The Rise of Rhythmic Structure in Bantu*

John Goldsmith

1. INTRODUCTION

My aim in this paper is to present a hypothesis concerning the development of the tonal system found today in the closely related Bantu languages, Kirundi and Kinyarwanda. These two languages are part of a larger group, the so-called Lacustrine languages, whose general tonal behavior is reasonably well understood. These two languages, however, present certain distinctive tonal characteristics that can be seen, once their tonal systems are studied in some depth. My hypothesis is that the basic cause of this change from the other Lacustrine languages was the incursion into these languages of a rhythmic pattern of alternating strong and weak syllable positions that encouraged tonal rules that were quite different from anything that had been seen before in the area. In particular, several tonal rules arose that shifted High tones one mora to the left or right in order to associate a High tone with a metrically strong position, shifting the tone from the position where morphological (as well as historical) considerations would lead us to expect the High tone.

This example is not the first one in which the origin of metrical structure has been explored from a diachronic perspective. Comparative work on the Northern Iroquoian languages by Wallace Chafe (1977) — using work by Chafe, Lounsberry, Woodbury, and others — has led to several recent theoretically oriented papers by phonologists working within the tradition of metrical phonology, including Stowell (1979) and especially Prince (1983). Chafe himself succeeds in tracing the shift of accent placement from a simple penultimate-accent rule in Proto-Northern Iroquoian to a more complex system in Onondaga, Cayuga, and Seneca. Others, following in his footsteps, have attempted to integrate these insights into current and rapidly developing generative theories. In section 2 of this paper, I will review the highlights of this work.

My primary goal, however, as I have noted, is to sketch a view of a similar development in one subgroup of Bantu languages, the group containing Kirundi and Kinyarwanda (and perhaps Gĩha, for which I have no relevant data), but excluding the other closely related Interlacustrine languages, such as KiHunde, Haya, Luganda, and Shi. I will call this small subgroup that holds our interest the “Middle Lake languages”. This sketch will be presented in sections 3–6. Unlike the situation in Northern Iroquoian, where the ancestral situation was the extremely simple pattern of penultimate accent, the prosodic situation out of which the Middle Lake languages developed a rhythmic system was an extremely complex tonal system, one which is, by good luck, one of the better studied families of tone systems.

* I am grateful to Fimard Sabimana for many enjoyable conversations on these subjects, without which this paper would not have been possible.
2. THE RISE OF RHYTHMIC STRUCTURE IN NORTHERN IROQUOIAN

The following description is based on the principles given in Chafe (1977), and to a somewhat lesser degree Foster (1982); much of the perspective of the analysis comes from Prince's (1983) diachronic scenario, though I have not utilized his proposal concerning the metrical grid here.

In Proto-Northern Iroquoian, accent was regularly placed on the penultimate syllable. This accent had tonal characteristics in some measure; the Strong rhythmic position was thus the spot where the High tone of the word was anchored. Let us schematize this as in (1), where a word-final binary Strong-Weak foot is placed over the end of the word. Regarding syllable structure, we may note that there were no underlying long vowels; a consonant was forced into the coda of a syllable on the left, closing the syllable, just in case it was followed by another consonant (while the latter consonant, of course, became the onset of the following syllable), or if it was a laryngeal glide (ʔ or h).

(1) oh ráʔ taʔ 'feather'

```
S \   / W
  \   / H
   \ / 
```

Both the High tone and the rhythmic structure are placed on the frame below the segments, but they should be thought of as being in separate frames extending out radially from the central core or skeleton.

At some point in the development of the Northern Iroquoian languages — possibly, in the light of these prosodic developments, at a point where Onondaga, Cayuga, and Seneca shared a unique common ancestor — a new kind of rhythmic structure arose in the language, a structure based not on binary right-to-left foot-construction from the end of the word, but rather of binary Weak-Strong foot-construction from left to right starting with the beginning of the word. This may be represented schematically as in (2), with the new left to right structure on what I will call the upper frame.

