɔu/, 'works' > Mod. W. gweithiau /gweiθjaɪ/; O.W. douceint /dɔukeint/, 'forty' > Mod. W. deugain /deɪgarn/ on the other hand.

- c) In non-final syllables, certain vowel reductions took place in the mid to late sixth century. The short vowels changing were: /ɪ, ɛ/ > /ə/, and /u/ > /ʊ/; e.g. (British iníssi, 'island' > ) O.W. inís > Mid. W. and Mod. W. ánis.

#### 4.3.1.2. Sound changes after the accent shift

It will be noted that sound changes occurring in ultimas and stressed monosyllables are of a 'strengthening' kind, while those occurring in non-final syllables and stressless monosyllables are reductions. The situation is reversed later in the history of the language; a few examples are given below.

- a) In final syllables, but not in stressed monosyllables, /aɔ/ monophthongised to /o/; e.g. British \*Caratācos (a proper noun) > O.W. Caratáuc > Middle Welsh Carádawc > Modern Welsh Carádog.
- b) Before final syllables, and before unstressed monosyllables, /h/ was lost; e.g. Mid. W. brenhin, 'king' > Mod. W. brenin, but Middle W. brenhines, 'queen' > Mod. W. brenhines.
- c) In final syllables and unstressed monosyllables, there is confusion between -n and -nn, also between -r and -rr, in texts from the eleventh century. In the same texts, there is no such confusion in non-final syllables and stressed monosyllables. The conclusion is that the new penultimate accent caused the reductions -nn > -n, -rr > -r, in newly unstressed syllables. Similarly, -nt becomes -n in these environments.

One point of interest about the two above sets of sound changes is that the sound changes before the accent shift concern vowel quality rather than length, while those after the accent shift mainly concern vowel length, and consonant and syllable quantity. This may reflect a change in the nature of the accent, which some scholars view as having altered not only its position but also its realisation.

#### 4.3.1.3. Nature of the old accent and its remnants today

A consideration of the phonetic nature of the Old Welsh accent may well shed some light on the situation in modern Welsh. The observation made at the end of section 4.3.1.2. gives a clue as to this phonetic nature. The old accent, or strategy for realising stress, seems to have made use principally of the cue of pitch, as well as unreduced vowel quality; these are seen in present-day ultimas and stressed monosyllables, and such cues are unlikely directly to influence vowel quantity. The Accent Shift, however, could have involved a shift in strategy away from pitch as the primary cue and towards the use of 'interval patterns' rather than 'length patterns' (see section 4.2.1.) and rhythm rather than inherent cues. Under such a strategy, unstressed syllables could still be pitch-prominent, since pitch was no longer a 'direct' stress cue but was solely concerned with intonation, while segmental length differences would become all-important and vowel shortening would ensue in unstressed syllables. This is the situation at the present day, involving penults, and to a certain extent, it seems, stressed monosyllables.

Some modern Welsh scholars make this distinction between the differing natures of the two types of accent. One early treatment of the old accent concludes that:

'C'est un accent de hauteur qui n'influe pas sur la quantité des voyelles sur lesquelles il se porte....Cet accent de hauteur devient intensif en gallois moyen et vers le XIII<sup>e</sup>- XIV<sup>e</sup>siècle réduit, surtout dans la prononciation, les diphtongues finales.' (Loth 1934)

D.M. Jones (1949) remarks on the high pitch frequently heard on ultimas, and gives examples where penults are lost in rapid speech, stating that: 'No doubt the frequently higher pitch of the final syllable plays a part in this process.' Whether it does so or not is still to be proven, but Jones, like Loth, links high pitch with the old accent, writing that the accent shift was not so much a shift of accent as a loss of the 'stress element' and a retention of the 'pitch element' in the case of final syllables. The penult, carrying only a 'stress element', is seen by Jones as having a stress, but not one of any great force, and he remarks on the notably small difference in stress between penults and ultimas. It is to be assumed that stressed monosyllables, in Jones' terms, retain both their 'pitch

element' and their 'stress element'.

