Apples – Journal of Applied Language Studies
Vol. 2, No. 1 , 2002, 1-24

The functional irrhythmicality of spontaneous speech: A discourse view of speech rhythms

Richard Cauldwell, University of Birmingham

Experimental evidence has fully refuted the stress- and syllable-timing hypothesis (SSH) of speech rhythms. However, it remains the prevailing view and still features in accounts of the rhythms of speech because no other hypothesis matches its deceptively bewitching power. This paper, written from a discourse perspective, proposes a replacement for SSH: spontaneous speech is functionally irrhythmic. Although the formal events of speech – phones, strong and weak syllables, words, phrases – occur 'in time' (they can be plotted on a time line) they do not occur 'on time', (they do not occur at equal time intervals). English is not stress-timed, French is not syllable-timed. The rare patches of rhythmicality are either 'elected' – as in scanning readings of poetry and the uttering of proverbs – or 'coincidental' – the side-effects of higher order choices made by speakers. Coincidental rhythmicality is most likely to occur where there are equal numbers of syllables between stresses. In spontaneous speech, the speaker's attention is on planning and uttering selections of meaning in pursuit of their social-worldly purposes, and this results in an irrhythmic norm which aids comprehension.

Introduction

The evidence from research is overwhelmingly against the hypothesis that languages are either 'stress-timed' or 'syllable-timed'. It is not possible to divide languages into either 'syllable-timed' or 'stress-timed' categories; it is not the case that stresses occur at equal time-intervals in 'stress-timed' languages; it is not the case that syllables occur at equal time intervals in 'syllable-timed' languages; so-called 'syllable-timed' and 'stress-timed' languages are alike in having variations in syllable-length; so-called 'syllable-timed' and 'stress-timed' languages are alike in having variations in inter-stress-interval length.

Astonishingly, the stress and syllable-timing hypothesis has survived its refutation, because in the absence of an alternative hypothesis, scholars continue to use the hypothesis in their accounts of the rhythms of English (e.g. Cruttenden, 1997; Ball & Rahilly, 1999; Rogers, 2000).

ISSN: 1457-9863
Publisher: Centre for Applied Language Studies, University of Jyväskylä

http://apples.jyu.fi

The purpose of this paper is to present a hypothesis of the rhythms of spontaneous speech which is more in accord with experimental evidence. The paper adopts a discourse perspective, the details of which will be made clear as the paper proceeds, and in more depth in Section 4. Briefly, the position is that spontaneous speech is speaker-controlled, purpose-driven, interactive, co-operative, context-related, and context-changing (Brazil, 1995: 26-39).

In Section 2 I identify those features which have enabled the stress and syllable timing hypothesis to survive its refutation. In Section 3 I review key aspects of the research evidence for the refutation of SSH. In Section 4 I outline the discourse approach adopted in this paper. Section 5 is a technical section: in it I analyse a short extract of spontaneous speech and argue that the majority of units of speech (tone-units) are too short to be rhythmical. In Section 6 I present an alternative hypothesis to stress and syllable timing: speech is functionally irrhythmic featuring rare patches of coincidental rhythmicality, and rarer still patches of elected rhythmicality. In Section 7 I consider the argument that stress and syllable timing is a perceptual phenomenon, and propose that irrhythmicality aids comprehension.

The Stress and Syllable Timing Hypothesis

Abercrombie's (1967) statement of the stress- and syllable-timing hypothesis (henceforth SSH) has been the most authoritative and influential. It features two mutually exclusive categories of language differentiated by contrasting types of isochronic rhythm:

> As far as is known, every language in the world is spoken with one kind of rhythm or with the other. In…*syllable-timed* rhythm...the syllables recur at equal intervals of time-they are *isochronous*. In…*stress-timed* rhythm…the stressed syllables, are isochronous... there is considerable variation in syllable-length in a language spoken with a stress-timed rhythm, whereas in a language spoken with a syllable-timed rhythm the syllables tend to be equal in length ... (Abercrombie, 1967, p97-98). [Italics are Abercrombie's]

Thus formulated, SSH presents a very clear-cut picture of phenomena that Abercrombie's predecessors had found far less clear-cut ('As far as is known...' is somewhat disingenuous). Whereas Abercrombie sees SSH as differentiating languages, Pike (1945) viewed stress-timing and syllable-timing as being in phonemic contrast *within the same language*; whereas Abercrombie asserts that speech is regular, Jones (1960, writing in 1918) had noted the 'extreme difficulty of describing or reducing to rules the innumerable rhythms heard in ordinary connected speech' (242).

SSH applies to verse and spontaneous speech alike: Abercrombie notes the 'close connexion between ordinary speech rhythm and verse rhythm'; for him, '[t]he rhythm of everyday speech is the foundation of verse...' (1967: 98). But his predecessors had found it necessary to qualify their assertions in relation to speech styles. Pike speaks of stress-timing being controlled 'strictly and mechanically in poetry – and possibly partially so in some types of elegant prose...' (34); Classe (1939) distinguished between three speech-styles – verse, prose, and speech: and said of speech that the 'necessary conditions' for

isochronous groups 'will generally not be present at the same time' (1939: 85-86 cited in Couper-Kuhlen, 1993: 12).

Abercrombie's predecessors provided SSH with a defence against research evidence: the 'interference view'. For Jones (1918), the agents of rhythm (the formal properties of the language – speech sounds and grammatical relations) 'interfere' with the general tendency to stress-timing. This interference view of the evidence from connected speech is a common theme: for Classe, isochrony 'frequently ... only remains as an underlying tendency of which some other factor at times almost completely *obliterates* the effects' (1939: 90 emphasis added). A consequence of the interference view is that there is always a ready defence against counter-evidence.

Another reason for the durability of SSH is that it is (apparently) easy to demonstrate. For example Underhill (1994: 71), writing for teachers of English as a second or foreign language, suggests working through the following sequence of utterances, speaking the prominent syllables at the same rhythm despite the increase in the number of intervening syllables:

a.	YOU	ME	HIM	HER
b.	YOU and	ME and	HIM and	HER
c.	YOU and then	ME and then	HIM and then	HER
d.	YOU and then it's	ME and then it's	HIM and then it's	HER

Demonstrations of this kind abound in the literature, and I shall argue (in 6.2 below) that they are only superficially related to timing – they are in fact demonstrations of the plasticity of speech.

SSH has survived because it is deceptively clear-cut, easily defended, easily demonstrated, and applies to all speech styles and all languages.

