Title

How Words can be Misleading:
A Study of Syllable Timing and "Stress" in Malay

Authors

Zuraidah Mohd Don, Gerry Knowles & Janet Yong
University of Malaya, Malaysia

Bio-Data

Zuraidah Mohd Don is Professor at the Department of English Language, Faculty of Languages and Linguistics, University of Malaya. Her research interests include English Language Teaching, pragmatics, Critical Discourse Analysis, Corpus Linguistics and prosody. She has published in international refereed journals and her most recent publications include two co-authored books entitled *Malay Word Class: A corpus-based Approach* and *Malay Adverbs: Problems and Solutions*

Gerry Knowles worked for many years in the Linguistics Department at Lancaster University in the UK, where he published in English linguistics, and developed interests in corpus linguistics and speech and language technology. More recently, he has worked in Malay linguistics, and developed the MALEX system, which provides precise linguistic information for applications in technology. He has co-authored two books on Malay, one on word class, and the other on adverbs and adverbials. He is now executive director of Miquest, a company based in Kuala Lumpur and specialising in the applications of technology in the teaching of English in Asia.

Janet Yong has worked as an Associate Professor in the Faculty of Languages & Lingusitics in University Malaya where she taught and researched in phonetics and phonology, sociolinguistics, research methodology and TESL/TEFL. Currently, she is a visiting lecturer in the School of Accounting & Finance in the Hong Kong Polytechnic University; she designs, teaches and does research in business communication.

Abstract

Duration and F0 were studied in a set of 111 Malay words produced by two female native speakers of Malay in order to identify the citation pattern. This preliminary study seemed to provide strong evidence for penultimate stress. Seen in a wider context, the evidence collapsed, and it became clear that Malay does not have word stress at all. The search for syllable timing led to doubts whether the syllable is a relevant unit in Malay prosody. The conclusion is that in view of the lack of complicating factors, Malay is an appropriate language to adopt for the study of prosodic structure, and for the development of automatic techniques for the analysis of spoken corpora.

Keywords Prosody, Malay, word-stress, syllable timing, spoken corpora, automatic annotation

1. Introduction

The longer-term aim of the work reported here is to produce a prosodic annotation system for a corpus of Malay, building on the analysis of an informal corpus of Malay (Zuraidah Mohd Don, 1996), and the prosodic annotation of the Lancaster / IBM Spoken English Corpus (Knowles, Williams & Taylor, 1996). The system has to be designed for the Malay language, and mark patterns which are readily identifiable in the data. In order to make possible the annotation of large amounts of natural data, the system needs to be robust, and as far as possible it should lend itself to automatic processing. In view of the known problems involved in annotating natural conversational data, we have restricted this preliminary investigation to simplified types, namely one-word utterances

produced in citation form.

Malay belongs to a group of languages which are among the most widely spoken (there are about 159 million Malay-Indonesian speakers) and least studied. These languages are generally considered to have a very simple phonology in that the vowel and consonant systems are simple, syllables are simple, and words appear to have a simple prosodic structure. What is not clear is whether an annotation system has to take into account anything in the Malay phonological system corresponding to English word stress. To an English ear it appears almost self-evident that Malay words have stress on the penult, and in some cases such as *kenapa* [kənapə] 'why' the pattern approaches the English prototype. On the other hand, this pattern is not at all clear to Malay speakers, who often find it difficult to say whether their language has word stress or not, and if so what its nature might be. There is also a widespread belief that Malay is a syllable timed language, and while this description seems to be meaningful to native speakers, it is not clear exactly what it means. Understanding patterns at word level is of course the prerequisite for work on the prosody of Malay texts, and here we are confronted with a pattern which is both confusing and elusive.

When the project began we were not in a position to design formal tests or set up formal hypotheses about word stress, and the aim was to find out enough to enable us to do this. The Malay word has two important properties perceived by native speakers which could be associated with stress. First, the final syllable appears to be lengthened, and secondly there is a high pitch towards the end of the word. There are no obvious variations in vowel and consonant quality which could be attributed to stress, and while central vowels certainly do occur, the central quality is part of the normal realisation of the vowel and has nothing to do with stress. In this respect Malay patterns like some North Indian languages. In the case of word citation forms, loudness would appear to vary with pitch, but in view of the complex role of loudness in texts (Zuraidah Mohd Don, 1996) where the correlation is not so close, it would be unwise to push the evidence of individual word forms too far. Another (but less clearly perceived) pattern is that the initial syllable seems to be given some kind of prominence.

