

Syllables, moras and accents in Beja

R. A. HUDSON

Department of Phonetics & Linguistics, University College, London

(Received 12 January 1972)

In his study of Japanese phonology, McCawley introduces an interesting typological classification of languages according to the kinds of rules needed for describing the distribution of prosodic features like stress and pitch (1968: 58-61). Whereas Trubetzkoy had made a simple distinction between 'syllable-counting' languages and 'mora-counting' languages, McCawley suggests that we should make two independent distinctions:

- (a) according to the 'unit of phonological distance' (the unit in terms of which the location of any accent is calculated) between 'syllable-counting' and 'mora-counting' languages;
- (b) according to the 'prosodic unit' (the unit that actually bears the accent) between 'syllable languages' and 'mora languages'.

For instance, it is possible for a language to be a 'mora language' in that 'there are two possible accentuations for a long syllable but only one possible accentuation for short syllables', so that, say, a long syllable might have either a falling or a level accent, whereas a short syllable could have only the latter; and the same language could also be a 'syllable-counting' language, in that the position of the accent could be calculated in terms of syllables, without any distinction between long and short syllables.

The Beja language, a Cushitic language of the north-eastern area of the Sudan, seems to fit into all FOUR categories, at least if we take a fairly surface view of its phonology: in some respects it is 'syllable-counting', and in others 'mora-counting', and in some respects it is a 'syllable language' while in others it is a 'mora language'. If this is so, McCawley's typological framework will need to be revised, especially given that he explicitly claims that it reflects universal constraints:

Implicit in this distinction [between syllable-counting and mora-counting languages] is the belief that within any language, all rules of forms such as 'put stress (high pitch, etc.) n units before (after) X' will refer to the same unit: either all distances in a language will be measured in syllables or all distances will be measured in moras.

However, I believe that the framework can be saved by reformulating it in terms of 'normal' prosodic patterns, rather than ALL prosodic patterns, and by

applying it separately to the underlying prosodic patterns and to the surface ones.

1. BEJA AS A MORA LANGUAGE

Beja is not a tone-language as this term is normally understood, but has a pitch-accent system in which syllables are either accented or not accented; the position of the accent is unpredictable, as witness:

[bárra:d] 'tea-pot'
[fɪndʒá:n] 'coffee-cup'

If a vowel is phonologically long, there are two possible manifestations of the accent: either a high-level pitch (represented as $\acute{}$), or a high-falling pitch ($\grave{}$):

[kà:m] 'camel'
[sú:g] 'market'

So far as these patterns are concerned, Beja appears to have a 'non-linear system with zero' (Hockett, 1955: 69-71), ranging it alongside Norwegian. In McCawley's terms, it counts as a mora language because long syllables (i.e. in this case, syllables containing a long vowel) have the two types of accent, level and falling, whilst short syllables do not. If we treat long syllables as composed of two moras, we can reanalyse the distinction between the two accents, in terms of position on the first or second mora of the syllable.

Given this reanalysis, Beja ceases to have two distinctive accents as such, and has simply one accent, realized by high pitch, occurring on either mora of a long syllable. Moreover, it also ceases to be a 'non-linear system with zero', since the non-linear relations between 'no accent', 'high-level accent' and 'high-falling accent' have been replaced by a linear relation between 'no accent' and 'accent'. However, we can still count Beja as a mora language, in McCawley's sense, since there are two different accentual possibilities (high accent on first or on second mora) on long syllables, compared with the one possibility on short syllables.

2. BEJA AS A MORA-COUNTING LANGUAGE

The position of an accent is determined by counting back either one 'place' or no 'places' to the left of a fixed point, which thus may or may not itself bear the accent. There can be up to six such fixed points in a word, each with its own overt accent, but I will consider here only those cases where the fixed point is the last mora of a noun word. Whether the accent is located on this fixed point or one place before it is determined by a complicated interplay of grammatical and lexical factors, but in many nouns the distinction between singular and plural is made by shifting the accent back one place:

[dɛ:t]	'mother'	[dɛ:t]	'mothers'
[kwá:t]	'sister'	[kwà:t]	'sisters'

In view of examples such as these, it seems that the rule for locating the accent must be formulated in terms of moras rather than syllables: the accent is on the last mora of the stem in the singular, and on the penultimate mora in the plural. (In other words, count back no moras in the singular, and one mora in the plural.)