(2) W \ S \ S \ S \ S \ S \ S \ S \ S \ S
    \ C V \ C V \ C V \ C V \ C V
    \ S \ W \ S \ W \ S \ W \ S \ W
    \ H \ H
```

```
```
When the word had an even number of syllables, both the penult and the antepenult were assigned a Strong position, though on separate frames (lower and upper, respectively, in (2)). This rhythmic conflict led to shifts in tone placement, length, and stress. The single most important shift is the one in (3), the Iroquoian Strong Pullback Rule (essentially Prince's rule 108, 'Backward H Spread').

(3) Iroquoian Strong Pullback Rule

There are some variations in the details of this rule as it is found in Onondaga, Seneca, and Cayuga. On the one hand, in both Seneca and Cayuga, the original association of the High tone is lost when the new leftward association is made, while in Onondaga, on the other, both syllables remain High; this was probably the original situation. Furthermore, there is a condition in Onondaga that the syllable on the left must be light and the syllable on the right must have a long nucleus (i.e., long vowel or vowel + laryngeal glide).

In Cayuga, main stress corresponds to the placement of High tone after the Strong Pullback rule. From the point of view of the rhythmic structure in the upper frame, erected from left to right, it always falls on a Strong Position; this is the penult in a word of an odd number of syllables and the antepenult of a word of an even number of syllables.

Other strong positions in the upper chart — other even-numbered syllables — present to the listener evidence of prominence, or secondary stress, and the weak positions undergo what has been called, following Lounsbury, laryngeal metathesis, with the sequence vowel + laryngeal glide — effectively the sole falling diphthong in the language — metathesizing to a glide + vowel sequence, or rising diphthong.

Seneca has innovated further by modifying the first foot formation rule on the lower frame, constructing a maximal unbounded foot from the right-hand end of the word where the first syllable of the foot contains a closed syllable — that is to say, the Strong position is the last closed syllable. But just as in Cayuga, when this position is in a metrically Weak position — which is to say, when it is in an odd-numbered syllable — then the High tone, or main stress, shifts back a syllable to a metrically Strong position.
Proto-Northern Iroquoian is interesting in that the origin of rhythmic metrical structure out of nothing is so clear here. In the case of the Middle Lake languages, the situation is different in that metrical structure arose out of an already extremely complex prosodic system, the complex tone system of the Bantu languages.

3. THE MIDDLE LAKE LANGUAGES: KINYARWANDA NOMINALS

The Middle Lake languages — primarily Kirundi and Kinyarwanda — have been the object of a great deal of study, both tonal and non-tonal; see, for example, Meeussen (1959), Coupez (1980), Kimenyi (1979), and Sibomana (1974). For the purposes of this paper, I will draw primarily on Goldsmith and Sabimana’s (1984) analysis of Kirundi, and Furere and Rialland’s (1984) analysis of Kinyarwanda.

Both of these treatments independently focus on the analysis of the verb into special units comprised of exactly an even number of moras. In the light of everything we know about the surrounding, related languages — the so-called Lacustrine group — this is an unexpected result. Let us reflect on how binary feet might arise in a language, and what would count as evidence for it.

There are two ways in which we can find direct evidence for binary metrical structure. On the one hand, we can find pairs of adjacent moras grouping together in some apparent way; on the other hand, we can find that the conditioning factor for some phonological rules (like the strong pullback rule of Northern Iroquoian) depends on counting an odd or even number of moras to the left or right of some phonological landmark in the word. The first type of evidence we might call local, the second distant.

In actual practice, the first kind of evidence, local evidence, is probably less significant. Makua, a Bantu language of Mozambique and southern Tanzania investigated by Cheng and Kisselberth, displays rightward tone doubling which should probably be analyzed in terms of local binary feet. No non-local or distant metrical effects have been identified in Makua, however, suggesting that local binary effects may not be a significant push towards a word-level rhythmic reanalysis.

