

THE PAST, PRESENT AND FUTURE OF ENGLISH RHYTHM

David Crystal

For over 50 years, scholars trying to describe the general rhythmical impression conveyed by a language have worked in terms of two broad categories. In one type, the overriding impression is of phonetic contrast between strong and weak beats in a sequence of syllables; in the other, it is one of sequences where the syllables are being uttered at similar levels of phonetic prominence. Several metaphors and analogies have been employed to characterise these different auditory impressions: terms such as 'bouncing', 'heart-beat', 'morse-code like', and 'tum-te-tum' have been used for the former; 'staccato', 'machine-gun like', 'pattering' and 'rat-a-tat' have been used for the latter.

More technically, these impressions have motivated the distinction (Pike 1945: 35) between a 'stress-timed' and a 'syllable-timed' language. The former term suggests that the stressed syllables fall at regular intervals, whether they are separated by unstressed syllables or not; the latter term suggests that all syllables occur at regular time intervals, whether they are stressed or not. English is cited as a typical example of the former (also Russian, Arabic, Portuguese, Swedish, Thai, German), and French of the latter (also Greek, Italian, Spanish, Hindi, Yoruba, Telugu, Indonesian). Syllable-timing appears to be much more widespread, in the world's languages.

The distinction, which has been widely cited in ELT manuals, is an extremely crude one, and in its bare form is almost certainly wrong. It is unlikely that all languages would fall neatly into any two simple types, whatever the domain of enquiry. And indeed, those who have looked carefully at the matter have found that languages vary greatly in the amount of stress-timing or syllable-timing they employ, and concluded that the distinction is more a continuum than a divide. It has also been pointed out that there is more to rhythmical

prominence than timing: segmental sonority, syllabic weight¹, and lexical stress are major factors in affecting auditory impressions of rhythm (Laver, 1994: 527). And even within a language, both stress-timing and syllable-timing can be heard in varying degrees. The more formal the speech, for example, the more rhythmical it is likely to be. Roach concludes (1982: 78; and see also, in the pedagogical domain, 1991: 120-1).

There is no language which is totally syllable-timed or totally stress-timed - all languages display both sorts of timing ... [and] different types of timing will be exhibited by the same speaker on different occasions and in different contexts.

Others agree. Dauer (1983), in an empirical study, concluded that the basis of the distinction is 'nothing to do with the durations of interstress intervals' (54), but is more a matter of 'what goes on within rhythmic groups, the characteristics of successive syllables and their interrelationships' (60). Fletcher (1991: 194-5) asserts: 'Many researchers now acknowledge that languages conventionally described as syllable-timed or stress-timed, share acoustic timing features originally thought to be characteristic of one timing typology as opposed to another'. And Pike himself, in introducing the original distinction, allowed both types within English:

English also has a rhythmic type which depends to a considerable extent upon the number of its syllables, rather than the presence of a strong stress, for some of its characteristics of timing...

He gives just one example: the 'spoken chant' of cat-calls ('Susie is a tattle tale'), and adds 'in English, however, the type is used only rarely'.

Despite these problems, the distinction continues to be widely cited, and one can see why. It has helped to give some structure to a domain which seemed previously incapable of

classification. It also has an intuitive appeal. Notwithstanding the variation which exists, and the phonetic complexity which lies at the heart of the notion, we do usually receive a coherent auditory impression of rhythmical consistency when we listen to a language, and our attention is often drawn to it. There is an immediacy of recognition which follows, for example, an English comedian's pastiche of a French accent (or, for that matter, a French comic's pastiche of a French accent, as in Antoine de Caunes in TV's Eurotrash), where the use of a prominent syllable-timed rhythm seems enough to convey identity, despite invariably execrable segmental accuracy. The distinction has also a major relevance to phonetics teaching, as Pike and others have pointed out, where failure to get to grips with the rhythmicality issue can lead to unintelligibility:

For Latin-Americans, the shift from their tendency toward a syllable-timing rhythm to the English normal stress-timing rhythm is highly necessary, but comprises one of their greatest problems...(Pike, 1945: 35)

Most Nigerian languages are 'syllable-timed' languages...However close a Nigerian speaker approximates to consonant and vowel qualities, if he uses 'syllable-timing' when speaking English, he may well be faced by total incomprehension on the part of any listener who is a native speaker of English. (Dunstan, 1969:29-30).