This mora-counting rule also covers cases where the accent does in fact shift from one syllable to another: in many nouns the last vowel shortens (i.e. loses a mora) in the plural, which means that the mora before the last mora in the stem is in the penultimate syllable:

[iindzú:n]	'coffee-cup'	[fíndʒan]	'coffee-cups'
[bikká:r]	'hut'	[bíkkar]	'huts'

3. BEJA AS A SYLLABLE-COUNTING LANGUAGE

However, there are other accent patterns in which the counting has to be in syllables rather than moras, since a short syllable counts for as much as a long one. For instance, among the nouns which shorten the final vowel in the plural there are also some which have the reverse movement of the accent between singular and plural: the accent is on the penultimate syllable in the singular, and on the final syllable in the plural:

[kíta:b]	'book'	[kítáb]	'books'
[bárra:d]	'tea-pot'	[barrád]	'tea-pots'

If we are to count this movement in terms of moras, we shall have a change of TWO moras for these nouns, paralleling the move (in the reverse direction) of only one mora for the earlier set of nouns. Moreover, there are NO instances (so far as I am aware) of two-mora changes involving two moras in separate syllables: in other words, this would be simply a rather complicated way of saying 'move the accent to the preceding syllable'.

There can be little doubt, incidentally, that it is correct to locate accents by counting back from a fixed point – in this case, from the last syllable in the stem – rather than by labelling the accentable syllables directly in the lexicon. One conclusive piece of evidence is that syllables can come and go, as the result of various morphophonological rules inserting epenthetic vowels, and that in these cases the only constant thing about the accented syllable is its position relative to the end of the stem, rather than its segmental make-up. For instance, as we have already seen, 'book' (or 'a book') is [kíta:b], whose accent must be on the syllable before the stem's last syllable. However, when the definite article [o:] is added, the [ɪ] drops, since it was needed simply to make the syllable structure fit the normal (C)V(C) pattern of Beja, which is now ensured by the

presence of [o:]. This makes [o:k] the syllable before the last syllable of the stem, and by virtue of this position [o:k] bears the accent: [ó:kta:b], 'the book'. By way of contrast, in [dɪwá:n], '(a) water-pot', the accent is located on the stem-final syllable itself, so the definite article does not have an accent: [o:dwá:n], 'the water-pot'.

4. SYLLABLE-COUNTING AND MORA-COUNTING

We have seen that there are some forms which need a mora-counting rule for accent-location, and others which need a syllable-counting one. The parallel between the two kinds of relation can be seen in pairs of singular/plural contrasts such as the following:

	<i>mora-counting</i>		<i>syllable-counting</i>	
	'camel'	'mother'	'book'	'bride'
singular	[kà:m]	[dé:t]	[kɪta:b]	[do:bá:t]
plural	[kám]	[dè:t]	[kɪtáb]	[dó:ba:t]

For nouns like 'camel' and 'book', high accent is on the penultimate mora and syllable respectively in the singular, but on the final mora/syllable in the plural; and for those like 'mother' and 'bride' the pattern is simply the reverse of this.

From examples such as these the reader would be right to conclude that the difference between mora- and syllable-counting reflects the number of syllables in the stem: in polysyllabic stems the counting is in syllables, but in monosyllabic stems it is in moras. In other words, the rule is: count in syllables if you can, but in moras if you can't.

However, there is more to it than that: it is not simply a question of the number of syllables in the STEM. We have already seen that this is not so in examples like [ó:kta:b], 'the book', where the accent is located one syllable (not one mora) back from the end of the stem, even though the stem itself only consists of one syllable (and a bit of the preceding syllable): [(k)ta:b]; in cases like this the article counts in with the stem's syllables.