The Middle Lake languages with which we are concerned, however, display both kinds of evidence of rhythmic structure. On the one hand, we find in Kinyarwanda (in almost all forms, verbal and nominal) that High tones do not appear alone on vowels, but rather always appear in groups or plateaus of two, four, or six High-toned vowels. This is a local metrical effect. Examples of this effect can be seen in (4), where the rightmost High tone is the etymologically original High tone. The principle at work in present-day Kinyarwanda, according to Furere and Rialland, is that the basic High tone extends leftward two pairs of vowels at a time; when this plateau of two or four vowels
extends as far as the edge of the nominal root, or beyond, then the plateau ends. What is crucial for our purposes is that the vowels are organized into pairs of vowels — binary feet, in the terms we used above — and these feet are used to assign tone. The anchor-point (to use Furere and Rialland’s term) is the first, and only, High tone in the word.

(4) Kinyarwanda nouns: all of the form \textit{Prefix}+\textit{Stem}

\begin{itemize}
\item a. \textit{CV}+\textit{CV} \quad \text{CV} \quad \text{H}
\item b. \textit{CV}+\textit{CV} \quad \text{CV} \quad \text{H}
\item c. \textit{CV}+\textit{CV} \quad \text{V} \quad \text{CV} \quad \text{H}
\end{itemize}

In the nominal system, in the overwhelming majority of tonally reconstructed noun stems, there is at most one High tone. This High tone acts, in Furere and Rialland’s words, as an “anchor-point”; from this point, binary feet must be constructed leftward (they may be constructed rightward, too; it does not matter, for they will not participate in any rules or realizations). The High tone is underlyingly associated with one particular vowel in the stem; it spreads leftward, however, to all vowels in feet contained (or partially contained) within the stem. We may perspicuously illustrate this by saying that these feet are joined together to form a constituent \textit{F'} (a “superfoot”, we might say), as in (5), and then the High tone spreads leftward across it, though this particular formalism is not crucial to the idea presented here.\footnote{Van der Hulst 1984 (chapter 4) contains a current discussion of the issues involved in different formal representations of rhythm, and the claims concerning phonological constituency in the different approaches.}

(5)

\begin{itemize}
\item \textit{F'}
\item \textit{F}
\item \text{V} \quad \textit{CV} \quad + \quad \text{CV} \quad \textit{CV} \quad \text{CV} \quad \text{CV} \quad \text{CV}
\item \text{H}
\end{itemize}

This \textit{particular} realization is undoubtedly a secondary one, that is, not to be reconstructed for Proto-Middle Lake; according to this pattern, for
example, a High tone on the first vowel of a stem is doubled back onto the last vowel of the prefix as well. This is true in Kinyarwanda, but false in Kirundi; it is an effect, as we have said, secondary to Kinyarwanda. But the construction of binary feet, starting from the only prosodically prominent vowel of the word, is an effect that should be included, I believe, in our reconstruction.

Let us step back for a moment and recall that in languages with clear alternating stress patterns, two general types of such stress patterns can be discerned (and, indeed, these two types can be found in the same language). The alternating stress pattern can take as its starting point ("anchor point") either some vowel with inherent prosodic prominence (typically syllable weight), or the pattern can begin with the left (or right) end of the word. If the former option is used, the rhythmic pattern frequently streams outward to both the left and right. (This latter case is somewhat different from the case of a left-to-right scanning, quantity-sensitive language where heavy syllables may not fall on a weak position, and where foot construction begins anew whenever a strong syllable is encountered.)

This is the principle, I would like to suggest, that arose in Proto-Middle Lake: binary metrical structure was erected both to the left and to the right of a prosodically prominent vowel. In nominals, this was the underlying High tone (in stems with High tones); this pattern then led to the leftward spreading particular to Kinyarwanda just described.

This is illustrated in (6). It is of great interest, in the light of the observations of Liberman, Prince, Hayes, and Halle that a grid and a foot system lead to quite different results if tone is to be considered spreading over the foot. A Perfect Grid, in Prince’s sense, can be spread easily in (6), but that leaves open whether it is through Strong-Weak or through Weak-Strong feet, as in (7a) and (7b), respectively.