2. Interpreting the Results of Previous Works

A problem in working on this paper is that previous work is not self explanatory. Research typically starts off with the assumption that Malay must have word stress like English, and that the task for the researcher is to find it and describe it. Thus it is claimed that word-stress in Malay is weak and so not very prominent (Yunus Maris, 1980), and that it involves length or loudness rather than pitch. Madzhi (1989) claimed to detect four degrees of word-stress, with primary stress falling on the last syllable of isolated and complex words; but this is to impose on to Malay the system put forward by Trager and Smith for English in 1959, and cannot be based on the empirical study of natural Malay data. There is no consistent understanding in the literature of what stress is, and to understand the claims being made it is first necessary to ascertain what notion of stress is being discussed.

In the present situation, instead of building on previous work, we have to explain previous results *post hoc*. Adam and Butler (1948) claimed that stress falls on the penult in Malay except when its vowel is a schwa, in which case stress falls on the final syllable. They also observed that stress shifts to the new penult when affixes are added to a word. Armijan Pane (1950), with the assistance of an Indonesian musician, developed a notation for stress and intonation using three pitch types: initial high pitch, pitch contour, and final pitch. He distinguished what he called syllable stress, word stress and sentence stress, and confirmed Adam and Butler's analysis, except that the interrogative enclitic particle *-kah* had no effect on the position of the stress. This description might sound a little bizarre, but *mutatis mutandis* it could possibly be reconciled with the findings of our own study; *-kah* could well be an additional item ignored by the final cadence.

In an instrumental study of 200 disyllabic words, Verguin (1955) found that the first (i.e. penultimate) vowel was longer in duration, higher in pitch and greater in amplitude than the final vowel. (If the measurements of the penultimate vowel are accurate, they indicate a variant of the word citation form which could belong to a different dialect.)

Kahler (1956) found the stress fixed on the penult when the root word is followed by an enclitic such as *-kah*, *-lah* or *-pun*. Alisjahbana (1967) distinguished dynamic stress, pitch stress and durational stress, claiming that word stress falls on the final syllable, except when it is a clitic pronoun *-ku* or *-nya*. Amran (1984) made a mingographic study of 140 words, and found word stress on the penult in isolated words, and on the final syllable in context. He described a typical stressed syllable as longer and louder than unstressed ones, and having a pitch contour containing a peak of pitch, although the initial pitch could be higher or lower than the final pitch.

Generative accounts (Cohn, 1989; Cohn & McCarthy, 1994; Kenstowitz, 1994; Halle & Idsardi, 1994) place primary stress on the penult, with secondary stress on alternating syllables to the left and on the initial syllable. Suffixes affect stress position, one suffix shifting the stress one syllable to the right; while if there are two suffixes, the first suffix takes the primary stress and a secondary stress falls on the penultimate syllable of the root. It is difficult to assess these claims because they have only a tenuous relationship with what one finds in Malay phonetic data, and in order to understand them one has to be familiar with the history of generative phonology.

3. Methodology

The data consists of a total of 111 Malay content words produced by two female native speakers of Malay. These were recorded, digitised and analysed using MAC Speech Lab 2. The set of words includes both simple and complex words and a few loan words from English, and words were selected to illustrate (a) the various possible occurrences of Malay CV(C) syllable structure, and (b) Malay word structure including simple words (containing from one to four syllables) and complex words encompassing prefix + root (e.g *mem+baca* 'read'), prefix + root + root (e.g. *ber+bagai+bagai* 'a variety of something'), prefix + root + suffix (e.g. *mem+baca+kan* 'read to somebody') and prefix + prefix + root + suffix (e.g. *mem+per+bagai+kan* 'to have variety'). The 111 words were segmented into 422 syllables, each syllable being annotated with its duration and a phonological representation. The data was then organised in the form of a database table.