A similar distinction is made by Watkins (1972a), who also remarks that present-day Welsh stress is very evenly distributed between syllables, 'particularly as between penult and final'. He notes the higher pitch of the unstressed ultima in Welsh and also on immediately post-stress unstressed syllables in the English spoken by some native Welsh speakers. Watkins states that, in Old Welsh, stress was evenly distributed on all syllables of polysyllables (which, he writes, would explain the retention of /h/ and of the opposition between single and geminate /n/ and /r/ in all syllables). The differences that developed between the vowels of final and pre-final syllables, according to Watkins, are attributable to the higher pitch of the ultima - this, again, needs to be proved. He writes that the accent shift was not so much a shift as a gradually-developing stress on the penult, while higher pitch remained on the ultima. Watkins links this new penultimate stress with the opposition of vowel length, writing:

'Yn gysylltiedig â phwysedd felly y mae'r cyferbyniad "hyd" a geir mewn Cymraeg llafar (fe'i ceir yn unig mewn unsillafion a sillafau gobennol).' (Watkins 1976:404) ['The opposition of "length" which is found in spoken Welsh is therefore attributable to stress (being found only in monosyllables and penultimate syllables).']

On the other hand, he finds developments in the modern dialect of Cwm Tawe, subsequent to conventionally-accepted developments, which seem to reflect the old accent:

'Mae'r ffaith fod cyferbyniad tôn rhwng sillafau terfynol ac unsill ar y naill law a sillafau cynderfynol ar y llall yn un rheswm pam y mae'r proses ffono-morffig "gwyriad" ("vowel mutation") wedi parhau yn ffenomenon gynhyrchiol yn y system lafar Gymraeg.' (Watkins 1976:404) ['The fact that there is an opposition of pitch between final syllables and monosyllables on the one hand and pre-final syllables on the other is a reason why the morphophonemic process of vowel mutation has remained a productive phenomenon in the sound system of Welsh.']

He later states that this process is beginning to be no longer productive in loan-words from English, and attributes this to the linguistic influence of English. Such an accusation may be unfounded,

since there is no reason not to suppose that these developments are the natural dying-out of the last remnants of the old accent that have remained in present-day Welsh.

A similar view of the phonetic nature of the old accent is taken by A. Thomas, who writes that

'...for Welsh, we may have to recognise two types of realization of accent:

rhythmic stress on the penultimate syllable,

pitch-prominence on the ultimate syllable.

And the ultimate syllable is often (perceptually) the stronger of the two, and always structurally the stronger.' (A. Thomas 1979:170)

While it may not be quite correct to accord both penult and ultima of the same word the status of 'stressed' (since stress is essentially culminative in function - see section 4.4.5.2), this observation nevertheless tallies with those quoted earlier. The consensus view seems to be that the old accent was signalled by pitch, the stressed ultima and stressed monosyllables having a higher pitch, and that it caused vowel mutation (reduction in vowel quality) in unstressed syllables; while the new accent is signalled by rhythm and by vowel and consonant length, causing reduction in vowel quantity in unstressed syllables. At the same time, however, a remnant of the old accent is seen in the phenomenon of vowel mutation (which, however, may no longer be productive) and in the frequently higher pitch of the ultima. Given such a situation, one may be forgiven for thinking that penultimate stress is not very firmly established in Welsh, and so other evidence - of a non-linguistic type - should also be considered.

#### **4.3.2 Penultimate stress; non-linguistic evidence**

##### **4.3.2.1. Traditional opinion**

The general consensus of traditional grammar-books has been that stress in Welsh is on the penult (e.g. Morris-Jones 1913, Rhys Jones 1977). This is supported by the judgements of native speakers in the experiments described in chapter 2 of the present study. This evidence is quite apart from the linguistic and methodological evidence presented in section 4.1., and for many would be the most convincing argument of all. Whatever may

be the phonetic prominence of the ultima, it seems to be outweighed by another factor determining stress.

#### 4.3.2.2. Music

In Welsh songs, the penult is the syllable that falls on the beat of the music. Similarly, in English songs, it is the stressed syllable that falls on the beat, in the better kinds of choral underlay. Thus it seems that the penult is indeed the stressed syllable in Welsh, and this conclusion also rests on evidence of a rhythmic nature.

#### 4.3.3 Final stress; non-linguistic evidence

##### 4.3.3.1. Poetic measure

On the other hand, some aspects of Welsh point towards a retention of some features of the old accent. For example, in Welsh poetry, the strict metres require a set number of syllables per line. This is similar to the French and Latin system of versification, rather than to the English and German system where merely a set number of feet per line is required, and the number of syllables may vary widely. Indeed, in earlier Welsh poetry, there is no concept of the 'foot' at all, since it does not enter into the extremely well-defined bardic rules for formal versification. It is only the more modern forms of verse (since the sixteenth century) that allow accentual composition in the sense of, for example, four or five accents or feet per line. Welsh poetry is extremely conservative, however, and this pattern of development may reflect a belated recognition of the nature of the new accent. The existing rules for 'syllable-timed' lines may be on the same level as vowel mutation; i.e. a remnant of an earlier stage of development which is no longer productive.