Research evidence

Abercrombie's SSH is, in fact, not a single hypothesis, but a collection of hypotheses. They include:

(a) all languages fall into one of two mutually exclusive categories: stress-timed or syllable timed

(b) in stress-timed languages, stresses occur at equal time-intervals (stress-isochrony)

(c) in syllable-timed languages, syllables occur at equal time intervals (syllable-isochrony)

(d) syllable-length varies in stress-timed languages, but not in syllable-timed languages

(e) inter-stress-intervals vary in length in syllable-timed languages, but not in stress-timed languages

These hypotheses are interdependent: (b) and (c) contain the defining characteristics (stress-isochrony, and syllable-isochrony) of the two categories that make up the binary distinction in (a). Thus if the research evidence shows that either one of stress-isochrony (b) or syllable-isochrony (c) does not exist, then hypothesis (a) is refuted. Hypothesis (a) would also be refuted if it were found that no language is characterised entirely by stress-timing, or if it were found that no language is entirely syllable-timed.

These hypotheses seem at first sight to be eminently testable, but as Roach (1982) makes clear, the methodological problems of testing the hypotheses are difficult to surmount. The problems include: (i) consistent identification of stresses across languages by the researchers and informants (ii) deciding where the start and end points should be for measuring inter-stress intervals (iii) how to allow for variations in tempo (iv) how to deal with pre-head and post-tonic syllables (74-76).

In the following sections we will review the evidence from the work of three scholars Roach (1982), Dauer (1983), and Couper-Kuhlen (1993).

Interstress intervals, and syllable length

Both Roach (1982) and Dauer (1983) addressed the issue of assigning a language to one or other of the categories 'stress-timed' or 'syllable-timed'. Roach used samples of two minutes of unscripted speech from six speakers, one for each of the languages listed by Abercrombie: French, Telugu, and Yoruba ('syllable-timed' languages) and English, Russian and Arabic ('stress-timed' languages). Dauer compared recordings in two stress-timed languages (English and Thai) and a syllable-timed language (Spanish), and two unclassified languages (Italian and Greek) of 'a passage from a modern novel or a play, in which a character is speaking in normal, everyday language' (52). I shall focus on her findings for Spanish and English.

Both Roach and Dauer examined inter-stress interval length. Roach found that the 'stress-timed' group of languages (against expectations) had greater variability in the length of inter-stress intervals than the 'syllable-timed' grouThus it would seem that inter-stress-interval-length differentiates between the two groups of languages, but *in the reverse direction* of SSH hypotheses (b) and (e) listed above: in other words, the 'stress-timed' group had greater variability in inter-stress-intervals than the 'syllable-timed' grouHowever, Roach attributes these differences to extreme values for one individual, and states that 'the figures...are better taken just as grounds for rejecting the hypothes[e]s' rather than evidence for calling the stress-timed group syllable-timed (77).

Dauer found that while there were no significant differences between languages, there were significant differences between speakers with extremes of speaking rate, even within the same language grouHer slow speaker of Spanish had significantly different results from her faster speaker of Spanish; her slow speaker of Greek had significantly different results from her fast speaker of Greek.

Dauer also found that English and Spanish were alike in that the timing of inter-stress intervals is proportionate to the number of syllables in both languages. This was what SSH predicts for Spanish (a 'syllable-timed' language) but it is against predictions for English ('stress-timed').

Roach, in addition, compared syllable duration across the two groups of languages and found similarities rather than differences: although the 'stress-timed' group showed variability in syllable-length (in line with expectations), the same was found to be true (against expectations) of the 'syllable-timed' group.

The evidence for refutation of SSH which emerges from these studies is that all the languages investigated showed variability in syllable-length, and variability in inter-stress-interval length. In other words, because of the inter-dependence of the hypotheses, the evidence is against the existence of the categories 'stress-timed' and 'syllable-timed'.

Perception of isochronous chains

Whereas Roach and Dauer used instrumental means to measure inter-stress intervals in different languages, Couper-Kuhlen (1993) used hearers' perceptions to identify 'isochronous chains' in just one language, English. She thus addressed hypothesis (b), stress-isochrony. Two informants analysed a two-minute extract from a phone-in programme broadcast on Radio Manchester (UK) consisting of 23 turns of varying length between the host and a caller. They identified the isochronous chains through repeated listenings, searching for stretches of speech sufficiently rhythmic for them to be able to tap a pencil, or nod their head to. The informants identified 48 isochronous chains in the recording, but there were some stretches of speech which did not form part of isochronous chains: 36% of all syllables, and 17% of stressed syllables, occur outside the 48 isochronous chains.

Couper-Kuhlen concedes that English is not 100% stress-timed: 'English speech is not uniformly isochronous over extended periods of time' (48 her italics). However, she qualifies this statement: 'But just as significantly, the passage is not wholly unisochronous either. In fact, allowing for discontinuities, a large portion of it is isochronous in one way or another' (48 her italics).

For Couper-Kuhlen, English is not isochronous when viewed from the macro-perspective of the entire temporal extent of a spoken text, but from the micro-perspective of the internal characteristics of each of the 48 chains it is isochronous.

The discussion of research findings

There are two common features of the discussion of evidence in SSH research: first, the very categories whose existence has been refuted are nevertheless required to facilitate discussion of the findings; second, scholars prefer to accommodate the refutation evidence in a revised version of SSH rather than abandon SSH in its entirety.

Roach (adopting a position similar to Pike 1945) concluded that 'there is no language which is totally syllable-timed or totally stress-timed – all languages

display both sorts of timing; languages will, however, differ in which type of timing predominates' (1982: 78). The wording of this conclusion is such that the categories 'stress-timed' and 'syllable-timed' remain necessary for discussing the rhythms of languages. In addition, it would be more logical to conclude, as the sub-hypotheses are interdependent (cf. Section 3.1), that the evidence of the investigation is that the entire SSH hypothesis is refuted.

Dauer seems to reject SSH in its entirety concluding that 'the difference between English, a stress-timed language, and Spanish, a syllable-timed language has nothing to do with the durations of interstress intervals', concluding that 'what these data reflect appears to be universal properties of temporal organisation in language' (1983: 54). However, in the later sections of her paper, Dauer continues to refer to the rhythmic differences as 'stress-timed' and 'syllable-timed' even as she advocating abandoning these terms: 'Many foreigners ... learning English use a syllable-timed rhythm' (60). She demonstrates the need for the terms as she argues for their abandonment.

Dauer proposes an avoidance of the word 'timing' – favouring adoption of the term 'stress-based' (following Allen, 1975 and O'Connor, 1973). For Dauer, a stress-based language is one in which stress plays a large role in word-stress, syllable structure and vowel reduction. It is important to realise that for Dauer, the term 'stress-based' constitutes a rejection of the notion of timing. She also proposes viewing languages as being placed along a 'dimension' (59) of 'more or less stress-based rhythm' rather than belonging to one or other of the mutually exclusive categories 'stress-timed' and 'syllable-timed'. It is important to note that the term 'syllable-based' does not feature on this dimension. However, a number of scholars (Laver, 1994 528; Dalton & Seidlhofer, 1994 42) seem to credit her with being the originator of the 'stress-based/syllable based continuum'.