(6) Grid-like representation

\[
\begin{array}{cccccc}
\times & \times & \times \\
CV & CV & CV & CV & CV & CV \\
\end{array}
\]

H

(7) Foot-tree representation

\[
\begin{array}{ccccccccc}
S & W & S & W & S & W & W & S & W \\
CV & CV & CV & CV & CV & CV & CV & CV & CV \\
\end{array}
\]

a. H

b. H
I think the evidence here suggests constituents of the sort seen in (7), and the evidence available is the foot as a unit of tone-spreading, as we shall see. In any event, the proposal is that the first High tone served as the "anchor-point" of the spreading of metrical structure, both to the left and to the right. In diagram (6) we have seen the kind of use that has been made of metrical structure to the left; we will turn now to evidence in the verbal system of metrical structure to the right of the anchor point. However, the vowel or the position that serves as the anchor point in the verbal system is probably quite different in the present day language from what it was in the earlier stage, the period of Proto-Middle Lake. Let us turn first to the situation in Kirundi.

4. THE KIRUNDI VERB SYSTEM

Now in Kirundi, we are struck by a rather different set of facts and rules from the nominal system of Kinyarwanda, but still one requiring a hypothesis of metrical or rhythmic structure. A rule virtually identical to the Iroquoian Strong Pullback rule applies (as it does, too, in Kinyarwanda, in fact) to shift a High tone one vowel to the left when the High tone falls on a metrical weak position. The notion of metricaly weak and strong positions is defined synchronically in Kirundi in a somewhat different way than was originally the case in Proto-Middle Lake, but I will begin with the current situation in Kirundi, and then present a reconstruction of the earlier situation. The Kirundi verb has the structure in (8).

(8) Finite Verb:
Subject Tense Focus Object Object Verb Suffix(es)
Marker Marker Marker Marker Radical
{........Base.................................}
{....Stem...........}

Infinite:
ku— +Base (defined above)
nominal
prefix

I have indicated two object markers in the verb; they are optional, of course, and in fact any number from zero to two can occur. The verb radical itself has an inherent tone, either Low or High. We shall return immediately to the question of what vowel that tone appears on.

When there is no object marker before the radical, the lexical tone falls on the first syllable of the radical, as in (9). (There are, indeed, cases where the lexical tone is overridden, but that is not relevant to our present concern.)
(9) Low tone verb: *ku rim a* 'to cultivate'
High tone verb: *ku bón a* 'to see'

When we add an object marker to the verb, in three tenses we find a lexical High tone pulled back from the verb radical to the preceding object marker. (This statement holds for the Low toned object markers, which includes all of the object markers except the reflexive, which is High toned, and possessed of a long vowel.) These three tenses include the infinitive, as illustrated in (10). The pullback does not occur, for example, in the Recent Past, as illustrated in (11).

(10) Low tone verb: *ku ki rim a* 'to cultivate it'
High tone verb: *ku ki bón a* 'to see it'

(11) Low tone verb: *n æa ki rim a* 'I cultivated it'
High tone verb: *n æa ki bón a* 'I saw it'

I would like to suggest that the leftward shift of tone in (10) is the result of the Strong Pullback rule, formulated much as in (3) above. However, in modern Kirundi, metrical structure is built on the base, as defined above; metrical Strong-Weak feet (not the Weak-Strong feet of Iroquoian) begin after the tense marker. The first position in the base — that of the first object marker, when there is one — is always a Strong position. Thus the Strong Pullback rule is as given in (12), but it applies only in three tenses: those where the tense marker is of the form CV (the infinitive, the present tense, and the -ka- tense). Rule (12), as formulated, does not yet restrict its own application to just those tenses. I shall suggest below that this condition has been morphologized; the condition is not a simple phonological one, though we shall take as our task in Section 5 to account for this generalization, a generalization which was once phonologically regular. We will propose that the earlier metrical structure was once responsible for determining which tenses display Tone Pullback.

(12) Strong Pullback rule: Kirundi