Malay spelling is almost phonemic, and just some minor adjustments convert it into a phonological representation. The digraphs *ng* and *ny* were replaced by symbols for velar and palatal nasals respectively, e.g. *orang* ® /oraN/ 'person'. Initial vowels were preceded by a glottal stop symbol, and a glottal stop or glide symbol was inserted between adjacent vowels, e.g. *istiadat* ® /?isti?adat / 'ceremony'; *buah* ® /buwah/ 'fruit'; *sosial* ® /sosijal/ 'social'. The phonological representation was then divided into CV(C) sequences. The outcome of this procedure matched intuitive syllable divisions except in the case of loanwords containing syllables with complex onsets, e.g. /kra/ in /de.mo.kra.si/. A problem also arose regarding the syllable break following prefixes with a final velar nasal, e.g. *meng-* /məN/. The nasal patterns regularly before a consonant, e.g. *meng* + *hadkan* ® /məN.had.kan/ 'to limit', but before a vowel it can either complete the first syllable as expected, e.g. /m@N.?is.ti.me.wa/, or begin the second syllable, cf **/m@.Nis.ti.me.wa.kan/**, in which case the glottal stop is elided. This remains to be investigated in spoken corpus data.

It was already clear from an initial study of the annotated waveforms that matters of prosodic interest cluster at the end of the word, and consequently the syllables were numbered from right to left, so that the final syllable was numbered 0, the penult 1, the antepenult 2, and so on. F0 plots revealed a characteristic pattern at the end of words, in that in typical cases there was a rise to a peak on the penult followed by a complete fall to the bottom of the speaker's range on the last syllable. These were encoded "RP" and "CF" respectively. There were very few incomplete falls ("IF") and final rises ("R"), each having only two occurrences.

In most cases the peak coincided with the syllable boundary, but occasionally it was reached earlier or later, and these were marked "-" and "+" respectively. For example, "RP-" marks a case in which the peak is reached before the end of the penult, and "RP+" marks a case in which the peak is reached after the end of the penult. No other pattern has been identified. Syllables before the penult were left unmarked for pitch. Table gives

the set of records for three sample words:
TABLE 1. Records for Three Sample Words

ORTHOGRAPHY	SYLLABLE	FORM	DURATION	F0
istimewa	3	?is	111	
istimewa	2	ti	269	
istimewa	1	me	143	RP
istimewa	0	wə	412	CF
teristimewa	4	tər	158	
teristimewa	3	?is	175	
teristimewa	2	ti	237	
teristimewa	1	me	116	
teristimewa	0	wə	454	RP
keistimewaan	5	kə	158	CF
keistimewaan	4	?is	159	
keistimewaan	3	ti	253	
keistimewaan	2	me	143	
keistimewaan	1	wa	222	
keistimewaan	0	?an	348	
				RP
				CF

4. Result

The pitch patterns proved simpler than expected. There were no exceptions to the rule that the pitch rises to a peak on the penult. The pitch fell to low except in the case of the words *boikot* 'boycott' and *memboikot* 'to boycott'. (While loanwords can have special segmental properties we have no evidence so far to suggest that they have any prosodic property of their own.) In some cases the fall is incomplete, and the pitch failed to reach the speaker's bottom pitch; but there were too few cases to indicate a pattern. In 92 of the 111 words the peak coincided with the syllable boundary, and this is clearly the default case. There were too few cases of early peak to give an explanation. The remaining 10 cases of delayed peak suggested an association with (a) foreign words, and (b) final syllables containing nasals, particularly the velar nasal. In the latter case, the syllable is rich in pitch-bearing segments, so that it is physically possible to delay the pitch peak and still hear the fall.

Syllable duration proved more complex, and we begin with the kind of distribution curve we might expect. If there are no other factors involved at all, we would expect a peak corresponding to the dominant CV syllable types, although CV syllables with an initial glide or glottal stop may be shorter. CVC syllables will of course be longer. We would expect a skewed unimodal distribution with a rapid rise to a peak and a slower decay. Strict syllable timing should produce a normal curve with a small standard deviation. A mere tendency to syllable timing should produce a compromise between these two distribution types.