##### 4.3.3.2. Bardic rhyme rules

In Welsh rhyming poetry, final syllables and stressed monosyllables carry the rhyme, despite the final syllables' supposed lack of stress. While penults may rhyme (though only with other penults), in such cases the respective final syllables also rhyme. Final syllables may rhyme with stressed monosyllables, but in more modern verse it seems that the vowel of the monosyllable is most often short and thus identical in quality to

that of the ultima. In earlier poetry, however, the most common pattern was a rhyme between an ultima and a stressed monosyllable, where vowel length was immaterial: e.g. gwenn /gwɛn:/, 'white', in the compound meinwenn, /mainwɛn:/, 'maiden', could rhyme with gwên, /gwem/, 'smile' (Loth 1900). The common pattern nowadays is either a final syllable rhyming with another final syllable, or (most commonly) a stressed monosyllable rhyming with another stressed monosyllable. Very rarely, a penult plus ultima may rhyme with ultima plus stressless enclitic.

This very abbreviated summary nevertheless makes clear the pre-eminence of the ultima over the penult in matters of rhyme. That such a situation can arise when the stress is supposedly on the penult is presumably due to the fact that the ultima is the more sonorous syllable, being longer and lacking schwa entirely.

#### 4.3.3.3. "Cynganedd"

A device much used in formal Welsh poetry is cynganedd (approx. 'consonant harmony', pronounced /kəŋanɛð/). This can be very complicated, and has been elaborated into several narrowly-defined types; in essence, though, cynganedd concerns consonants and their judicious ordering to produce a pleasing effect of alliteration. For example, the consonants in one half of a line may be repeated in the same order in the second half, irrespective of word or morpheme boundaries. Such a device seems to disregard entirely the special status of the consonant after the stressed penult, treating this consonant in the same way as any other consonant. In this respect, at least, the 'new' accent seems immaterial in Welsh poetry: an accent that depends so crucially on the length of a particular type of consonant is not likely thus to ignore it.

#### 4.3.3.4. "Gair cyrch"

A poetic device concerning rhyme is the gair cyrch (approx. 'seeking word', pronounced /gair kɪrx/). This is a word at the end of the first line of a stanza, immediately after the rhyme-word. The gair cyrch rhymes not with the end of a line but (very often) with a word in the middle of the next line; this is a case where a stressed monosyllable must have an ultima as its rhyming word. Again, penultimate stress does not enter into

it. Even more intriguingly, in the absence of linguistic stress on the ultima, a special poetic device continues in use that makes much of the syllables on which the old accent fell, while completely ignoring the syllables on which only the new accent falls.

#### 4.3.3.5. "Accentuated cynghanedd"

A blend of the old and the new appeared in the work of a seventeenth-century Welsh poet, Huw Morus, who composed verses to be sung to tunes with a regular beat. As Parry writes,

'In the traditional metres it was the type of cynghanedd used in each line that determined the rhythm of the line, and the charm of this kind of poetry is due, in no small measure, to the almost infinite variety of rhythms that cynghanedd makes possible. In the Huw Morus type of verse the rhythm of the line is, of course, determined by the tune, and the result is accentuated cynghanedd, which, as an exercise in versification, is perhaps more difficult than the original system.' (Parry 1962:xiv-xv)

The tunes used by Morus were often English, and thus had been composed for accentual verse rather than for the traditionally 'syllable-timed' Welsh verse. Parry notes two metrical developments resulting from this practice:

'Firstly, the musical phrase was always very strictly observed and the words made to conform to the framework of the tune. Hence the intricate pattern of the stanzas. Secondly, cynghanedd was abundantly introduced, though it belonged traditionally to the strict metres and was, by its very accentuation, alien to the free metres. Full use was also made of rhyme to denote the end of musical phrases and even to mark a break in the phrase.' (Parry, op. cit.:552)

Unfortunately, Morus' successors were less skilled craftsmen in their fitting of words to a tune, and the genre of 'accentuated cynghanedd' never achieved the recognition of the traditional metres. It may perhaps be regarded as a transitional form between the purely 'syllable-timed' metre of the traditional bardic forms and the purely 'stress-timed' free metre that is becoming increasingly common nowadays. The latter type is seen both with and without cynghanedd or rhyme, and often approaches the timbre of English verse in its rhythm and its flexibility (which unkind critics would term looseness) of structure. It may be that the traditional

metrical forms are only artificially kept alive, preserving as they do in a petrified form certain features of the old accentual system, and that the newer metres are a reflection of the new accent which is destined to increase in influence even in this most conservative of fields in what is a most conservative language.