The distinction therefore retains a pragmatic value, and insofar as it helps us to see language phenomena more clearly, it will stay until a more-refined classification of rhythmical types arrives on the phonetic scene. Recent accounts indeed are beginning to develop alternatives, recognising other categories (e.g. mora-based, to account for Japanese). And Laver, for example, having considered the theoretical possibilities, suggests replacing *syllable-timed* by *syllable-based*, and *stress-timed* by *stress-based* (1994:528-9) - a sensible suggestion which I will follow here.

The past

Most people would be happy to accept Pike's judgement that English is essentially stress-based, with just occasional use of syllable-based speech. English seems to have been this way for at least 500 years. We actually know next to nothing about the early history of English rhythm, and what we do know is inspired by phonological analyses of word stress. We can deduce from the processes of segmental sound change what rhythmical contrasts must have been operating within a word (particular sequences of stressed and unstressed syllables), but this is phonological, not phonetic reasoning. We know that a contrast was there, but not exactly how it sounded. Similarly, analyses of the metrical structure of poetry suggest what must have been happening with the poetic line - we know that most Anglo-Saxon lines of poetry comprised four stressed elements divided into two two-stress phrases linked by alliteration - but we can say nothing about the extent of the phonetic contrast between those syllables, nor do we know exactly how these patterns were read aloud (a wide range of musical possibilities exist) or how these rhythms were related to those of everyday speech.

In Indo-European and early Proto-Germanic, the strongest stress (or accent) in a word was not restricted to a single position (it was 'free'). 'Free' does not mean that the location varied randomly: it means that there was no single location in word structure for the accent. (Modern English has a 'free' accentual system, in this sense; Welsh, by contrast, does not, for there most polysyllabic words have the accent on the penultimate syllable.) One word might have its accent on the first syllable of its root; another on its last; a three-syllable word might have it on the penultimate; in some cases a prefix or suffix could attract it. We can assume, on general auditory phonetic grounds, that such a language must have had a broadly syllable-based character. Indo-European relied greatly on inflectional endings, but endings need a respectable level of prominence if they are to carry grammatical weight. Indeed, the falling away of inflectional endings in English

has long been attributed to a change of accentual weight within the word.

In early Germanic, it has been established that there was an accentuation shift (sometimes called the Germanic Stress Rule) which led to the accent falling on the first syllable of the lexical root, and rarely on prefixes and suffixes ('*gode*, *ge'worden*). This rhythm applied regardless of the length of syllable or the number of syllables in the root. The consequences of a strong primary accent led to a weakening of other vowels - diphthongs became monophthongs, long vowels became short, short vowels disappeared - and by the time Old English came to be written down, the reduction in auditory distinctiveness was producing considerable confusion over the way inflectional endings were spelled (a given word ending might be spelled *-an*, *-on*, or *-en*, for example). Anglo-Saxon speech must have sounded increasingly stress-timed, as a consequence.

From Old English on, and especially after 1066, Romance words entered the language in ever-increasing numbers: indeed, from a lexical point of view, English eventually came to be more a Romance than a Germanic language. The word-stress system became vastly more complex, as a result. Most of the new words reflected a different accentual system, and the Germanic Stress Rule soon found itself in a minority. What replaced it was a Latin-based system, in which the main accent was assigned to the heaviest syllable closest to the end of a word. In the development which has since been called the Romance Stress Rule, final heavy syllables were accented (*ar'rest*); if the final syllable was light, then the penult was accented (if it was heavy, as in Chaucer's *pilgrimage* / *pilgrɪ'ma:ʒə* /); if not, then the antepenult was accented, regardless of weight (*bachelor*). This was the foundation of the Middle English stress system (for further details, see Lass, 1992: 83, ff).