FIGURE 1. HISTOGRAM FOR DURATIONS OF SYLLABLES

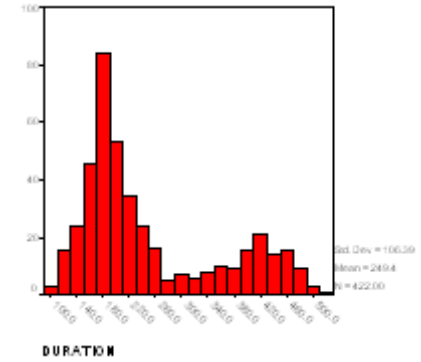
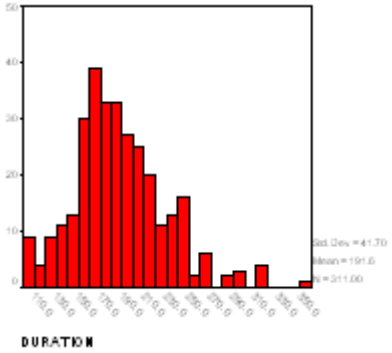


FIGURE 2. HISTOGRAM FOR DURATIONS OF SYLLABLES 1..7



The histogram in Figure 1 displays the actual distribution of syllable durations for all positions in the word. It is clearly bimodal, and looks like two separate distributions (one centred at approximately 180ms and the other centred at approximately 420ms) overlapping in the region of 280ms. This suggests that the sample is not homogeneous, and that its elements come from two different populations. The relevant factor proves to be word position, all final syllables except one having a duration greater than 280ms, and nearly all non-final syllables having a duration below 280ms, This was confirmed when the data for final syllables was filtered out, leaving the unimodal histogram displayed in Figure 2. From this initial overview it is immediately clear that final syllables are longer than non-final ones, and the shape of the curve in Figure 2 does not lead to high expectations of syllable timing.

If final position is associated with a special duration, then perhaps other syllable positions are too. Table 2 displays the mean durations (in ms) of syllables in the eight different positions. As expected, the mean of 411ms for final syllables (0) was much greater than the means for non-final syllables (1..7) ranging from 146 to 210.

TABLE 2. Mean Syllable Durations for Differene POSITIONS IN THE WORD

Syllable	N	Mean
0	111	411.41
1	110	198.15
2	93	191.05
3	62	181.37
4	31	187.35
5	11	204.82
6	2	146.00
7	2	209.50
Total	422	249.40

A one-way Analysis of Variance (ANOVA) was used to test the hypothesis that there was a relationship between syllable duration and the position of a syllable in a Malay word. ANOVA was used because more than 2 groups (i.e. 8 groups each representing 8 syllable positions) were being tested. The results of the analysis (displayed in Table 3) reveal that the means for the 8 groups are significantly different from each other ($p<0.05$; the significant level being set at 0.05). The probability of the F value ($F(7, 414)=295.82$, $p<.05$) occurring by chance is .000, thus providing reasonable evidence that the relationship between syllable position and syllable durations does exist.

TABLE 3. The First F-Test of Difference of Group Means

	Sum ofSquares	Df	Mean Square	F	Sig.
Between Groups	3971448.0	7	567349.708	295.815	.000
Within Groups	794019.56	414	2175.846		
Total	4765467.5	421			

However, since ANOVA only looks for overall difference between groups, these results do not tell us where the differences lie. In order to determine which value groups of the independent variable have the most to do with the relationship, we conducted post-hoc comparison tests that compared each mean to each other mean. Pairwise comparisons between syllable positions and syllable durations through Tukey HSD post-hoc analyses revealed that the mean differences were only significant between pairs containing a word-final syllable ($p<0.05$; the significance levels for post-hoc analyses being set at 0.05 for all post-hoc analyses). There was no significant difference between pairs which are made up of either both word-initial or word-medial positions or word-initial and word-medial positions. This confirms that the durations of the final syllables are significantly different from those of the non-final syllables.

5. Discussion

The first impression given by these measurements is that they provide overwhelming confirmation of a stress on the penult. The fact that the final syllable and only the final syllable is lengthened suggests further that the overall pattern associated with the stress includes the following long syllable. Patterns of this kind are also found in English, e.g. *widow* or *tissues*. While this explanation is consistent with the experimental data available so far, it is actually only one of a set of possible explanations. And since this particular explanation is closely related to what we already know about stress in English and other languages, there is at least a possibility that beliefs about English have been unintentionally transferred to Malay, where they may be less appropriate.