#### 4.4 Implications for the theory of stress

##### **4.4.1 Two trends**

###### **4.4.1.1. Statistical basis**

In the experimental work described in chapter 2, two trends were discovered in relation to stress in Welsh. These may be termed trends towards 'syllable-timing' and 'stress-timing'. The intriguing finding is that it is the weaker trend statistically speaking (towards 'stress-timing') which is the determiner of stress in the modern language. In such a situation, it is necessary to consider not the absolute lengths of units, but the direction in which changes in length are made; the model must be a kinetic rather than a static one. As Barnwell writes (see page 19 of this study), what is heard as indisputably rhythmic 'may be really the interpretation of changes in duration in the direction of true rhythm' (Barnwell 1971:88). Similarly, Classe sees isochrony as an undercurrent phenomenon rather than one manifested on the surface, writing

'...frequently, it only remains as an underlying tendency of which some other factor at times almost completely obliterates the effects' (Classe 1939:90).

###### **4.4.1.2. The situation in English**

In English, these two trends may also be discerned, though their relationship may well be different from their relationship for Welsh. The syllable defined as stressed according to rhythmic factors, in English, is also the only syllable eligible for pitch-prominence: herein lies the crucial difference between English and Welsh. The lack of mismatch in the syllables picked out by these two factors makes for a less confusing

situation, certainly, but it also carries with it the danger that stress and pitch could be viewed as the same phenomenon, when, in fact, they must be defined separately (see in particular section 1.2.4.1.).

This lack of mismatch, therefore, could be one reason why the independence of stress from pitch has sometimes not been recognised; since most work is carried out on English, it is natural for the results thus obtained to be assumed to apply also to other languages with seemingly similar features. It is here that Welsh comes into its own, pointing out the separateness of stress and pitch by being the exception that spotlights assumptions.

#### 4.4.1.3. The situation in Welsh

In modern Welsh, stress is based on rhythm, the pulse being marked out by segmental durational features. However, there remain certain features of the older accent, which seems to have been based on pitch alone. While these latter features seem to be no longer productive, their presence in the modern language makes the stress situation somewhat complicated and can also interfere with intonational patterns. While the old accent preserved the quality of stressed vowels, the new accent frequently alters vowel quality (simultaneously with vowel length) and can even fall on the completely reduced vowel, schwa - clearly vowel quality is not an important part of the new accent system. Instead, what remnants there are of the old accent are allowed free rein with vowel mutation and retention of the ultima when the penult is dropped (when weakly stressed); it is likely that such latitude is possible precisely because vowel quality has no bearing on the new accent, and thus cannot 'threaten' it directly.

On the other hand, vowel and consonant length are crucial to the realisation of the new accent, and it is clear that the penult is master here; the vowel length rules, and consonant lengthening and shortening/cluster simplification, point to the prominence of the penult at the expense of the ultima. Thus, while the prominence of the ultima may condition the quality changes that constitute vowel mutation, these changes never affect the length of the mutated vowels; this is reserved for changes conditioned by the penult alone.

The present state of the language offers a rich field for research into stress, particularly as the situation seems fluid. The vowel mutation process is no longer productive, according to Watkins (1976:405), while the penultimate stress is not strong enough to prevent some English words being loaned with their original stressing unchanged (Watkins 1972a). It will be interesting to observe which stress system eventually wins the day in Welsh.

#### **4.4.2 Stress accent and pitch accent**

##### **4.4.2.1. Conclusions for Welsh**

The general conclusions obtained from the experimental work described in chapter 2 are two in number, as has been outlined in section 2.4.2; they will merely be summarised here.

The first conclusion is that stress must be defined independently of intonation, or indeed of any pitch phenomena at all. This is an obvious feature of Welsh: the pitch-prominence of the unstressed syllable was noted by several earlier scholars, who nevertheless felt no need to relocate 'stress' on the syllable thus defined as prominent. In effect, they were defining stress on separate grounds, as have traditional grammar-books and the present study. The perceptual experiments in section 2.3. present evidence against defining stress purely in terms of pitch-prominence, while the theoretical apparatus developed in chapter 3 presents a descriptive framework within which stress and intonation may remain essentially separate while interacting in certain closely-defined ways.