Many of the complicating features of present-day English stress had their origins during this period (see Poldauf, 1984). Romance words

had their stress shifted towards the beginning, but kept a remnant of their original stress at the end (as a secondary stress), e.g. *catalogue*, *quadruped*. The trend was not totally consistent, and was conditioned by various factors, such as word class (it was especially strong on nouns). Major segmental sound changes were taking place during the period, and several of these affected syllable structure. The presence of certain types of affixes also restricted the shift (*enter'tain*, *intro'duce*). The trend was also stronger in British English than in American English, as can be heard today in AmE *pre'miere*, *ba'ton*, *ca'fe*, *mat'i'nee*, *pla'teau*, *ver'mouth* and many other words. We can conclude that the rhythm of Middle English speech must have been much more varied, as a result - to use the above terminology, the language became more 'stress-based' than 'stress-timed'. And so it stayed.

The present

The dramatic expansion of English around the world during the 20th century has begun to alter this scenario (Crystal, 1995b). Most (? all) of the new varieties which have emerged in second language contexts seem to be syllable-based, as the following quotations suggest:

Most African languages have a strong preference for syllables of the structure /CoV/ (zero or more consonants plus a vowel, with no final consonant).... For those Africans whose first language is syllable-timed (as many are), the resultant pronunciation of a word such as *society* [*sə.sai.je.ti*] is very different from what is heard in England or America... (Wells, 1982: 642)

.. in African English as a whole it is very common for pronouns, auxiliary verbs, prepositions and so on to be stressed in running speech; and a further consequence of this is that no weak forms are used. Compare the sentence *she's a rascal, you know* in its West African syllable-timed form ['ʃi ze 'raskal 'ju no] ..and its English-English stress-timed form [*ʃizə 'rɑ:skəl ju nəʊ*]. The use of tone

rather than stress, and of syllable-timing rather than stress-timing, combine to make some African English strikingly different from other varieties in pitch and rhythm. (ibid.: 643-4).

One of the most prominent features of Singaporean English is the use of syllable-timed rhythm, as against the stress-timed rhythm of other accents. This is something that applies throughout all that is said. (ibid: 646).

Rhythmically, Standard Filipino English is syllable-timed rather than stress-timed. It is spoken with 'a delivery which gives the impression of a clipped or staccato rhythm ...' (ibid.: 647)

Rhythm [in Hawaiian English] varies from the stress-timing usual in English to the syllable-timing characteristic of much Hawaiian Creole ... it is often criticised by local teachers of English as 'choppy' or 'staccato', and involves not only timing but also stress. Carr ... gives the example of the sentence to *see if I could get into law school*, with ... a creole syllable-timed rhythm involving stress on every word ..(ibid: 651).

Type B Indian speakers [i.e. those with a rhotic accent] sometimes use patterns of accentuation that are different from the patterns in native English. The rhythm is also different from the stress-timed rhythm of native English. They sometimes omit the accent on words that are normally accented and place the accent on common structural ... words that should not be accented. ... Prabhakar Babu ... reported that the rhythm in Indian English was neither syllable-timed nor stress-timed. Dhamija ... reported that in Rajasthani English stressed syllables occurred too frequently in connected speech, thus producing a kind of halting and jerky movement. (Bansal, 1990: 227).

SABE [South African Black English] maintains an unchanging rate of syllable utterance (tempo) over given periods of time, unlike SAE [South African English] where rate of utterance is likely to vary between one tone unit and another ... In this particular respect, therefore, SABE shows

'syllable-timed' characteristics. (Lanham, 1990: 250)

These general impressions must be interpreted cautiously. Several of the observers point out that the situation is not as simple as the single-label description of 'syllable-timed' suggests. Also, as more-detailed studies emerge, it is likely that descriptive generalisations will need to be refined. For example, the impression that structural words are stressed in syllable-timed speech will need qualification, as some such words tend to attract stress more than others. This is reported for Singaporean English by Deterding (1994), for example, who notes that the stressing of demonstratives and modal verbs is a notable feature of this variety. Another study reports the use of individual strategies in coping with timing in second-language English (Bond & Fokes, 1985). However, the impression that there is some kind of syllable-based speech among second-language English learners is widespread, and apparently affects all areas where new varieties are emerging, in Africa, South Asia and South-East Asia.