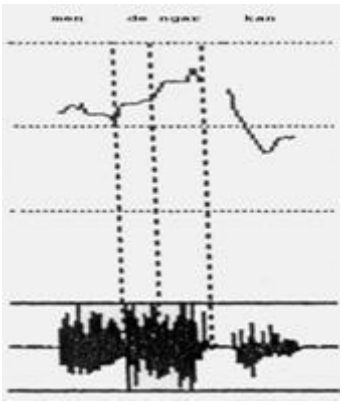
Our data contains different morphological types, but we have assumed that morphology is irrelevant. On the other hand, we have taken for granted that syllables are relevant, and that our syllabification procedure is the right one. Either or both of these assumptions could be false. Moreover, a list of 111 English words stressed on the penult would not prove that penultimate stress is the norm in English. The initial explanation must therefore be regarded as provisional until the necessary checks have been made. A suitable comparator in this connection is found in a study of stressed syllables in Arabic words by de Jong and Zawaydeh (1999). The "words" studied are not lexical items but phrases which happen to be written solid in the Arabic writing system. The phonological representation used is a hybrid of phonemic transcription and a transliteration using the Latin alphabet. Varieties of Arabic are known to vary in the alignment of F0 and in vowel reduction, and yet the data contains a mixture of colloquial and standard forms. The domain of stress is the syllable, but the statistical analysis deals only with the vowels. The conclusion that Arabic is like English in this area is rather surprising for any practical phonetician familiar with both languages, and cannot safely be drawn without further confirmation. In order to avoid comparably contentious conclusions for Malay it is essential to begin with a rigorous organisation of the data, and to put the findings in a wider context.

The measurements of individual words clearly tell us something interesting about the prosody of Malay, but exactly what they tell us is not clear. For that reason, the observed pitch pattern and the final lengthening were used as input to subsequent studies of spoken Malay, but the drawing of any conclusions was delayed pending further investigations. In the following discussion, we supplement the experimental findings with insights gained from more recent work on connected speech, involving in particular speech synthesis and the annotation of waveforms.

6. The Citation Pattern

The analysis of word duration failed to confirm the perception of some kind of prominence on the first syllable. However, a closer look at the F0 plots (see the illustration in Figure 3 using the word *mendengarkan*) revealed a pattern consistently beginning on that syllable. The pitch is set on the first syllable to somewhere in the middle of the speaker's range, and it then rises to a peak on the penult before falling to low on the long final syllable. This can be thought of as the word citation pattern.

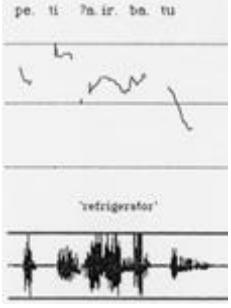
FIGURE 3. The Citation Pattern on the Word *MENDENGARKAN* 'LISTEN TO'



The citation pattern is the pattern that occurs by default, when there is no reason to do anything else, when words are cited out of any communicative context. This is why this pattern was used almost exclusively for the test words uttered in isolation.

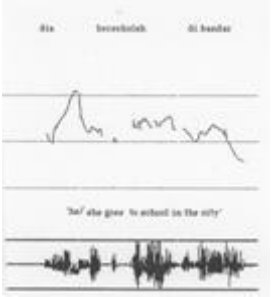
If we now examine phrases read aloud, we find that the word citation pattern is in fact made up of two elements, one initial and the other final, as shown in Figure 4. The initial pattern involves a rise in pitch to a peak at the end of the first word (henceforth the onset rise), and the pitch then remains mid until finally falling away at the end (henceforth the final fall). The important point is that the high pitch is not a property of the end of the speech interval (as might be assumed when only individual words are examined), but of the beginning.

FIGURE 4. The Citation Pattern on the Phrase *PETI AIR BATU* 'REFRIGERATOR'



A similar pattern can be seen in sentences read aloud, as illustrated in Figure 5. The middle section here is quite long, and the pitch seems to move up and down in a random fashion; in speech synthesis it has been found that this movement can be successfully emulated using a random function.

FIGURE 5. The Citation Pattern on the Sentences *DIA BERSEKOLAH DI BANDAR* 'She goes to school in the city.'