The second conclusion is that the independent basis of the definition of stress in Welsh is rhythm. This conclusion is not as immediately obvious as the first, but is the only one possible given the evidence presented in chapter 2. The second conclusion augments the first, in that pitch and duration are essentially different acoustic phenomena which have no great influence on each other in the natural course of events; thus the separation of stress and intonation on the linguistic level is seen to rest on a similar clear separation on the phonetic level, and the explanation is thereby considerably simplified.

#### 4.4.2.2. Conclusions for English

While this study has been conducted in terms of Welsh, it is naturally hoped that some at least of its conclusions will prove to be applicable to other languages also, particularly English. In the case of the first conclusion (that stress is to be defined separately from intonation) there is already a precedent for English, in the shape of the traditional 'British' and 'American' schools of thought on suprasegmentals. These both view stress as separate from intonation, though they place their emphasis differently, and in this they are distinguished from the 'accent' analyses, such as that of Bolinger, which seeks to define stress in terms of pitch movement. The new 'metrical' theory of stress likewise separates stress and intonation, and finds that the theoretical description of each is simplified by so doing. Thus the first conclusion, formulated for Welsh, can equally well be applied to English.

The second conclusion (that stress is to be defined in terms of rhythm) is probably applicable to English also. This, however, rests at the moment on assertions of 'subjective isochrony' for English which are in need of somewhat more rigorous experimentation. The metrical theory finds it theoretically more elegant to view English stress in terms of rhythm. It may well be that the 'pulse' in English is marked by intensity rather than by the segmental durational features used in Welsh; however, this is a minor point not central to the question of whether the rhythm exists. For the present, it must be concluded that, while it seems likely that stress in English is based on rhythm, there is a need for more empirical research on this point.

#### 4.4.3 The notion of stress

##### 4.4.3.1. Form

What is termed 'stress' may take different forms in different languages, or in different historical stages of the same language. The form need not necessarily be inherent acoustic cues, as has been largely assumed up to now. A distinction is often drawn between 'stress accent' and 'pitch accent', the former viewed as prominence (of one syllable per word) based on intensity, and the latter viewed as this prominence based on pitch.

This distinction may well still be useful, and can help draw a distinction between stress in languages such as French (pitch accent) and languages such as English (stress accent). On the other hand, the terms are misleading in that 'stress accent' may cover more than one strategy for realising stress, as has been seen: it may refer to intensity or to rhythm, and also to both together or possibly to yet another factor.

#### 4.4.3.2. Function

The essential function of stress seems to be to denote 'prominence' in some sense related to semantic importance in the particular context. As Hyman (1977, 'On the nature of linguistic stress') writes, the culminative function of 'stress accent' was emphasised as a definitional requirement by European phonologists such as Trubetzkoy, Jakobson, Martinet, and Garde. The basic principles of a stress language, he states, are that only one syllable per word will receive primary stress, so that each stress signals a separate word (or other accentual unit) and, it is claimed, acts as an aid in processing utterances (Hyman, *op. cit.*:38). Any function that stress may have in facilitating the processing of utterances has yet to be investigated fully; for the moment it will suffice to define stress as a property of the word, assigned in the lexicon to a particular syllable of that word. This syllable then constitutes a potential locus for the phonetic manifestation of accent. Factors such as rate of speech, given and new information, and surface syntactic structure may determine whether or not the stressed syllable is actually accented. Hyman describes the relationship between stress and accent in the following terms:

'...stress is an underlying mental phenomenon, which however must somehow be picked up by the hearer. I therefore choose to regard the above perceptual cues as "strategies" for the realization of stress.' (Hyman, *op. cit.*:40)

It is quite possible, as has been seen, for 'strategies' to vary from language to language (e.g. English versus French), or even from an earlier state of a language to a later state (as happened in Welsh). While the form taken by stress in Welsh has changed, its function has not, in that it is still a culminative word-based phenomenon. It is a mistake to attempt to define stress solely in terms of its form, as has been the general rule to date; Welsh points out this fallacy clearly. Thus, in the definition

and description of stress, it is also necessary to take account of the functional and semantic aspects, and this is what has been attempted in the present study.