Nor can we use the WORD as the relevant unit, in place of the stem: there can be a number of accents in the same word, and some can be of the mora-counting kind, as [bà:ká:ri:fhè:b], 'without meeting me' ([bà:], 'not, adverbial'; [ká:ri:f], 'meet'; [hè:b] 'me'). The only way in which we could account for the falling accent on [hè:b] would be by treating it as a separate word, so that the mora-counting rule could take over; but there are good grammatical reasons for not doing this, but keeping it as a suffix to the verb.

For these reasons, among others, it is necessary to postulate a larger phonological unit above the syllable, which I shall call the 'accent-unit'. It is the number of syllables in the ACCENT-UNIT that decides whether counting shall be in syllables or in moras: if there is a syllable between the relevant fixed point and the

preceding accent-unit boundary, then counting is done in syllables, but otherwise it is done in moras. Thus there must be an accent-unit boundary, for instance, immediately to the left of [hè:b] 'me', to prevent its accent from being realized in the previous syllable.

Given that we have accent-units, then, the rule for placing the pitch accent is: count back zero or one syllable from the fixed point if there is a syllable before it in the same accent-unit, but if not, count back zero or one mora. Whether one counts back zero or one place depends on the lexical and grammatical circumstances, but the advantage of this formulation of the rule is that it simplifies the rules relating lexico-grammatical patterns to accentual ones, as I shall now explain.

As we represented the four examples of singular:plural pairs of nouns given above, there were three accentual patterns: high-level pitch on the last syllable ([kám], [dé:t], [kítáb], [do:bá:t]), high-level pitch on the penultimate syllable (kíta:b], [dó:ba:t]), and high-falling pitch on the final syllable ([ká:m], [dè:t]). However, we can now subsume the second and third pattern under a single more abstract accent pattern, whose phonetic manifestation depends on the number of syllables in the accent-unit. We can call this accent-pattern 'penultimate accent', since it is manifested as a high pitch either on the penultimate syllable, or on the penultimate mora, before the fixed point. The other accent can be called 'final accent'.

If we postulate these two kinds of accent, where do we locate the penultimate accent itself? We can locate it either on the mora in which it is manifested, or on the one which counts as the 'fixed point' from which the position of the high pitch is calculated. It is clear that there is everything to be gained from taking the deeper of the two approaches for locating the accent on the mora from which the position of the phonetic pitch is calculated. For one thing, the two accents will always occur in the same positions in the underlying representation, so that where one alternates with another (as in the singular/plural contrast) we can specify where the accent occurs independently of saying WHICH accent it is. Secondly, the penultimate accent will fall on the same position in the word irrespective of whether it is manifested on the previous mora or syllable.

The transcription which has been used so far has been a relatively surface one – hence the square brackets, suggesting a broad phonetic transcription. It does not reflect the new analysis which I have been arguing towards, in which long vowels are considered to comprise two moras and there are two accents (apart from the lack of an accent), located on the fixed point.

Instead of representing the long vowels as [a:], [i:], etc. we can represent them as /aV/, /iV/, and so on, where /V/ stands for the half of a vowel which occurs in the second mora of a syllable, but takes its quality entirely from that of the preceding vowel. This notation has obvious advantages over the earlier notation, not least of which is that it can be taken as a transcriptional represent-

ation of a distinctive feature matrix, with two separate columns corresponding to each long vowel. Another advantage is that it simplifies the morphophonemics: most consonantal suffixes have the effect of lengthening a stem-final vowel, as in [dó:ba:t] 'brides' (absolute form) as opposed to [dó:ba] 'brides' (form used after articles or adjectives), and this can easily be stated by including /V/ in the suffix's realization - /Vt/, for instance. In other words, there is a morpheme boundary in the middle of the long vowel, so we treat the latter as bimoric, as Trubetzkoy suggested (1969: 173).

As far as the accents are concerned, we can represent them as ' (final) and ˇ (penultimate), located in both cases on the 'fixed point', which, in all the eight nouns we contrasted above, was the final mora of the word. We can now represent these words as follows:

	<i>mora-counting</i>		<i>syllable-counting</i>	
	'camel'	'mother'	'book'	'bride'
singular	/kaŃm/	/deŃt/	/kitaŃb/	/do:baŃt/
plural	/kám/	/deŃt/	/kitáb/	/do:baŃt/

Together with some phonemicization of the vowel segments, the two changes we have introduced have the effect of simplifying the morphophonemics considerably. There are now just two classes of noun stems differing in the relation between singular and plural:

- (a) those in which the singular has /ˇ/ and the plural /'/, and the plural only has one mora, corresponding to the singular's two moras, in the last syllable;
- (b) those where /'/' is in the singular form and /ˇ/ in the plural, but there is no shortening.