Thus, while presenting the counter-evidence, scholars find the categories of stress-timing and syllable-timing too tenacious, attractive, and convenient to abandon. For Laver (1994) the tenacity of the concept of stress-timing is an indication of an underlying truth (524); Crystal (1996) finds it convenient to use for lack of anything better – the distinction between stress and syllable-timed languages 'is an extremely crude one and in its bare form is almost certainly wrong' (8) but 'it will stay until a more-refined classification of rhythmical types arrives on the phonetic scene' (9). Dalton and Seidlhofer (1994: 110) find SSH attractive: while acknowledging the difficulties with stress-timing and syllable timing they state 'It cannot be denied ... that ...stress-time still represents an appealingly neat categorisation, so that references to stress-time (especially with regard to English) are still frequent'.

A discourse view

A discourse view of speech, which takes into account those factors which are immediately relevant to speakers communicating in real time, in context, offers an account of speech rhythms which is more in line with the research evidence. The principles of the discourse approach are outlined in 4.1; and in 4.2 there is an explanation of the tone-unit, a unit of speech central to the arguments of this paper.

Against interference

The view of spoken discourse adopted here is that spontaneous speech is speaker-controlled, purpose-driven, interactive, co-operative, context-referenced, and context-changing (Brazil, 1995, p26-39). The choices that speakers make, and their reactions to the making sense to their hearers in context in real-time, are central concerns.

The contention is that speakers are the agents of rhythm. The suprasegmental choices that speakers make (speed of delivery, size of tone-unit, pitch-height, tone-choice, volume), and performance factors inevitable in unscripted speech (pauses, restarts, etc) are the dominant factors in determining the rhythm of an utterance. Crucially, these factors (for the purpose of this paper they will be known collectively as 'discourse factors') are more influential than a syntactic sequence of word-accent patterns.

Consider the sentence 'My *cou*sin his *dau*ghter is *re*cently *wi*dowed'. This sentence has four polysyllabic words, each with the word-accent on the first syllable (italicised), and there are two unstressed syllables between each word-accent. If we view this sentence as an isolated language sample containing a sequence of citation forms, it would appear that we have favourable conditions for rhythmicality: a sequence of four word-accents separated by an equal number of syllables (recall that Dauer 1983 found that the length of inter-stress intervals is proportional to number of syllables it contains); and it is possible to read out this utterance in a way which illustrates rhythmicality. Spoken thus, the sentence seems to offer evidence that English is stress-timed. It seems a reasonable next step to view rhythm as inhering somewhere between the word-accent shapes of the lexicon, and the syntactic rules governing the linking of these forms into possible utterances in the language.

This view contrasts strongly with the discourse view which starts with used language – language that has happened. The sentence discussed above actually occurred thus:

9	▲ AND my COUsin ♦♦	3.9	2.9
10	►▼ HIS DAUGHter ♦♦	3.8	2.5
11	▲ ERM	3.3	3.3
12	►▼ is Recently WIdowed ♦♦	5.1	3.5

The numbers in the first column are reference numbers which refer to a longer text 'Moving Again' which will be treated in more detail below. The transcription follows (broadly) the conventions of Discourse Intonation (Brazil, 1997). Each line contains a separate tone-unit; the arrows signify the tone-choice which starts on the underlined syllable; upper case letters indicate prominent syllables; the double diamond a long pause. The last two columns indicate articulation and speaking rate in syllables per second.

It is immediately apparent that the 'sentence' we have been discussing belongs to a larger unit of some kind: tone-unit 9 begins with *and* indicating that it has been preceded by other material; also, there is no falling tone – indicating that this sequence of tone-units does not yet constitute an act of telling (Brazil, 1995) – there must therefore be more material – with falling tone – to follow.

Now that we are no longer imagining a reading of an isolated language sample, we have to note the discourse factors at work: there is a filled pause in tone-unit 11, and silent pauses after tone-units 09, 10, & 12; articulation rate and speaking rate vary between highs of 5.1/3.5, and lows of 3.3/2.5 syllables per second respectively; grammatical words are unexpectedly made prominent: *and* (09), *his* (10); there are level tones in 09 and 11, and fall-rise tones in 11 and 12.

The imagined reading of the isolated sentence led us to believe that conditions were favourable for rhythmicality. However, the discourse factors at work in the four tone-units result in a sequence which was not perceived as rhythmical (cf. Cauldwell, 2000).

The traditional view has been that these discourse factors are 'interfering' or 'obliterating' an underlying inherent isochronic rhythm. However, discourse factors are a necessary part of spontaneous speech and they are central to the issue of perception of rhythm. Discourse factors, far from interfering or obliterating isochrony, are the 'stuff' of rhythm. They most often result in irrhythmicality which (as I shall argue in Section 7 below) has an important function in spoken communication.

The interference view, implies that at some point in its preparation, every embryonic utterance exists complete in an isochronous state, which is then changed (interfered with, obliterated) by other factors as the utterance moves towards actualisation in performance. This view is tantamount to believing that every utterance has a psycholinguistic antecedent which has verse-like isochrony. The discourse view is that speech is produced linearly tone-unit by tone-unit, and with Laver (1970) holds that the tone-unit is the most likely candidate for neuro-linguistic preassembly. Filled pauses (*erm* and *er*), silent pauses, and level tones are symptomatic of the speakers' focus of attention being split between future, present, and past demands: speakers are deciding what to say next, while producing the current utterance, which they have just decided to say.

Prominences and tone-units

The majority of speech can be divided up into tone-units of three sizes: single-prominence, double-prominence, and triple prominence tone-units. Tone-unit boundaries are notoriously difficult to define (cf. Brown, Currie & Kenworthy, 1980; Knowles, 1991). The particular discourse approach adopted here – Discourse Intonation (Brazil, 1997) henceforth 'DI' – applies both internal and

external criteria. The important internal events occur on syllables that speakers choose to make *prominent*, and particular importance is associated with the *tonic prominence*. Having made a syllable prominent, a speaker has the option of making it *tonic;* once a tonic prominence has occurred, the tone-unit is complete (barring post-tonic non-prominences). Any subsequent prominences occur in a new tone-unit. Thus a tone-unit is complete when a tone has occurred – there can only be one tone per tone-unit.

The pause is an external criterion relevant to determining the extent of a tone-unit: wherever they occur, pauses are regarded as marking the end of tone-units, even if the result is an incomplete tone-unit. Thus, wherever there is a pause, there is a tone-unit boundary – but tone-unit boundaries can occur where there is no pause, and in these cases they occur (as described in the previous paragraph) *between* the tonic prominence and any subsequent prominence, though the precise location of the tone-unit boundary is not of importance.