#### 4.4.4 Uncertainty in the description of stress

##### 4.4.4.1. What?

When examining stress in a given language, one cannot be certain what the realisation of stress will be. Synchronically, this realisation can differ between two languages, while diachronically it can differ between two historical stages of the same language. Thus it is a mistake to begin such an examination with assumptions concerning the likely nature of stress. This is seen clearly in the case of Welsh; if a researcher were to assume from the outset that stress would be realised by high pitch and pitch glides, as is the case in English, then the results would be completely in error. It is necessary to take into account other phenomena, such as sound changes, morphological processes, and vowel and consonant length rules, since stress is a feature of language that has repercussions on many levels of the linguistic analysis.

##### 4.4.4.2. When?

The present analysis has been deliberately flexible, attempting to incorporate a probabilistic element into the theoretical framework. This is in contrast to the 'always/never' approach of traditional descriptions, the rigidity of which could become a barrier to greater elegance of description. Often, one cannot with certainty predict when a particular rule will apply; for example, whether or not a 2-column in the metrical grid will count as stressed. It is true that the probabilistic element in this theoretical framework is but crudely sketched in as yet; however, the principle is there, and it is hoped that development along these lines will lead to a closer approximation to the truth of the situation.

##### 4.4.4.3. Why?

It seems that at some point in its historical development, the Welsh language changed its strategy of stress realisation. We do not know why it did so; moreover, we cannot even be sure what kind of factor to look

for as a cause of such change. One might attempt to 'explain' the change in terms of the influence of English; but this is a mere relocation of the problem, and would be implausible if the 'explanation' were formulated in terms of a cognitive shift in the mode of perception from 'length patterns' to 'interval patterns', since the latter clearly do not occur in English. If the 'cognitive shift' theory were espoused, one would still be hard put to it to find a catalyst for such a shift, when nothing comparable seems to have occurred in neighbouring languages of the same family.

It is perhaps a mistake to expect a very high degree of regularity in linguistic 'explanation'; languages are the creation of people, and people are notoriously irregular. In fact, a too-rigid formulation may lead the investigation away from language itself and nearer to the more abstract realms of speculative logic - which is an admirable pursuit, certainly, but not one that can be the primary concern of linguists, whose professed task is to search for and explain, if possible, the regularities and irregularities that occur in language as it actually is and not as some may consider it ought to be.

## REFERENCES

- ABERCROMBIE, L. 1923 Principles of English prosody. London: Secker.
- ADAMS, D. 1980 The Restaurant at the End of the Universe. London: Pan Books Ltd.
- ALLEN, G.D. 1975 Speech rhythm: its relation to performance universals and articulatory timing. Journal of Phonetics, 3: 75-86.
- ALLEN, M.R. 1975 Vowel Mutation and Word Stress in Welsh. Linguistic Inquiry, 6: 181-201.
- ALLEN, W.S. 1973 Accent and rhythm. Cambridge: Cambridge University Press.
- BARNWELL, T.P. 1971 An algorithm for segment durations in reading machine context. Technical Report 479. Cambridge, Mass.: M.I.T. Research Laboratory of Electronics.
- BELL, A. 1977 Accent Placement and Perception of Prominence in Rhythmic Structures. In Hyman (ed.) 1977.
- BLOOMFIELD, L. 1933 Language. New York: Holt; London: Allen and Unwin, 1935.
- BOLINGER, D.L. 1958 A Theory of Pitch Accent in English. Word, 14: 109-149.
- BOLINGER, D.L. 1965 Pitch Accent and Sentence Rhythm. In I. Abe & T. Kanekujo (eds.): Forms of English: Accent, Morpheme, Order. Cambridge, Mass.: Harvard University Press.
- BOLINGER, D.L. 1972 Accent is predictable (if you're a mind-reader). Language, 48: 633-644.
- BUXTON, H. 1983 Temporal Predictability in the Perception of English Speech. In A. Cutler & D.R. Ladd (eds.) Prosody: Models and Measurements. Berlin: Springer-Verlag.
- CATFORD, J.C. 1968 Articulatory possibilities of man. In Malmberg (ed.) 1968.

CHOMSKY, N. 1970 Deep structure, surface structure and semantic interpretation. In D.D. Steinberg & L.A. Jakobovits (eds.) 1971 Semantics. Cambridge: Cambridge University Press.

CHOMSKY, N. & HALLE, M. 1968 The Sound Pattern of English. New York: Harper & Row.

CLASSE, A. 1939 The Rhythm of English Prose. Oxford: Basil Blackwell.

CRYSTAL, D. 1969 Prosodic systems and intonation in English. Cambridge: Cambridge University Press.