In all nouns the accent is on the last mora of the stem.

What effect does this reanalysis of the accent system have on the way in which we classify Beja according to McCawley's typological framework? First of all, let us consider the question of how the location of the accent is calculated. If we take this as a question about the UNDERLYING accents, Beja is NEITHER a syllable-counting NOR a mora-counting language, since the accent is located directly on the reference-point - say, the end of the stem - so that no counting at all is needed. If, however, we take the question to refer to the SURFACE accents, Beja is BOTH syllable- AND mora-counting.

Let us take the question in the latter sense: what now of McCawley's claim that languages always count consistently in terms of either syllables, but never mix the two units? Strictly speaking the facts of Beja do seem to contradict this claim. On the other hand, it is not the case that both kinds of counting are on an equal footing in Beja: syllable-counting is the norm. This is clear from the way in which the rule for locating surface accents is formulated: syllables are to be

counted except where moras are counted (and not vice versa). This being so, the distinction between syllable-counting and mora-counting languages can be saved by means of a reformulation in terms of UNMARKED accent-patterns; and similarly the claim about languages never using both syllables and moras could be corrected (and, unfortunately, weakened) to read: no language uses both moras and syllables as the NORMAL unit of prosodic distance.

To summarize, it seems that Beja has the following characteristics as far as the distinction between syllable-counting and mora-counting is concerned:

- (a) in terms of surface patterns, it is a SYLLABLE-COUNTING language, in the sense that the syllable is the unmarked unit for calculating the position of surface accents;
- (b) in terms of underlying patterns, it is neither syllable-counting nor mora-counting, since accents are located directly on particular places in word structure.

5. BEJA AS A SYLLABLE LANGUAGE

The other change which our reanalysis effects in the typological status of Beja is that we have now reverted to the earlier analysis in which there are two contrastive accents, and not just one accent (high) which can occur on either mora of a long syllable. However, whereas in our first analysis only long vowels were subject to the contrast between the two accents, now both long AND short vowels allow both accents, since they are now distinguished primarily by the LOCATION of the surface-accent, rather than by the contrast between level and falling pitch. Presumably we must conclude from this that Beja is not after all a mora language, but is a syllable language, since there is no accentual difference in the underlying representation between long and short syllables.

Thus, as far as the second distinction, between mora languages and syllable languages, is concerned:

- (a) in terms of surface patterns it is a MORA language, since there are two possible accents (high falling and high level) for long syllables, but only one (high level) for short syllables;
- (b) in terms of underlying patterns it is a SYLLABLE language, since there are no differences between long and short syllables: both can bear either of the two accents, called 'final' and 'penultimate'. However, paradoxically, the distinction between long and short syllables is made in terms of moras, the long syllable containing two moras to the short syllable's one.

This paradox may appear to be simply an artifact of the notational device introduced above, whereby all long syllables are shown in the deep representation

as containing some vowel followed by V, representing a second mora. Would it not have been better to keep to the non-segmental length marks (/a:/ or /ā/, say) in the underlying representation, and keep the mora notation for the surface representation?

The main reason for preferring the analysis with the paradox of moras in a syllable language is that not all long monosyllables behave like /kaṽm/ 'camel'. There are some which have the penultimate accent, just as /kaṽm/ does, but which do not have a falling surface accent where /kaṽm/ would have one. For instance, [ó:su:g] 'the market', has the accent on the article, in contrast with [o:fá:r], 'the flower', so we can identify the penultimate accent on the former, and the final accent on the latter. This being so we should expect a falling accent on the word for 'market', when there is no article, to match [kà:m], '(a) camel'. But the accent is in fact the level accent: [sú:g] '(a) market', just like [fá:r], '(a) flower'.