Using the occurrence of tonic prominences and pauses as tone-unit boundary markers is useful for the investigation of rhythmicality in English, because the phenomena that occur on a tone (pitch change, increase in amplitude) are likely to co-occur with added lengthening of the tonic syllable and any subsequent syllables. Such boundary phenomena are likely to bring to an end any rhythmic pattern set up by a preceding sequence of prominent syllables. Indeed Roach (1982) chose to discount post-tonic elements (and pre-heads) from his investigation precisely because they created measurement problems (76).

Making sense of the counter-evidence

From this point onwards, this paper will focus on the presentation of an alternative hypothesis of the rhythms of speech. I shall confine my examples to English, but I would expect a substantial proportion of the argument to hold for spontaneous speech in any language. This section presents an analysis of a twenty-second extract of spontaneous speech. It is the most technical section of the paper, preparing the way for the statement of the irrhythmicality hypothesis in section 6.

First, a note on terminology. It is possible to discuss rhythms of speech in both isochronous and non-isochronous terms. The isochronous view looks for the co-occurrence of speech events with regular time-intervals: SSH takes an isochronous view of speech. It is possible however to discuss the rhythms of English in a non-isochronous way, where rhythm is seen 'as a pattern of events related to one another in terms of salience' (Couper-Kuhlen, 1986: 51). One can speak of rhythmic patterns of alternation of weak and strong stresses or of rules such as the intermediate accent rule (Knowles, 1987: 125) without necessarily holding that isochrony is a factor.

From this point on I shall use the terms *rhythm* and *rhythms* to refer generally to patterns of language events in speech (of whatever kind): this phrase will be neutral in relation to timing. I shall use the terms *rhythmical* and *rhythmicality* to refer to cases of perceived isochronic patterns.

Spontaneous speech: 'Moving Again'

It is now necessary to look at a text 'Moving Again' of which we have seen an extract above. 'Moving Again' consists of 21 seconds of speech, and is taken from a larger extract 'Houses in New Zealand' lasting 1 minute 52 seconds (Cauldwell, 1997). The speaker – Gail – is answering a question about an uncle living in New Zealand, who has a hobby of buying houses, doing them up and selling them. The shaded rows (7 & 8, and 19) indicate tone-units identified by two informants (cf. Cauldwell, 2000) as being rhythmical: they will be discussed below.

The transcription follows (broadly) the conventions of Discourse Intonation, (Brazil, 1997) – henceforth *DI*. Each line contains a separate tone-unit; upper case letters indicate prominent syllables; the underlined syllables show the syllable upon which the tonic movement starts; the tone itself indicated by the arrows which precede the tone-unit; the single diamond denotes a short pause, the double diamond a long pause. Column C gives the articulation rate (this excludes pauses) of each tone-unit in syllables per second; Column D gives the speaking rate (this includes pauses) in syllables per second.

Table 1. 'Moving Again'.

a)

A	*B*	*C*	*D*
1	▲ he DOESn't ♦♦	4.1	3.1
2	▲DO it ERM ♦♦	3.6	2.2
3	►for the MOney he's going to MAKE	8.9	8.9
4	▲but he DOESn't LOSE	6.3	6.3
5	► when he SELLS his house ♦♦	4.5	2.7
6	▲ ER ♦♦	2.0	0.9
7	▼he's CURrently THINking of MOving aGAIN	6.9	6.9
8	▼he's EIGHty TWO ♦♦	4.6	2.3
9	▲AND my COUsin ♦♦	3.9	2.9
10	►▼ HIS DAUGHter ♦♦	3.8	2.5
11	▲ ERM	3.3	3.3
12	►▼ is REcently WIdowed ♦♦	5.1	3.5
13	▲AND ♦	2.4	2.0
14	▲ has JUST	5.1	5.1
15	►MOVED in to LIVE with them	5.9	5.9
16	►having SOLD her house ♦	4.0	2.6
17	▲ ERM	3.7	3.7
18	►▼ he's HAving a house BUILT for her	6.0	6.0
19	▼which she'll MOVE into in FIVE month's TIME ♦	4.3	3.8
20	▼and when SHE moves OUT	4.8	4.8
21	▲TO	2.2	2.2
22	► live in her NEW house ♦	4.3	3.6
23	▼they will MOVE HOUSE ♦	4.0	3.5
24	► to be NEAR her	5.1	5.1

Thus, reading across row 5, the downward arrow tells us that this tone-unit features a falling tone, that the syllables *when he* and *his house* are non-prominent, that *sells* is both prominent (upper case) and tonic (underlined), the falling tone starts on this syllable, and continues over the last two syllables of the tone-unit. The two diamonds indicate that there is a long pause between this speech unit and the next; the last two columns tell us that the articulation rate was 4.5 syllables per second, and the speaking rate was 2.7 syllables per second.

The 24 tone-units of MA exhibit features which are common in spontaneous speech: 14 (i.e. a majority of the twenty-four tone-units) are not co-terminus with a clause; clauses are split between tone-units (e.g. 'he doesn't // do it erm // for the money he's going to make'); they are characterised by parataxis (e.g. 'and my cousin his daughter is recently widowed and has just moved in to live with them') rather than hypotaxis; there are main clauses ending with rising tones (07 & 08) and subordinate clauses ending with a falling tone (20-22).

Thirteen of the tone-units are followed by pauses, four tone-units feature filled pauses with level tone (02, 06, 11, & 17), and six other tone-units have level tone in which the speaker rests momentarily on words while deciding what to say next (01, 04, 09, 13, 14, & 21). The other tones to occur are the falling tone (03, 05, 15, 16, 22, 24) which DI associates with telling and the two types of rising tones, rise (07, 08, 19, 20, & 23) and fall-rise (10, 12 & 18) which DI associates with referring (cf. Brazil, 1997).

Rhythms of speech in 'Moving Again'

The minimum requirement for speech to be perceived as rhythmical is that there should be two events of some kind which match, or are perceived to match, in some way. Scholars have typically investigated two different types of potentially rhythmical event: interstress intervals (Roach, 1982; Dauer 1983); and metrical feet (Abercrombie, 1964; Halliday, 1967, 1994). What I want to demonstrate in this section is that most tone-units are too short to contain a sufficient number of matching events to be perceived as rhythmical.

Tempo variations and pauses.

It was clear from Table 1, (and Roach, 1982 and Couper-Kuhlen, 1993 noted the same) that the tempo of speech changes constantly: the articulation rate (Column C) featured a high of 8.9 syllables per second (tone-unit 03) and a low of 2.0 syllables per second (tone-unit 06). The stream of speech thus features constant fluctuation in articulation rate around (in the case of MA) an average of 4.5 syllables per second. An additional factor in the variation of tempo is the occurrence of pauses: thirteen of the twenty-four tone-units are followed by short or long pauses. This results in a speaking rate (Column D) which fluctuates between 8.9 and 0.9 syllables per second, around a mean of 3.9. Because of the fluctuations in tempo and the occurrence of pauses it is less likely that rhythmicality will be perceived <u>across</u> tone-units than <u>within</u> tone-units.