7. Timing In Malay

It is difficult to characterise syllable timing precisely (cf Laver, 1994: 156-7), but it is usually taken to mean that all syllables are subjectively equal in duration. This is clearly not true of Malay. The distribution in Figure 1 does not point to isochronous syllables, and the systematic difference between final and non-final syllables undermines the conventional belief that Malay is a syllable timed language. We could argue for a model with final lengthening imposed on underlying syllable timing. But the syllables are not even equal within the two subgroups. Final syllables vary from 514ms to 261ms, and non-final syllables from 359ms to 106ms, and a range of this kind is far too great for the members of the two sets to be regarded as equal.

Lengthening ratios vary considerably. The syllable /ca/ averages 402ms finally, and the one non-final token measures 262ms, a ratio of 1.53:1, while /gaj/ averages 477ms finally and 229ms non-finally, a ratio of 2.28:1. It might seem that /aj/ lends itself to lengthening in a way that /a/ does not, but this does not explain /da/ which averages 414 and 196, giving a ratio of 2.10:1. These ratios are much higher than found appropriate in speech synthesis where 1.2:1 or 1.3:1 gave realistic results.

The study of speech intervals (i.e. intervals of speech bounded by pauses) in texts read aloud shows that non-final words lack the final lengthening that they have when spoken in isolation. This suggests (as might be expected) that the lengthening is a property of the speech interval rather than a property of the word itself. It is the consequence of a reduction in tempo, and as in the case of a rallentando in music, the degree of slowing down is not fixed, so that the durations of final syllables varies in an unpredictable manner.

8. Words, Syllables and Stress

In designing our initial experiment to measure durations in words, we took for granted that the word is a relevant prosodic unit in Malay. We have found evidence that the first word in a speech interval has a corresponding prosodic pattern, as does the word aligned with the final cadence. We have no evidence so far that the word in between the first word and the word with the cadence have any prosodic status at all. For the purposes of speech synthesis, we treat these words as a single string of phones. In this way, the prosody associated with a typical short speech interval has a beginning, a middle and an end, and in the middle the word has at best a doubtful status.

We must also raise the question whether the syllable is a genuine prosodic unit in Malay. The reason we might want syllables at all in our description of Malay is to account for phoneme distributions. As in other languages, consonants typically overlap the beginnings and ends of vowels, and contiguous consonants typically overlap different vowels. The sequence VCCV is accordingly divided VC.CV. In this respect a word like *demokrasi* is something of an anomaly, since VCCV is here divided V.CCV and the sequence /kr/ is grouped with the following vowel. This is because when /k/ overlaps the preceding vowel it is realised as glottal constriction, and usually described as a glottal stop; but in this word it is a voiceless stop and characteristic of /k/ overlapping the

following vowel. The grouping /kra/ at first sight involves a kind of consonant overlap that is alien to the Malay language, except in so far as it has been imported along with English loan words. However, in rapid speech, it is normal for many vowels represented in the writing system (and therefore pronounced in citation forms) to be elided, and this brings together groups of consonants very similar to those which for English are regarded as syllable onset clusters. What is anomalous about the string /kr/ in the word *demokrasi* is that it has no vowel between the consonants even for the citation form.

The non-significant differences among the means indicate that there are no other factors affecting duration which relate to syllable position (counting right to left). There are other factors, but these are distributed across the positions. The most obvious factor is syllable structure, for CVC is presumably longer on average than CV. Possibly because of inequalities in syllable duration, the duration of the speech interval shows a better correlation with the number of phonemes than with the number of syllables. Durations can in fact be predicted remarkably well by distinguishing long, medium and short phonemes. The syllable has nothing to do with the onset rise, the domain of the final fall can be specified without reference to the syllable, and the syllable seems to have nothing to do with speech timing. This leaves the syllable with at best a marginal role in Malay phonology.

9. The Final Cadence

The slowing down in tempo coincides with the final fall, and together these form a final cadence. The onset rise is spread over the first word of the phrase, but in the case of a one word phrase, it is limited to the first part of the word not taken by the final cadence. The combination of the peak of the onset rise followed by the cadence gives considerable prominence to the end of the word, and since the pitch peak typically coincides with the end of the penult, the penult itself is made to stand out. This combination can give the impression of a word stress almost exactly like that of English, especially when the penult is surrounded by central vowels as in a word like *kelapa* [kəlapə] 'coconut'.