Not all long syllables are to be treated the same, then, as far as accents are concerned. Some are like [kà:m], 'camel' and [-hè:b], 'me', and others are like [sú:g], 'market'. Indeed, the former are much rarer than the latter among morphemes. The following list may be exhaustive for monomorphemic morphemes with long vowels like [kà:m], which can have a falling accent:

- (a) a few monosyllabic noun-roots ending in a vowel (to which /-ṽt/ can be suffixed): /de-ṽt/ 'mother(s)', /kwa-ṽt/ 'sister(s)', /a-ṽt/ 'milk';
- (b) the comparative suffix /-ka/, which always has an accent-unit boundary immediately before it, and can be followed by concordial suffixes like /-ṽt/, as in /fagar+ka-ṽt/, 'more diligent (fem. plur.)';
- (c) the noun-root /kaṽm/, 'camel' and the numeral root /gaṽl/, 'one';
- (d) the object pronoun suffix /heṽb/, 'me';
- (e) various negative prefixes found in certain forms of the verb: /kaṽ-/ 'I don't . . .', /baṽ-/ 'without'.

If these few morphemes were removed from Beja, practically all the complexity of the accent system would also disappear, since there would be no need for moras and Beja would be a pure syllable language.¹ Be that as it may,

[1] There are in fact a set of words in which falling accents do occur on the surface, but where these do not result strictly from the lack of a syllable for the penultimate accent to be realized on. In all cases there is a penultimate accent on a vowel which, in the underlying representation, follows /ʔ/ or /h/ preceded by a short vowel. It ought to be this short vowel which bears the surface accent, but there is a rule which automatically prevents short vowels from having a surface manifestation before /ʔ/ or /h/, so the accent is thrown back again on the syllable following /ʔ/ or /h/, where it results in a falling surface accent. For instance, [kɪlʔò:b] is the surface form of /kɪlaʔoṽb/, 'shells' paralleling [ragáda:b], from /ragadaṽb/, 'legs'. To take account only of these words, we should have to say that Beja was a pure syllable language in its deeper phonology, although it was a mora language on the surface.

these few morphemes do exist, and have to be distinguished phonologically from those which do not permit a falling accent.

One way to make the distinction would be to restrict the two-mora analysis to those cases where a falling accent is possible, and to mark other long vowels in the usual way, as long single moras: /ka^ˇm/, 'camel', versus /sú:g/, 'market'. Given this analysis, we should have to formulate the vowel-shortening rule, which shortens final vowels in some plural nouns, in such a way that it removes either the length marker or the V representing the second mora. This is one disadvantage of the analysis, albeit a small one, and another is that we would probably have to say that the vowel-lengthening associated with a consonantal suffix involved a length-mark in some cases, like /do:ba-ʔt/, 'brides', and a second mora in others, like /de-^ˇt/, 'mothers'.

The analysis which I prefer would treat all long syllables as bimoric, containing some vowel followed by V, and would make the distinction between words like [kà:m], 'camel', and [sú:g], 'market', in terms of the location of the deep accent on the first or the second mora: /ka^ˇm/ versus /sùVg/. This would rule out a falling accent on /sùVg/, there being no mora before the accent-bearing one. Thus whereas the high pitch in /ka^ˇm/ can be mapped onto the mora before the accent-bearing one, allowing a falling pitch, this is not possible for /sùVg/, since there is no such mora.

If we adopt this analysis, we shall need to complicate the rules for mapping deep accents onto surface pitch-accents, and these will now be as follows:

- (a) A (deep) 'final' accent is mapped onto a surface accent on any mora in the surface syllable (corresponding to the deep syllable) in which the accent-bearing mora occurs.
- (b) A deep 'penultimate' accent is mapped onto any mora which is in the same accent-unit as the accent-bearing mora, and which:
 - (i) is in the syllable before the accent-bearing mora (/kíta^ˇb/, 'a book', /oVka^ˇm/, 'the camel'); otherwise,
 - (ii) is before the accent-bearing mora (/ka^ˇm/, 'a camel', /de^ˇt/, 'mothers'); otherwise
 - (iii) is in the syllable containing the accent-bearing mora (/sùVg/, 'a market', /wín/, 'big').