Single-prominence tone-units

The majority of tone-units in *MA* are either single or double-prominence tone-units (there are eleven of each): only two tone-units (07 with four, and 19 with three) have more than two prominences. Single-prominence tone-units are too short to provide a sufficient number of matching events. Table 2 shows the properties of a single-prominence tone-unit, with an example from MA. A single-prominence tone-unit has three elements to its phonological structure: a proclitic element, a tonic element, and an enclitic element. The shaded column indicates the prominent syllable, the unshaded columns (2 & 3) represent non-prominent syllables.

Table 2. Properties of a single prominence tone-unit.

	1	2	3
element	*proclitic*	*tonic*	*enclitic*
words/syllables	when he	SELLS	his house
duration (ms)	186	369	517

The last row of the table shows the duration of the elements in milliseconds. Although both first and last elements are bi-syllabic, the figures for duration show that the first element is spoken nearly three times faster than the third element, which features final lengthening. This difference in duration/speed between first (proclitic) and last (enclitic) is a typical one (cf. Cruttenden, 1997: 21): and it militates against these elements being perceived as rhythmical on their own. This is why Roach (1982) removed such elements from his data before commencing measurement.

Double-prominence tone-units

As with single prominence tone-units, double-prominence tone-units are too short for perceived rhythmicality. The explanation of why this is so begins with the properties of a double-prominence tone-unit, which are shown in Table 3.

Table 3. Properties of a double-prominence tone-unit.

		1	2	3	4	5
1	*elements*	*proclitic*	*onset*	*interval*	*tonic*	*enclitic*
2	*words/syllables*	he's	HA	ving a house	BUILT	for her
3	*duration (ms)*	148	122	493	216	442
4	*duration of trochees*	xxx	615		658	
5	*duration of iambs*	270		709		xxx

Table 3 shows that a double-prominence tone-unit has a structure of five elements: two compulsory prominent elements (*onset, tonic*) and three optional non-prominent elements (*proclitic, interval, enclitic*). In the example (tone-unit 18 from MA), all the elements are realised. Duration is shown in row 3, and row 4 shows the duration of a metrical analysis of this tone-unit in trochaic feet: element 1 has to be omitted because the trochee has to start with a salience[1]. The figure of 615 represents the duration of the trochee that includes both the onset and interval; the figure 658 represents the duration of the trochee that includes both the tonic and the enclitic.

In a double-prominence tone-unit, there is only one inter-stress interval (element 3). If therefore we take the threshold for rhythmicality to be the occurrence of two matching events, and that these events should be interstress intervals, it is clear that the double-prominence tone-unit is too short to be rhythmical.

However, the five elements make it possible for prominent and non-prominent elements to pair up into metrical feet (the trochees in row 4, the iambs in row 5). Note that such pairings will (in a tone-unit with all five elements realised) leave out one non-prominent element – either the proclitic or the enclitic element. The durations of the trochaic feet (615 & 658 ms) are sufficiently close for them to be perceived as similar in length, the 043 ms difference is not sufficiently large for it to be noticeable (cf. Lehiste, 1979). The duration for the second iambic foot is well over twice the length of the first iambic foot – and they are unlikely to be heard as matching events.

Despite the existence of two matching trochaic feet this tone-unit was not identified as rhythmical in the study conducted by Cauldwell (2000). We will discuss why this might be so after looking at the structure of a triple-prominence tone-unit in the next section.

Triple-prominence tone-units

The structure of a triple-prominence tone-unit, with a sample tone-unit from MA (19) is given in Table 4.

Table 4. The structure of a triple prominence tone-unit.

		1	2	3	4	5	6	7
1	*elements*	*proclitic*	*onset*	*interva1 l*	*middle*	*interval 2*	*tonic*	*enclitic*
2	*words/syllables*	which she'd	MOVE	into in	FIVE	months	TIME	[pause]
3	*duration*	416	168	387	270	334	414	280
4	*duration of trochees*	xxx	555		604		694	
5	*duration of iambs*	584		657		748		xxxxx
6	*sps*	4.8	6.0	7.8	3.7	3.0	2.4	

As can be seen from Table 4, the triple-prominence tone-unit has a structure of seven elements: three compulsory prominent elements (*onset, middle, tonic*) and four optional non-prominent elements (*proclitic, interval 1, interval 2, enclitic*).

On this occasion, the sample tone-unit does not have all the elements realised: the tonic prominence *time* is the last element, the enclitic element is not filled.

This tone-unit happens to be followed by a pause, so I have included the duration of the pause (280 ms) in the last column of the table – though it is by no means certain that the pause had any effect on the perception of rhythm.

This tone-unit was perceived to be rhythmical, and there are three candidates for the durational correlates for this perceived rhythmicality: interstress intervals (of which there are two); trochaic feet, and iambic feet (of which there are three each). As far as inter-stress intervals are concerned, interval 2 is shorter than interval 1 by 53ms, but this difference is within the limit for just noticeable differences established by Lehiste (1979) therefore they could be heard as equivalent in duration.

Both sets of metrical feet, the trochees and the iambs, increase successively in duration: the trochees in steps of 49 and 90 ms; the iambs in steps of 73 and 91ms. It is possible either that: (a) these feet are heard as equal in duration, or that (b) the progressive increments in duration are heard as a pattern that is interpreted as rhythmical.[2]

In this triple prominence tone-unit, any of the three candidates for rhythmicality could constitute the matching events which trigger perceptions of rhythmicality: which of the three it might be is not the concern of this paper. What is clear is the fact that the triple-prominence tone-unit is long enough to provide ample material to trigger perceptions of rhythmicality.

It is necessary to return to the question of why the threshold for perceived rhythmicality should be three – rather than two – events. One reason is suggested in Lehiste's (1979) research into the perception of differences in lengths of non-speech sounds. One of her findings was that people were most sensitive to durational changes of the third interval in a series of four: perhaps it is the case that rhythms do not become perceptible until the third element (either a iambic or a trochaic foot) occurs to confirm the matching of the first two elements. Another reason comes from teaching music: when practising the rhythm of two beats, you need to play the third in order to get the duration of the second one correct: in the case of rhythmicality of speech, there has to be a third event – either itself matching a preceding pair of events, or simply to mark the end of a second matching event. The occurrence of the third prominence in a triple-prominence tone-unit is thus crucial to rhythmicality: it marks the end of the second interstress interval, it starts the final trochaic foot, and it ends the final iambic foot.

The rarity of triple-prominence tone-units

An association between perceived rhythmicality and triple-prominence tone-units would mean that the occasions on which rhythmicality can be perceived in speech are relatively rare. This is because, as can be seen from Table 5, triple-prominence tone-units account for only a small percentage of tone-units.

Table 5. Percentage of different sizes of tone-units.