In rapid Malay speech, interconsonantal central vowels tend to be elided, and perhaps for that reason do not always take the pitch peak. In words such as *sebut* [səbut] 'mention' and *betul* [bətul] 'true', where the penult contains a central vowel, the pitch peak (and so the prominence) is delayed to the final syllable, which can give the impression that the "stress" has shifted. This impression is confirmed when the central vowel is elided in a word like *terus* [tərus] 'continue', so that [tərus] become [trus], almost exactly as English [pəli:s] 'police' can become [pli:s].

Another case that has come to light concerns the diphthong [ai] in a word such as *baik* 'good', when it is the only vowel in the word and is followed by a consonant. In this case the diphthong is fractured, and the two elements in the resulting form [bajik] are linked by a palatal glide, the pitch rising to a peak on the first element, and falling to low over the lengthened second element. This fracture, too, has its counterpart in English stressed syllables, e.g. the Texan [bæjəd] 'bad'.

The situation is more complicated in interactional data (Zuraidah Mohd Don & Knowles, 2006). Although the cadence is normally delayed until the last word, there are cases in which some final words, including vocatives and the deictic particle *ini* or *ni* 'this' when used to refer anaphorically, are ignored and treated as a kind of tail. In such cases the cadence begins on the penultimate word, e.g. *Macamana, doctor?* 'How, doctor?' and *logam ni* 'this iron ore'. These tails are similar to tails at the ends of tone units in English resulting from the deaccentuation of similar kinds of item.

The final cadence can also be highlighted with high pitch and loudness of its own, creating a wide final fall to low pitch, and this pattern seems to constitute a signal to the listener to take over the turn. The initial pattern can be masked by high pitch and loudness extending over the whole word; this pattern is used by an interrupter seeking to take over the turn. The highlighted cadence can result in a pitch peak on the last syllable, again possibly giving the impression of a "shift of stress". In the case of a word like *kenapa* [kənapə] 'why', in which the final vowel happens to sound rather like a schwa, the "stress" seems to have shifted on to the wrong syllable.

It is an interesting fact that a carefully selected sample can give the impression that

Malay has a stress system rather like English. But a similar argument can prove that English has tones like Mandarin Chinese: tone 1 for many onsets, tone 2 for elliptical questions, tone 3 for 'non-final' items, and tone 4 for primary stress. But English prosodic patterns have to be interpreted in the context of the English prosodic system, and the imposition of Chinese categories would lead to gross distortion. We would argue that exactly the same principle applies to Malay.

10. Conclusion

The first conclusion to be drawn is that there are no phenomena in spoken Malay corresponding to what phonologists call stress, and that the whole notion of stress is completely irrelevant in the description of Malay. The pitch may go up and down, loudness and tempo may increase or decrease, and on occasion the effect may be superficially similar to that produced by stress in a language like English; but these phenomena are all accounted for independently of stress.

There is also reason to doubt the need to use syllables in Malay prosody. The evidence of words like *demokrasi* is weak, and has no consequences beyond the overlapping of consonants on vowels. Syllables are important in the phonology of languages like English and Latin because they account for different patterns at different levels. At this stage we know of no high level prosodic patterns in Malay that require reference to the syllable at all. Unless and until evidence emerges to the contrary, we shall treat the syllable as irrelevant. It is difficult to believe that syllables are the cause of anything in Malay, and they are certainly not responsible for the timing of the spoken language.

It is important to note that these conclusions could not be drawn from our initial experiment. The measurements and other observations have to be interpreted, and it is significant that the simplest interpretation is one that seems to confirm the conventional view that Malay words have stress on the penult. The fact that this does not appear to be the case only emerges when we investigate in detail what happens when words are put together in continuous speech.

We therefore disagree with earlier researchers who have claimed stress on the penult. A methodology which involves the study of individual words is likely to reach this conclusion because the prosody of individual words cannot distinguish the onset rise from the final cadence. The error is therefore one of finding apparent patterns in the data which do not exist in reality. Our approach combines individual words with annotated naturally produced data. If stress and syllables really exist in the data, it is difficult to see how we could have failed to spot them.

The consequences of this finding are potentially disturbing for research in prosody. It seems a matter of common sense that if we want to find out about the prosody of a language we should start with individual words before investigating connected speech. Our evidence suggests that this may not be correct. The problem is that we cannot utter individual words with just their word level properties. When we utter a word we have to treat it as a sentence on its own, or at least as an item in a list, and this means assigning it higher level prosodic patterns. The task for the researcher is to separate out properties of the word from properties of the speech interval, and that cannot be done by examining isolated words alone.