```
S  W
V  V
-----
Z  H
```
This is illustrated in (13).

(13) ku ki hon a 'inf.-it-see-suffix'

\[ \begin{array}{c}
\text{Z} \\
\text{H}
\end{array} \]

The possible tenses which we shall look at are the following (giving here the Kirundi forms):

(14) Tenses where Pullback occurs Tenses without Pullback

| Present:  | -ra-   | Recent Past:  | -aa-   |
| Narrative: | -ka-  | Far Past:    | -d-ra- |
| Persititive: | -ki- | Future:      | -zbo- |

We are keenly aware that metrical theory very much underdetermines the ways in which rhythmic structure should be erected to produce the placement of Weak and Strong positions so that the first object marker, for example, will always fall on a Strong Position. One could propose, as we have, that metrical structure begins there, calling this unit the base; but how do we know that the Strong-Weak feet do not start, for example, on the tense marker itself, or at the beginning of the word?

I will present below my reason for believing that rhythmic structure begins with the base, i.e. with the first object marker; however, this appears to be an innovation in Kirundi itself, as I will suggest when I present a reconstruction of an earlier stage.

The tonal behavior of verbs in negative, subordinate and subjective tenses is essentially identical in Kirundi, but distinct from that of the main clauses that I have mentioned so far. In these other tenses (which I will simply call "subordinate verbs"), the lexical tone of the verb is lost, the first mora of the verb stem is Low, and the second syllable, in most cases, is raised to High, as we see in (15), for example.

(15) kó n-a- ki-rim- ír- a

that I recent is cultivate-for-suffix

past

kó n-a- rim- ír- a

that I recent cultivate-for-suffix

past

When the vowel of the first syllable of the verb stem is short, then it is always the second syllable that receives the High tone, as in (15), regardless
of whether there is an object marker or not; there is no apparent effect of any Strong Pullback, at least not onto an object marker; see (16).

(16) Metrical structure

However, when the vowel of the first syllable is long, then the placement of the subordinate verb High tone varies, depending on the presence or absence of an object marker — or, in the terms of our analysis, depending on metrical structure.

(17) a. \( n a \  ki \ raab \ ir \ a \)
I past it see for suffix
b. \( n a \  raab \ ir \ a \)
I past see for suffix

See (18):

(18) a.

b.

Now this rightward shift to a strong position in the subordinate forms occurs independently of the form, shape, or number of moras in the tense marker. Hence the rhythmic or metrical structure we need for these subordinate
forms must begin strictly with the base, i.e. with the first object marker. This conclusion, based entirely on the subordinate forms, is entirely consistent with the position sketched above, for the main clause forms and the formulation of the Strong Pullback rule given in (9).²

The point of all this, then, is to motivate an analysis of Kirundi that relies on metrical structure, but structure which consists of Strong-Weak feet established from left to right, beginning at the beginning of the "base", i.e. with the first object marker or with the stem, if there are no object markers.

But so far we have no explanation for why the Pullback rule is operative only in the infinitive, the present tense -ra- tense, and the consecutive -ba- tense. What do these tenses have in common, and why does this rhythmic rule apply only in these tenses?

§ 8. RHYTHMIC STRUCTURE IN THE VERB IN PROTO-MIDDLE LAKE

Now the Pullback rule of Kirundi discussed in Section 3 reflects a situation it shares with Kinyarwanda; in essentially all verb forms, the Kinyarwanda tonal pattern can be formed by adding a single High tone to the left of the Kirundi form. This is, of course, the same leftward tone doubling mentioned immediately above, which we said was an innovation particular to Kinyarwanda.

We will concentrate on the Kirundi verbal forms, therefore, recognizing that the Kinyarwanda forms will, ultimately, be derivable. But why, as we have asked, is the Pullback rule restricted to three tenses?

Linked to this question is the other, so far unasked question: why should metrical structure be erected in the verb from left to right starting with the "base"? We recall, after all, the hypothesis that I dropped in section 3, the

² The unusual behavior that Kirundi and Kinyarwanda show in the shifting placement of the "second mora High tone", as we may call it, can be shown to have arisen before metrical structure came into the picture. Shi, studied by Poliak-Bynon, is a tone language of the Group J (which includes the Middle Lake languages) that is much more similar tonally to Luganda or Haya than it is to the Middle Lake languages. We can see in Shi, however, the origin of the shifting placement of the second-mora High tone for long-vowel stems. All the Group J languages have certain tenses in which a High tone is placed on the second mora of the stem, and Shi follows this generalization. In Shi this High tone is apparently placed on the second syllable, not the second mora, when the first syllable has a long vowel, just as we saw in Kirundi. A shift in the location can in one environment be found in Shi, however, with this High tone back on the second mora of a long vowel in the verb stem, paralleling the situation in Kirundi. But in Shi this shifting of the second-mora High tone occurs only when the object marker that precedes it is High toned, and this process undoubtedly harks back to more complex tonal phenomena that the Middle Lake languages have lost — primarily because they have lost the tonal contrast in all object markers except the (long voweled) reflexive marker.