EVANS, D.S. 1964 A Grammar of Middle Welsh. Oxford: Dublin Institute for Advanced Studies.

FRY, D.B. 1958 Experiments in the perception of stress. Language and Speech, 1: 126-152.

FRY, D.B. 1968 Prosodic phenomena. In Malmberg (ed.) 1968.

GIMSON, A.C. 1956 The Linguistic Relevance of Stress in English. Zeitschrift für Phonetik und allgemeine Sprachwissenschaft, 9: 143-149.

GIMSON, A.C. 1962 An Introduction to the Pronunciation of English. London: Arnold.

GRIFFEN, T.D. 1979 On Phonological Stress in Welsh. Bulletin of the Board of Celtic Studies, XXVIII: 206-212.

HYMAN, L.M. (ed.) 1977 Studies in Stress and Accent. Southern California Occasional Papers in Linguistics, 4. Dept. of Linguistics, University of Southern California. Los Angeles, California.

HYMAN, L.M. 1977 On the nature of linguistic stress. In Hyman (ed.) 1977.

JACKSON, K.H. 1953 Language and History in Early Britain. Edinburgh.

JONES, D. 1950, 3rd edn. The Pronunciation of English. Cambridge: Cambridge University Press.

JONES, D.M. 1949 The Accent in Modern Welsh. Bulletin of the Board of Celtic Studies, XIII: 63-64.

JONES, M. & THOMAS, A.R. 1977 The Welsh language: studies in its syntax and semantics. University of Wales Press for the Schools Council.

JONES, R.O. 1967 A structural phonological analysis and comparison of three Welsh dialects. University College of North Wales, Bangor: unpublished M.A. dissertation.

KINGDON, R. 1958 The Groundwork of English Stress. London: Longmans.

LADD, D.R. 1980 The Structure of Intonational Meaning. Bloomington and London: Indiana University Press.

LADEFOGED, P. 1967 Three Areas of Experimental Phonetics. London: Oxford University Press.

LADEFOGED, P. 1975 A Course in Phonetics. New York: Harcourt Brace Jovanovich.

LEHISTE, I. 1970 Suprasegmentals. Cambridge, Mass.: M.I.T. Press.

LEHISTE, I. 1975a The syllable nucleus as a unit of timing. Proceedings of the Eleventh International Congress of Linguists. Bologna - Florence, 1972. L. Heilman, (ed.) Pp. 929-932. Societa editrice il Mulino Bologna.

LEHISTE, I. 1975b The perception of duration within sequences of four intervals. Paper presented at the Eighth International Congress of Phonetic Sciences, Leeds, August 21.

LEHISTE, I. 1977 Isochrony reconsidered. Journal of Phonetics, 5: 253-263.

LIBERMAN, A.M., COOPER, F.S., HARRIS, K.S., MACNEILAGE, P.F., & STUDDERT-KENNEDY, M. 1963 A motor theory of speech perception. Proceedings of the Speech Communication Seminar, Stockholm 1962. Stockholm: RIT.

LIBERMAN, M.Y. 1975 The intonational system of English. Unpublished doctoral dissertation, Cambridge, Mass. Reproduced by Indiana University Linguistics Club, June 1978.

LIBERMAN, M.Y. & PRINCE, A. 1977 On Stress and Linguistic Rhythm. Linguistic Inquiry, 8: 249-336.

LIEBERMAN, P. 1960 Some acoustic correlates of word stress in American English. Journal of the Acoustical Society of America, 32: 451-454.

LIEBERMAN, P. 1965 On the Acoustic Basis of the Perception of Intonation by Linguists. Word, 21: 40-54.

LIEBERMAN, P. 1967 Intonation, Perception, and Language. Cambridge, Mass.: M.I.T. Press. Research Monograph No. 38.

LOTH, J. 1900 La Métrique Galloise: Introduction au Livre Noir de Carmarthen et aux vieux poèmes gallois. Paris: Ancienne Librairie Thorin et fils. 2 vols. Appears as H. d'Arbois de Jubainville, Cours de Littérature Celtique, no. 9.

LOTH, J. 1934 Remarques à l' "Historia Britonum" dite de Nennius (suite). Revue Celtique, 51: 1-31.

MALMBERG, B. 1963 Phonetics. New York: Dover.

MALMBERG, B. (ed.) 1968 Manual of Phonetics. Amsterdam: North-Holland Publishing Company.

MORRIS-JONES, J. 1913 A Welsh Grammar. Oxford.