			Size of Tone-units				
Texts	Length	Tone-units	incomplete	single	double	triple	quad
Moving Again	*0:21*	*24*	0%	46%	46%	4%	4%
Houses in New Zealand	*1:51*	*96*	1%	45%	47%	6%	1%
Voices in the University	*29:36*	*1603*	6%	46%	43%	5%	0%

Table 5 gives the percentages for the different sizes of tone-units in three related texts of increasing size, *Moving Again* (21 seconds), *Houses in New Zealand* (1 minute 51 seconds), and *Voices in the University*, (30 minutes).

Table 5 shows that for all three texts, single and double prominence tone-units account for very close to 90% of all tone-units: triple prominence tone-units account for only 5% of the total. The quadruple prominence tone-unit in *MA* disappears from the percentages in the last row as it is only one out of a total of 1603 tone-units.

Thus to the extent that rhythmicality is associated with large tone-units, it is likely to be rare. I would emphasise that it is possible for large tone-units not to be associated with perceived rhythmicality, and for other sizes of tone-units either alone (larger, quadruple-prominence tone-units) or in combination (single and double-prominence tone-units run together) to be so associated (cf. Cauldwell, 2000). Nevertheless, larger tone-units, and combinations of smaller tone-units with the right characteristics for perceived rhythmicality are themselves at least as rare as triple prominence tone-units.

A caveat

It is unlikely that a different methodology for identifying rhythmicality, and a different definition of the tone-unit would lead to replication of the evidence reported above. Couper-Kuhlen (1993) allowed her informants unlimited re-listening for a text of a similar length to that analysed in Cauldwell (2000) – *Houses in New Zealand:* Couper-Kuhlen's informants identified 48 isochronous chains, Cauldwell's informants (allowed only two listenings) identified only 8 rhythmical patches – adoption of Couper-Kuhlen's methodology would have resulted in a much greater number of rhythmical patches.

A different definition of the tone-unit, such as those adopted by Crystal (1969), Halliday (1967, 1994) or Brown at al. 1980 would also produce very different outcomes to both the figures and the discussion above. Crystal and Halliday allow for more than one tone per tone-unit, and allow pauses to occur within tone-units. Brown et al. used the concept of the pause-defined unit. So analyses using their approaches would have resulted in fewer, longer tone-units.

Therefore the reader should be aware that the evidence discussed is to some extent an artefact of the discourse approach. However, I contend that this approach (in its attempt to account for real-time perceptions, and to incorporate discourse factors in the analyses) has greater validity vis-à-vis spontaneous speech than other approaches.

Spontaneous speech is irrhythmic

Spontaneous speech is irrhythmic: it occurs in a series of short bursts – tone-units – most of which (close to 90%) are too short to trigger perceptions of rhythmicality. Each tone-unit has a different tempo from its neighbours, and its boundary is marked by tempo-disrupting phenomena (pauses, lengthening of tonic and post-tonic syllables) therefore whatever incipient rhythmicality there might be in one tone-unit is disrupted by boundary phenomena and the incipient rhythmicality of the one that follows.

The main determinant of the rhythms of the stream of speech are the decisions made by speakers concerning the lexical choices and how to package them into tone-units. Rhythmicality can occur in two ways: it can either be *coincidental* (as suggested by Classe, 1939) or it can be *elected*.

Coincidental rhythmicality

Coincidental rhythmicality is a short-lived unintended side-effect of speech which is pursuing social purposes. It typically occurs in triple-prominence tone-units (or larger) which provide optimum conditions for the perception of rhythmicality (cf. Section 5). These conditions occur as a result of higher-order discourse decisions: prosodic – the division of the stream of speech into tone-units; and lexico-syntactic – the choice of wording to realise meanings.

Division into tone-units

The following clause was produced as a quadruple-prominence tone-unit in MA and is rhythmical:

7	▼he's CURrently THINking of Moving aGAIN

With four prominences, this is an unusually large tone-unit. The speaker could have uttered this clause in two double-prominence tone-units:

7a	►▼ he's CURrently THINking
7b	► of MOving aGAIN

If she had done so, then (for reasons explained in Section 5 above) the clause is far less likely to be perceived as rhythmical. Moreover, the speaker could equally well have produced this clause as three tone-units:

7c	▲ he's CURrently ♦♦
7d	▲ THINking of ERM ♦♦
7e	► MOving again

Rendering a clause into three tone-units may seem an unlikely choice, but such a rendition would parallel the speaker's choices in the opening tone-units of MA:

1	▲ he DOESn't ♦♦
2	▲DO it ERM ♦♦
3	► for the MOney he's going to MAKE

The three tone-unit version (7c-e), with pauses, would be even less likely to trigger perceptions of rhythmicality. In traditional accounts of the rhythms of English the four-prominence tone-unit would be regarded as a 'normal' way of packaging the clause. From the discourse perspective, it is a highly unusual way of doing so. The large tone-unit seems to indicate the successful delivery of a pre-planned 'chunk' which may have been uttered before, in telling other people about the uncle in New Zealand. Had she packaged these words differently – more 'normally' – then it is far less likely that they would have been perceived as rhythmical.

Other words, other word-accents patterns

For this tone-unit the speaker, Gail, could have chosen other words with other word-accent patterns to realise existentially equivalent meanings (Brazil, 1997). The meaning of *currently* could have been realised by 'now', and the meaning of *moving again* by 'doing it all over again', thus producing the different word-accent patterns shown in Table 6.

Table 6. Differing patterns of word-accents in existentially equivalent clauses.

word accents	he's	CUR rent ly	THIN king of	MO ving a	GAIN
Rhythm	x	X x x	X x x	X x x	X
word accents	he's	NOW	THINking of	DOing it all over a	GAIN
Rhythm	x	X	X x x	X x x x x	X

The latter version is less likely to be perceived as rhythmical, particularly as there are unequal numbers of syllables (respectively none, two and four) in the intervals between the prominences. And as Dauer (1983) and Halliday (1994) note, the length of an inter-stress interval is proportional to the number of syllables it contains. It is of course possible to speak the latter version in a such a way that it will be heard as rhythmical, but to do so would require the speaker to devote attention to counteracting the natural flow of speech by resisting the pressure to make inter-stress intervals proportional to the number of syllables – it would require a conscious focus on producing a timed utterance.

Elected rhythmicality

Such conscious attention on producing rhythmical utterances results in elected rhythmicality. Examples of elected rhythmicality occur in scanning readings of verse 'i WILL aRISE and GO now'; or when speakers come close to reciting in reading aloud the titles of books 'the seLECted LEtters of PHIlip LARkin'; or uttering ; idiomatic, or semi-idiomatic material such as 'he can TURN his HAND to ANything'. Notice that these examples feature triple-prominence tone-units or larger, with the potential for producing a sufficient number of matching events to trigger perceptions of rhythmicality (cf. Section 5).