A preliminary auditory description of a few sentences of syllable-based speech, taken from these areas, illustrates the kind of work which needs to be done. In each case, there is a general impression of syllable-timing, but the reasons for this impression vary greatly. At one extreme (illustrated by **example 5**, from India) there is a rapid and fluent speech style, in which the unstressed vowels (as would occur in an accent such as RP) are produced with somewhat increased tension, resulting in a greater evenness of articulation than would be found in stress-based speech. At the other extreme (illustrated by **example 1**, from Ghana), there is an erratic speech style, in which only certain words and phrases receive a strongly syllable-based articulation. In this example, there are several such items because of the metalinguistic subject-matter: the speaker is specifically drawing attention to particular words and phrases, and these, being semantically foregrounded, are articulated with

a greater level of syllabic force: 'hea'vy, 'beau'ti'ful, 'pre'tty, 'ex'pres'sive, 'bru'tal'.

Not all of the syllable-based items in this extract are the result of this process. The words 'hy'bri'dized and 'sy'no'nym' seem to stand out simply because they are polysyllabic technical terms (the speaker is a university lecturer). By contrast, the everyday items 'girl'friend, 'somebody, and Gha'naian have less syllabic prominence. This suggests a frequency or familiarity factor: the rarer the word, the more likely it is to be syllable-based; the more familiar the collocation, the less likely it is to be syllable-based. In **example 2**, from Guyana, the syllabic impression is entirely a function of the specialised vocabulary, and because the words *transformation*, *transmutation*, and *transcendence* occur in quick succession, the cumulative syllabic effect is considerable. **Example 3**, from West Africa, also shows the more specialised words attracting extra syllabic prominence ('ele'va'ted, 'rele'ga'ted, o'bliv'ion; whereas such items as 'English 'language, 'African, inter'national and 'national have a more noticeable stress-based rhythm.

The remaining examples in this selection are there to show the relevance of other factors. There is a major contrast in educational background distinguishing **examples 4 and 5**, both from India; the former is a clerk in a business office, the latter is a university lecturer. The main features of **4** are varying speech rate and reduced articulatory clarity in certain words, such as *movements*, *register* and *confidential*, both of which complicate the analysis of an otherwise dominant syllable-based element in this speech (and which also contribute to comprehension difficulty). **Example 6** is the opening of Benjamin Zephaniah's rap 'Crucial the Cat', which I include chiefly to illustrate the remarkable syllabic pressures which are put on individual words when circumstances demand it: 'Sagi'tari'an, 'fishi'tari'an, 'belly, 'territo'ry.

The future

Does it matter which norms are used? It matters insofar as the interaction between stress-based speakers and syllable-based speakers gives rise to problems of comprehension. The point has been made by Bansal, Lanham, Tiffen (1974), and others. Individual words can be misinterpreted by listeners used only to a stress-based system because of a failure to identify phonological structure. Grammatical patterns can be misheard because of the unfamiliar stressing of structural words. The examples are so far all to do with the difficulty faced by stress-based speakers understanding syllable-based speakers. It is unclear whether there are difficulties in the opposite direction, or whether syllable-based speakers misunderstand each other any more than stress-based ones do. But there is no doubt that a problem exists when these varieties come into contact, and that it can often be severe.