O'CONNOR, J.D. & ARNOLD, G.F. 1961, 2nd edn. 1973. Intonation of Colloquial English: A practical handbook. London: Longmans.

O'CONNOR, J.D. & TRIM, J.L.M. 1953 Vowel, consonant and syllable - a phonological definition. Word, 9: 103-122.

OHALA, J.J. 1977 The Physiology of Stress. In Hyman (ed.) 1977.

ÖHMAN, S. & LINDQVIST, J. 1966 Analysis-by-synthesis of prosodic pitch contours. Speech Transmission Laboratory Quarterly Progress and Status Report, 4/1965: 1-6. Stockholm, Sweden.

PALMER, H.E. & BLANDFORD, W.G. 1924 A Grammar of Spoken English on a Strictly Phonetic Basis. Cambridge: W. Heffer & Sons.

PARRY, T. (ed.) 1962 The Oxford Book of Welsh Verse. Oxford: Clarendon Press.

PIKE, K.L. 1945 The Intonation of American English. Ann Arbor, Mich.: University of Michigan Press.

RAPP, K. 1971 A study of syllable timing. Speech Transmission Laboratory Quarterly Progress and Status Report, 1: 14-19. Stockholm, Sweden.

REES, J.M. 1977 Aspects of Welsh intonation. University of Edinburgh: unpublished doctoral dissertation.

RHYS JONES, T.J. 1977 Living Welsh. London: Hodder and Stoughton Ltd.

SELKIRK, E.O. 1980a The Role of Prosodic Categories in English Word Stress. Linguistic Inquiry, 11: 563-605.

SELKIRK, E.O. 1980b On Prosodic Structure and its Relation to Syntactic Structure. Unpublished paper presented to the Sloan Workshop on the Mental Representation of Phonology, University of Massachusetts, November 18-19, 1978. Reproduced by Indiana University Linguistics Club, 1980.

SHEN, Y. & PETERSON, G.G. 1962 Isochronism in English. University of Buffalo, Studies in Linguistics, Occasional Papers, 9: 1-36.

SOMMERSTEIN, A.H. 1977 Modern Phonology. London: Arnold.

STETSON, R.H. 1928 Motor Phonetics. Amsterdam: North-Holland Publishing Company.

SWEET, H. 1878 Handbook of phonetics. Oxford: Clarendon Press.

THOMAS, A. 1979 A lowering rule for vowels and its ramifications, in a dialect of North Welsh. In: Occasional papers in linguistics and language learning, no. 6: Papers in Celtic Phonology. The New University of Ulster.

THOMAS, C.H. 1967 Welsh Intonation - A Preliminary Study. Studia Celtica, 2: 8-28.

THOMPSON, H.S. 1980 Stress and Saliency in English: Theory and Practice. Unpublished doctoral dissertation: Palo Alto Research Center, California.

TRAGER, G.L. & SMITH, H.L. Jr. 1951 An outline of English structure. Studies in Linguistics, Occasional Papers, 3. Washington: American Council of Learned Societies; Norman, Okla.: Battenburg Press.

VANDERSLICE, R. 1967 Larynx versus lungs: cricothyrometer data refuting some recent claims concerning intonation and archetypality. Working Papers in Phonetics, University of California at Los Angeles, 7: 69-79.

VAN KATWIJK, A. 1974 Accentuation in Dutch. An experimental linguistic study. Utrecht. [Cited by Ohala 1977.]

WATKINS, T.A. 1953 The Accent in Cwm Tawe Welsh. Zeitschrift für Celtische Philologie, XXIV: 6-9.

WATKINS, T.A. 1972a The Accent in Old Welsh - its Quality and Development. Bulletin of the Board of Celtic Studies, XXV: 1-11.

WATKINS, T.A. 1972b The Accent-Shift in Old Welsh. In H. Pilch & J. Thurow (eds.) Indo-Celtica: Gedächtnisschrift für Alf Sommerfelt. Pp. 201-205. In series: Commentationes Societatis Linguisticae Europaeae. Munich: Max Hueber Verlag.

WATKINS, T.A. 1976 Cyfnewidiadau Seinegol sy'n Gysylltiedig â'r 'Acen' Gymraeg. Bulletin of the Board of Celtic Studies, XXVI: 399-405.

WILLIAMS, B.J. 1980 Some features of modern Welsh intonation. Unpublished MS.

WILLIAMS, B.J. In press. An approach to the Welsh vowel system. Bulletin of the Board of Celtic Studies, May 1984 edn.