Thus speech can be made rhythmical as the result of a conscious decision to recite, as with a style of verse reading known as a 'scanning' reading (Jakobson, 1960; Cauldwell, 1994); but rhythmicality is rarely the focus of speakers' and hearers' attention. Scanning readings of verse, conventional demonstrations of stress-timing, and classroom pronunciation drills (such as that by Underhill, 1994 mentioned above) are, in actuality, demonstrations of the plasticity of speech – they are *not* proof that language is stress-timed.

Speech is *plastic* in the sense that at every moment the speaker is propelling and shaping the flow of speech, it can be shaped in an infinite variety of ways, within the limits of the requirements for comprehensibility.

Perception revisited: the function of irrhythmicality

Scholars often concede the evidence against SSH in speech production, but then typically argue that it is a phenomenon related to perception. They do so with justification. As long ago as 1977 Lehiste argued that although most studies of isochrony in speech production had found only counter-evidence, isochrony still had a role in perception. Perception evidence had to be taken into account because 'sentences that are not produced with absolutely isochronous intervals between stresses may still be perceived as if the interstress intervals were identical' (1977: 258). In experiments Lehiste (1979) found that hearers could not perceive differences in length of sounds less than 30 milliseconds, but in certain circumstances sounds had to differ in length by 100 milliseconds before hearers could perceive differences.

Following Lehiste's work, it is now generally accepted that with language, as with any 'sensory material in the time domain', the human cognitive system imposes 'a constructed rhythm' (Laver, 1994: 524). We thus have to consider the possibility that production-SSH can be replaced by a perception-SSH: '*As far as is known, every language in the world is perceived with one kind of rhythm or with the other ...*'.

Two of the major issues concerning rhythm in speech perception research are first, how hearers pick out words from the stream of speech ('speech elementation'); second, the use of rhythmic expectations to predict the location of accents as an aid in processing meaning (the 'attentional bounce hypothesis'; Pitt & Samuel, 1990). Research typically focuses on subjects' judgements of many short samples of specially recorded and edited speech under laboratory conditions.

Typical statements concerning speech segmentation are that hearers of English expect trochaic rhythmic patterns in speech (Allen, 1975; Echols,

Crowhurst & Childers, 1997); that hearers work on the assumption that there is a word boundary before each stress (Cutler & Norris, 1988); and that hearers of French expect iambic rhythmic patterns (Allen, 1975:78).

The assumptions on which such research is based are familiar: it is a common assertion in studies of perception that perceived rhythm differs from language to language (e.g. Allen, 1975 78; Cutler, 1994: 80). Allen, in talking of 'languages with strong tonic accent (e.g. English and German)' and 'languages with accent based on duration (e.g. French)' reveals that he is an adherent of some form of SSH, as does Cutler (1994: 80). Readers of the literature on speech perception could be forgiven for thinking that it is a fundamental assumption – a given – and is not an issue that it is necessary to investigate.

Perception-SSH is thus largely (though not entirely) a mirror image of production-SSH, and thus falls victim to the same arguments. This is because underlying the view that speakers of a particular language have a unique way of perceiving that language, is the fact that the shape of this perceptual predisposition is determined by the input they get from speakers of that language. So the reason that native speakers of French (say) perceive French as syllable-timed (even when exposed to non-timed input) is because they have become attuned to its 'phonological syllable-timed-ness' through exposure to French speakers. In other words, perception-SSH and production-SSH are different manifestations of the same phenomenon, and the arguments that hold for one, hold for the other.

However, the arguments against production-SSH do not entirely do away with issues of perception of rhythm; we have to account for the ability of the human perception system to impose rhythm on irrhythmic material.

Thus speech production is characterised by irrhythmicality, and speech perception is characterised by rhythmicality brought about by constructive ordering. I want to suggest that the production/irrhythmic perception/rhythmic balance provides a necessary tension for effective communication.

If the majority of utterances in English were spoken rhythmically, it would be difficult for hearers to attend to speech as a connected set of units of meaning. The rhythm would draw attention to itself and distract the hearer's attention from meaningful choices: it would, in other words, be English in oblique orientation (cf. Brazil, 1997). Listeners to Halliday's (1970) 'rhythmical prose narratives', and to scanning readings of verse will know that the presence of a perceptible rhythm attracts attention to itself and away from the processing of the text as meaning. This is because – as Bolinger (1986: 47) argues – in allowing 'the mechanical phenomenon of even rhythm ...[to]...assert itself…' speakers will be heard to be speaking 'routinely and mechanically'.

Bolinger notes that 'stylized intonation' (e.g. it's NEver too LATE to MEND) has this routine and mechanical feel to it. But 'stylized intonation' is a special case of elected rhythmicality (cf. 6.2 above), and is therefore not an appropriate speech-style on which to base generalisations about spontaneous speech. He expresses the worry that 'this sort of sing song is just the kind of intonational frame that a classroom drill is apt to fall into', and suggests that the use of such drills 'has helped to make us see English accentual rhythm as more regular than it really is' (48).

Bolinger goes on to suggest two reasons why spontaneous speech is not 'routine and mechanical': first he states that 'one thing the adjustment is never allowed to interfere [with] is our meaning' (1986: 47); and secondly, 'the words we want to emphasize are often irregularly spaced, which means that the

number of syllables may be radically different from measure to measure' (1986: 47). These views have been borne out by the research reported in the preceding sections.

The view of irrhythmicality in speech being functional goes counter to the attentional bounce hypothesis. Martin (1972, cited in Allen 1975: 84) suggested that temporally patterned, and therefore temporally predictable speech, aids perception by enabling attention to be cycled between input and processing, whereas in the absence of patterning, perception 'would seem to require continuous attention'. Recent findings relating to the attentional bounce hypothesis provide counter-evidence (cf. Cutler, Dahan & van Donselaar, 1997, for a summary) because, it is now recognised, prosodic structure 'might only rarely be such as to produce the sustained regularity which ... listeners need if they are to exploit the predictability' (Cutler, Dahan & van Donselaar, 1997, p173-174).

A *lack* of a regular rhythm in speech production is essential for effective communication. In other words the irrhythmicality of spontaneous speech is functional. If the rhythms of speech were not fleeting and ever-changing, speakers might find it difficult to hold the attention of hearers: because, instead of attending to selections of meaning they would be distracted – by the pattern of an established rhythm – from attending to the communication of meaning which is the purpose of most speech. The *non-occurrence* of a continued rhythm of any sort could therefore be viewed as a necessary feature of any co-operative purpose-driven spontaneous speech.

Conclusion

Allen (1975) asserted 'No one doubts that spoken language has rhythm' (75). With SSH refuted, we now have to doubt that speech has isochronic rhythm. Allen's opening reveals a bias towards confirmation of rhythmic hypotheses, rather than a refutation stance. Along with most scholars he confirmation-minded. A scholar biased the other way, a refutation-minded scholar, would be much happier with the research evidence.