The significance of this development for the future of English is now beginning to be appreciated. Large numbers of speakers are involved. The total number of L2 speakers in these areas is now some 350 million (Crystal, 1995b). This is less than the 400 million L1 speakers around the world; but the population growth of the L2 countries is three times that of the L1 countries; so, if current population and learning trends continue, the balance will soon change. Within 10 years there will be more L2 speakers than L1 speakers, and with 50 years there could be up to 50% more. Add to these the varieties of English as a mother-tongue which are syllable-based (creole varieties, such as in Caribbean and West Africa), plus the growing numbers of speakers of English as a foreign language whose speech is syllable-based (probably the majority), and it is plain that very soon stress-based speech will be a minority variety of English, spoken in a small number of world locations.

Does any of this have long-term consequences for L1 speech? Is it likely that, one day, the standard English of Britain or the USA will become syllable-based? As things stand, it is

unlikely. The L2 varieties are not sufficiently prestigious for them to become models for L1 speakers - though some syllable-based speech has come to be part of young people's phonological repertoire (notably in the rap chanting of popular song and in play renditions of alien, Dalek-like speech: see Crystal, 1995a). And it would take only a small number of social changes for the situation to alter - for example, the appointment to high office in L1-speaking countries of people with strong Hispanic or African-American Creole accents. But at present, in most of the L2 countries, the L1 (stress-timed) models are still the prestige models, and there are signs of these being reinforced, as the increased availability of satellite television (e.g. in India) makes access to them more routine. The implications for language learning are considerable: certainly, it suggests a move away from the typical learning situation where students (whose mother-tongue is syllable-based) find themselves regularly taught by teachers (whose English is already syllable-based) with very little opportunity to hear mother-tongue stress-timed speech. If this situation is now in the process of change, we may see an end to the fostering of syllable-based norms through the traditional reliance on second-language pedagogical models.

Whether, in the long term, stress-based speech will replace syllable-based speech, or vice versa, is impossible to say. But attention should also be paid to a third possibility - that L2 learners will become competent in both kinds of speech, continuing to use syllable-based speech for local communication, as a signal of national identity, and switching to stress-based speech for international communication, as a means of ensuring intelligibility. Multi-dialectism already exists in many socio-linguistic situations, and it would be a natural development for it to eventually incorporate rhythmality. Rhythm, after all, is always present in speech - and is therefore much more available as a signal of identity than are segmental phonemes, nuclear tones, lexical items, and other putative markers of style.

Whatever its phonetic basis, its sociolinguistic future seems assured.

1 In phonology, weight is a concept used to distinguish levels of syllabic prominence, based on the segmental constituency of syllables. Syllables can be metrically *heavy* or *light*: a light (or 'weak') syllable is one whose rhyme comprises a short-vowel nucleus alone or followed by a coda of no more than one short consonant (in terms of phonological length, a mora); a heavy (or 'strong') syllable is any other type (its phonological length being greater than one mora). Syllables of structure CVVC or CVCC are sometimes referred to as 'superheavy'.

Examples

1. *'Heavy', in hybridised Ghanaian English means 'beautiful', 'pretty' or 'expressive'. So somebody, your girlfriend, is 'heavy' or 'he has a heavy car' ... and a synonym of that is 'brutal'; 'your girlfriend is brutal'.*

2. *But out of that pain comes transformation and transmutation and transcendence and, ultimately, rejoicing.*

3. *Came to realise that African writers in English language, you know, had come to be elevated. Whereas those African writers who had always been using African languages were relegated to a certain kind of international and even national oblivion.*

4. *Well, I am, there is all the incoming letters and outgoing letters, like, then um file movement, movements in the register and, er these confidential letters and secret letters, like that.*

5. *We've always spoken English, I guess we've taken it for granted, we never really considered it on that level. But it certainly does help, you know, when you are communicating internationally, that you can speak English.*

6. *I used to have a cat that was really smart, not even the government could keep us apart, he was ginger and white, this cat was all right, he would always take to the streets at night, born in December a Sagittarian, loves to eat fish, a fishitarian, when he feels good, he rolls on his belly, makes a funny smell to mark his territory, Crucial the cat, Crucial the cat, used to know a cat, who knew where it was at.*