hypothesis that metrical structure was erected wave-like in both directions from the first High tone of the word, a hypothesis that was easily supported in the nominal system.

I would like to suggest that the basic generalization is this: tenses in which the Pullback rule does not apply are tenses where the tense marker bears a High tone or a long vowel, and that this inherently prominent syllable was consistently chosen as the anchor point for the metrical structure. The following syllable, then, which would be the beginning of the base, would always be a Weak position. This is illustrated in (19).

\[ S \quad W \quad S \quad W \]
\[ \sigma \quad \text{bu} \quad \text{zoo} \quad \text{ki} \quad \text{tem} \quad \text{a} \quad \text{they-Future-it-cut-suffix} \]

(19)

Verbs with tense markers with short, Low toned vowels — the infinitive, the -kar- tense, the present -ra- tense (see (14)) — were assigned metrical structure from the beginning of the word, as in (20).

\[ S \quad W \quad S \quad W \]
\[ \text{bu} \quad \text{ru} \quad \text{ki} \quad \text{tem} \quad \text{a} \quad \text{they-Press-it-cut-suffix} \]

(20)

This pattern was then eventually morphologized: it cannot be maintained as a synchronically valid principle. Consider, for example, the Far Past, which is a tense with a High toned syllable in the tense marker (-ora-): by the account just offered, we would expect the following syllable to be a weak position, and since the Far Past does not show Pullback, we would expect this weak position to be the first syllable of the base. But since the Far Past marker is (synchronously) bisyllabic, an alternating rhythmic would lead to the base-initial position being strong, not weak. However, comparative work (see Goldsmith and Sabimana 1984, for example) shows that in the Far Past, a Focus Marker position was created between the Tense Marker and the object marker. Before that Focus position arose, the Far Past marker would fall phonologically in with markers like the Future marker in (19), and we can infer that the morphological tense-limitations on Pullback arose before the focus marker arose in the language.
6. CONCLUSIONS: CAUSES OF THE RISE OF METRICAL STRUCTURE

We have sketched a historical picture in which metrical structure arose in Proto-Middle Lake out of what must be called a pure tone language, that of an earlier stage of Proto-Lacustrine Bantu. The prosodic system that served as the basis for this innovation was one in which the prosodic system already served many grammatical functions, quite unlike the case in a simple system such as that of Proto-Northern Iroquoian.

At some point in Proto-Middle Lake, a pattern of alternating weak and strong positions was established in the word. The “anchor point” for this rhythmic pattern varied, the point outward from which the rhythmic pattern arose, and for this there were three related subcases. In nouns that had a High tone, that High tone was the anchor point. For verbs, in certain tenses the tense marker served as the anchor point. This was the case when the tense marker had inherent prosodic prominence, being either High toned or long; furthermore, a long tense marker (-ző-, -aa-) acted as a single unit as far as assignment of weak and strong units was concerned. Finally, in verbs where the tense marker did not have inherent prominence, rhythmic structure was assigned starting from the beginning of the word.

When historically appropriate and morphologically expected placement of High tone led to the assignment of a High tone to a metrically weak position, rules arose that shifted the High tone — sometimes to the left, sometimes to the right — to achieve a structure on the surface where in most cases High tones were associated with strong positions.

What factors led to the development of this metrical system? This is a hard question to answer at this point, but certainly an important question to keep on trying to answer. Two likely factors are:

1. The rise of a one-High-tone-per-word principle. We see in these languages a drift of some sort towards a surface restriction to one tone or string of Highs, unlike a true tone languages. This is probably related to ‘culminative function’ of accent (Trubetzkoy, Hyman), and perhaps to the rise of a metrical system.

2. Rhythmic pattern of HLHL in the two object marker case. We can infer from comparative evidence that it is likely that the tone pattern on the verb when two object markers were present was High Low High Low, over the two object markers and the verb radical; it is possible that this tonally rhythmic pattern contributed to the reanalysis that we have observed.

Finally, I would emphasize the guiding hypothesis that lies behind the analyses of these Middle Lake languages, and to some extent the Northern Iroquoian systems discussed earlier: that formally odd, and apparently unmotivated, rule systems may be the result of the interaction, both synchronically and diachronically, of more than one prosodic system (here, rhythm-
mic and tonal), with a strong tendency on the part of the language to encourage the different prosodic systems to line up, and to harmonize with each other.

REFERENCES