English has been the language with which I have outlined the hypothesis for the functional irrhythmicality of spontaneous speech. I expect that analyses of other languages using a tool such as Discourse Intonation – or any other tool sensitive to discourse factors – will show that spontaneous speech in any language will have the features identified for English.

Language learners, teachers, and scholars are quick to attribute the causes of rhythmic phenomena to SSH. This is because in speech, many language events (elements, syllables, words, stresses, non-stresses) occur in quick succession: these events happen in a temporal dimension, they are therefore amenable/vulnerable to being constructed into rhythmical units by the hearer. It is likely that such hearers interpret differences as 'timing-related' or 'rhythm-related' (that is attribute the differences to some aspect of SSH) because they prefer the clear-cut shorthand of SSH to the more complex explanations that the evidence of spontaneous speech requires. It is easier and more acceptable to attribute inter-language differences to SSH than to differences in syllable structure, word-accent, and vowel reduction (Dauer, 1983: 55) or 'elemental sonority, syllabic weight and lexical stress in the lexicon of the language, and of

the pragmatic use of the lexicon in the utterances of that language' (Laver, 1994: 527).

The continued presence of the refuted hypothesis, that has become hard-wired into our thinking, is an obstacle to progress in understanding the nature of spontaneous speech: long-refuted, it should be now discarded. Life without the stress and syllable-timing hypothesis will be more difficult, but it should make possible real advances in the understanding of spontaneous speech.

Endnotes

[1] In Hallidayan approaches to rhythm, the analyst typically adds a 'silent ictus' to the analysis, but as the tone-unit is not preceded by a pause such an addition seems unwarranted in this case.

[2] Lehiste (1979) did not study reference durations larger than 500ms, and the durations in question range from 0.584ms to 0.748ms.

References

Abercrombie, D. (1964) Syllable quantity and enclitics in English. In D. Abercrombie, D.B. Fry:A.D. MacCarthy, N.C. Scott & J.L.M. Trim (Eds.), *In honour of Daniel Jones: Papers contributed on the occasion of his eightieth birthday.* (p216-222). London: Longman.

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Allen, G. D. (1975). Speech rhythm: Its relation to performance universals and articulatory timing. *Journal of Phonetics, 3,* 75-86.

Ball, M. J. & Rahilly, J. (1999*). Phonetics: The science of speech*. London: Arnold.

Bolinger, D. (1986). The English beat: Some notes on rhythm. *Studies in Descriptive Linguistics, 15,* 36-49.

Brazil, D. (1995). *A grammar of speech.* Oxford: Oxford University Press.

Brazil, D. (1997). *The communicative value of intonation in English.* (2nd Edition). Cambridge: Cambridge University Press.

Brown, G. (1990). *Listening to spoken English.* (2nd edition.). Harlow: Longman.

Brown, G., Currie, K. & Kenworthy, J. (1980) *Questions of intonation*. London: Croom Helm.

Cauldwell, R. T. (1994) Discourse Intonation and recordings of poetry: Philip Larkin reads 'Mr. Bleaney'. [Doctoral dissertation]. The University of Birmingham

Cauldwell, R. T. (1996). Stress-timing: Observations, beliefs, and evidence. *Eger Journal of English Studies, 1,* 33-48

Cauldwell, R. T. (1997). *Voices in the university.* [Book and Cassette]. Birmingham (UK): The University of Birmingham, English for International Students Unit

Cauldwell, R. T. (2000) Perceived rhythmicality in spontaneous speech: A discourse view. [paper submitted for review].

Classe, A. (1939). *The rhythm of English prose.* Oxford: Blackwell.

Couper-Kuhlen, E. (1986). *An Introduction to English prosody*. London: Edward Arnold.

Couper-Kuhlen, E. (1993). *English speech rhythm: Form and function in everyday verbal interaction.* Amsterdam: John Benjamins.

Cruttenden, A. (1997). *Intonation*. (Second Edition). Cambridge: Cambridge University Press.

Crystal, D. (1969). *Prosodic systems and intonation in English.* Cambridge: Cambridge University Press.

Crystal, D. (1996). The past, present and future of English rhythm. *Speak Out, Newsletter of the IATEFL Pronunciation Special Interest Group, 18,* 8-13

Cutler, A. (1994) The perception of rhythm in language. *Cognition, 50,* 79-81.

Cutler, A., & Norris, D. G. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 113-121.

Cutler, A. Dahan, D. & van Donselaar, W. (1997) Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40 (2),* 141-201.

Dalton, C. & Seidlhofer, B. (1994). *Pronunciation.* Oxford: Oxford University Press.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11,* 51-62.

Echols, C. H., Crowhurst, M. J., & Childers, J. B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language, 36,* 202-225.

Halliday, M. A. K. (1967). *Intonation and grammar in British English.* The Hague: Mouton.

Halliday, M. A. K. (1970). *A course in spoken English: Intonation.* Oxford: Oxford University Press.

Halliday, M. A. K. (1994). *An introduction to functional grammar.* (2nd Edition). London: Edward Arnold.

Jakobson, R. (1960). Closing statement: Linguistics and poetics. In T. A. Sebeok (Ed.), *Style in language*. Cambridge, MA: MIT Press, 350-377.

Jones, D. (1960). *An outline of English phonetics.* (9th Edition; 1st Edition 1918). Cambridge: Cambridge University Press.

Knowles, G. (1987). *Patterns of spoken English: An introduction to English phonetics.* London: Longman.

Knowles, G. (1991). Prosodic labelling: The problem of tone group boundaries. In S. Johansson & A. Stenström (Eds.), *English computer corpora: selected papers and research guide*. Berlin: Walter de Gruyter, 149-163.

Laver, J. (1970). The production of speech. In J. Lyons (Ed.) *New horizons in linguistics.* Harmondsworth: Penguin.

Laver, J. (1994*). Principles of phonetics.* Cambridge: Cambridge University Press.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics, 5,* 253-263.

Lehiste, I. (1979). The perception of duration within sequences of four intervals. *Journal of Phonetics, 7,* 313-316.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behaviour. *Psychological Review, 79,* 487-509.

O'Connor, J D. (1973). *Phonetics.* Harmondsworth: Penguin.

Pike, K. L. (1945). *The intonation of American English.* Ann Arbor: University of Michigan Press.

Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human perception and performance, 16 (3),* 564-573.

Roach: (1982). On the distinction between 'stress-timed' and 'syllable-timed' languages. In D. Crystal (Ed.) *Linguistic controversies, Essays in linguistic theory and practice..* London: Edward Arnold, 73-79.

Rogers, H. (2000). *The Sounds of language: An introduction to phonetics*. Harlow: Longman

Underhill, A. (1994). *Sound foundations.* London: Heineman