Malay has a reputation as a "simple" language, whatever that means. It certainly does not mean it is an easy language to study, and it does not lend itself to generalisations about human language based on English and the languages of Europe. One of its more interesting properties is that it has no stress or tone or any other intermediate level of prosodic organisation. This makes Malay an ideal language to work on for the investigation of prosodic structure. It is certainly a better language for this purpose than English, in which stress and accentuation obfuscate prosodic patterns at discourse level. The prosodic relationships between words and texts are remarkably clear and simple in Malay; but as we hope we have shown in this paper, taking advantage of the clarity and recognising the simplicity can be exceedingly difficult.

The preliminary findings reported here raise a number of important theoretical issues. The lack of a stress system in Malay raises the question of how prominence should be approached. Work so far indicates that many of the prosodic patterns of Malay are

phonetically similar to those of English, the difference being that these patterns are not constrained or motivated by stress. A task for future research is to develop a theoretical framework to account for these prosodic patterns directly. Secondly, if there is no stress in Malay we have to find another explanation for the distribution of schwa and for vowel deletion, both of which seem superficially to constitute evidence for a stress system. Finally, if there is no evidence for the syllable as a relevant unit in Malay phonology, we need an alternative theory to account for the opening and closing of the vocal tract in spoken Malay corresponding to sequences of vowels and consonants. What we need is a more general theory of which the syllable theory is a specific instance.

References

Alisjahbana, S. Takdir. (1957). *Dari Perdjuangan dan Pertumbuhan Bahasa Indonesia*. Djakarta: Pustaka Rakyat.

Amran, H. (1984). *Intonasi dalam hubungannya dengan sintaksis Bahasa Indonesia*. Jakarta: Penerbit Djambatan. (PhD dissertation, University of Michigan, 1967).

Cohn, A. (1989). Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory*, 7, 167-216.

Cohn, A., & McCarthy, J. (1994). Alignment and parallelism in Indonesian. Ms, Cornell University and University of Massachussetts, Amherst.

de Jong, K., and Zawaydeh, B. A. (1999). Stress, duration and intonation in Arabic word-level prosody. *Journal of Phonetics*, 27, 3-22.

Halle, M., & Idsardi, W. J. (1994). General properties of stress and metrical structure. In J. Goldsmith (Ed.), *Handbook of phonological theory* (pp. 403-443). Oxford: Blackwell Publishers.

Kahler, H. (1956). *Grammarik der Bahasa Indonesia*. Weisbaden: Otto Harrossowitz.

Kenstowitz, M. (1994). *Phonology in generative grammar*. Cambridge, MA & Oxford: Blackwell Publishers.

Knowles, G., Williams, B. J., & Taylor, L., eds. (1996). The Lancaster/IBM Spoken English Corpus: A corpus of formal British English Speech. London: Longman.

Knowles, G., & Zuraidah Mohd Don. (*in press a*). *Word class in Malay: A corpus-based approach.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Knowles, G., & Zuraidah Mohd Don. (*in press b*). The notion of a "lemma": Headwords, roots and lexical sets. *International Journal of Corpus Linguistics*.

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

Madzhi, J. (1989). *Fonologi Dialek Melayu Kuching (Sarawak)*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Pane, A. (1950). *Mentjari Sendi Baru Tata Bahasa Indonesia*. Djakarta: Balai Pustaka.

Tassilo, A. & Butler, J. P. (1948). *Grammar of the Malay language*. New York: Roskin Photo Offset Corp.

Verguin, J. (1955). L'accentuation en Malgache-merina et en Malais. *Orbis*, 4/2, 522-528.

Yunus, Maris. 1980. (*The Malay sound system)*. Kuala Lumpur: Fajar Bakti.

Zuraidah Mohd Don. (1996). *Prosody in Malay discourse: An analysis of broadcast interview*. Unpublished PhD thesis, University of Malaya.

Zuraidah Mohd Don, & Knowles, G. (2006). Prosody and turn-taking in Malay broadcast interviews. *Journal of Pragmatics,* 38,490-512.