If there is a surface accent on the first mora of a syllable, but not on the second, this represents a high-falling pitch; otherwise the surface accent always represents a high-level accent. Thus in monosyllables a 'penultimate' accent will be indistinguishable, phonetically, from a 'final' accent if the underlying accent is not on the second mora, either because it is on the first of two moras, as in /sùVg/ ([sú:g]~[ó:su:g]), or because it is on the only mora, as in /wín/, 'big' ([wín]~[ó-win]).

I have tried to show why it is necessary to treat long syllables in the underlying

representation as comprising two moras, in spite of the paradoxical fact that both the deep accents can occur on both long and short underlying syllables. However, in the light of the analysis introduced in the last two paragraphs, this paradox disappears again: although both accents occur on short as well as on long syllables in the underlying structure, there are two places where they can be located within a long syllable (on the first or on the second mora) whereas there is only one in a short syllable. So once again there is a distinction between long and short syllables, as far as their accentual possibilities are concerned, and Beja can be regarded as a true mora language, even in its underlying phonology.

6. CONCLUSIONS

In order to fit Beja into McCawley's typological framework, we need to make two sets of distinctions: between 'deep' and 'surface' phonology, and between 'unmarked' and 'marked' patterns.

(a) *Surface phonology*

- (1) Beja is a SYLLABLE-COUNTING LANGUAGE, in that the syllable is the unmarked 'unit of prosodic distance', in terms of which one counts back from the deep accent-bearing mora to the mora with a high surface accent.
- (2) Beja is a MORA LANGUAGE, in that there is a contrast between high-level and high-falling pitch-accent on long syllables, but not on short ones.

(b) *Deep phonology*

- (1) Beja seems to be neither syllable- nor mora-counting, since accents are located directly by the morphophonological rules; however, they are located on moras rather than on syllables, so we might reinterpret McCawley's distinction and treat Beja as a deeply MORA-COUNTING LANGUAGE.
- (2) Beja is a MORA LANGUAGE, since there are more accentual possibilities for the long syllable than for the short one, and, moreover, long syllables are analyzed (for other reasons) as comprising two moras.

It looks as though we have resolved the conflict, at least in the sense that we no longer need to say that Beja is a mora language in some respects and a syllable language in others. However, it is important to realize how different the deep and surface patterns are, so that the apparent harmony between them should not be taken for granted. On the surface we have THREE possible accented syllable types: short, long level, long falling; in the deep analysis there are SIX: short, long (first mora accented), long (second mora accented), all bearing either the 'final'

SYLLABLES, MORAS AND ACCENTS IN BEJA

or the 'penultimate' accent. On the surface, the accent and the fixed point from which its position is calculated may or may not coincide; in the deep analysis, they always do. These discrepancies can be seen in the following examples:

	'market'	'the market'	'camel'	'the camel'
<i>deep</i>	/sũVg/	/o:sũVg/	/kaŧm/	/o:kaŧm/
<i>surface</i>	[sú:g]	[ó:su:g]	[kà:m]	[ó:ka:m]

The general implications for phonological typology are clear, though scarcely surprising. Just as in syntactic typology we have to distinguish surface patterns from deep ones, so that it is conceivable that we might treat English as a 'deep' V-S-O language (McCawley, 1970), although it is a 'surface' S-V-O language; so in phonological typology we have to distinguish surface and deep patterns. To hazard a rash guess, I should expect less discrepancy between surface and deep patterns in phonology, and it may be significant in this respect that Beja does turn out to be a *mora* language in its underlying phonology as well as in its surface patterns, in spite of the complex set of rules mapping one onto the other.

REFERENCES

- Hockett, C. F. (1955). *A manual of phonology*. Baltimore: Indiana University.
 McCawley, J. D. (1968). *The phonological component of a grammar of Japanese*. The Hague: Mouton.
 McCawley, J. D. (1970). English as a VSO language. *Lg* 46. 286-99.
 Trubetzkoy, N. S. (1969). *Principles of phonology*. Tr. Baltaxe, C. A. M. Berkeley: University of California Press.