Tonal density index in Mande and beyond

Valentin Vydrin

INALCO – LLACAN, Paris

Tonal studies in African linguistics since Diedrich Westermann's times

- Westermann was one of the pioneers of tonal studies in African languages: Kpelle, Ewe, Igbo… (cf. M. Delafosse's opinion on Mande languages as non-tonal).

- In the first half of the XX century, no distinction between underlying and superficial tones was made; the idea of „one to one" correspondence between tones and syllables predominated (cf. Pike 1948).

- Nowadays, it is evident that the great majority of the languages of Subsaharan Africa are tonal.

- It is also clear that the types of tonal systems vary very much.

„Traditional" comparison of tonal systems

- By the number of tones.
- By type (level tones vs. contour tones).
- By function (lexical vs. grammatical tones).

Carlos Gussenhoven: The tonal density

- In some languages, (nearly) all the tonal bearing units (TBU) have distinct tones. These are languages of a high tonal density.

- In some others, only some TBU carry distinctive tones. These are languages of a low tonal density.

- This parameter allows a quantitative comparison of tonal systems (including „pitch accent" types); their difference will be expressed numerically.

- However, Gussenhoven has not developed a method to define the tonal density in particular languages.

Tonal Density Index (TDI)

- We (Elena Perekhvalskaya & me) introduce a Tonal Density Index (TDI) to show more precisely the importance of tones for coding values in a particular language.

- The TDI is the ratio between the number of tones or tonemes (i.e. meaningful tones, or melodies) and the number of segmental units.

- TDI = number of tones (tonemes, melodies) per 100 syllables (or marae) in a random text.

Key notions necessary for the calculation of the TDI

- toneme
- tonal domain
- marked tone and default tone
- basic segmental unit (BSU)

Toneme (melody)

- A toneme is a meaningful tone, i.e. a tone (tonal contour) which is relevant for the contrast of lexical or grammatical meanings.

- Contextually and positionally conditioned (and, therefore, predictable) tones are not considered as tonemes.

- In some languages, distinction between tones and tonemes is irrelevant (all tones are meaningful and can be regarded as tonemes).

Tonal domain

- Tonal domain is a segmental chain on which a toneme is realized.
- E.g., in Bambara, tonal domains of variable length:

cɛ̀ ´ man'

cɛ̀kɔrɔ ´ old man'

cɛ̀kɔrɔba 'adult man'

- In a language where tones and tonemes are not distinguished, it is unnecessary to postulate tone domains.

Marked and default tones

- In many languages with two contrastive tone levels, one tone (most often, high) can be regarded as marked, and the other one, as default; this situation can be regarded as follows: only high-toned syllables carry tones, other syllables are toneless. In such languages, only marked tones should be counted.

- In the languages where only marked tones are counted, the notion of tonal domain seems irrelevant.

- It would be wrong to think that any 2-level tone system are to be interpreted in the terms of "marked tone vs. toneless syllables". There are languages with two tones where no tone can be regarded as void.

Basic segmental unit (BSU)

- Gussenhoven and Hyman suggested to count the tonal density per morae.

- Theoretically, both types of TDI make sense: the syllabic TDI and the moraic TDI.

- In syllable-counting languages (where syllable = mora), the numeric values are the same in both cases.

- In mora-counting languages, the values of the syllabic TDI and the moraic TDI will diverge.

Problems concerning the syllabic TDI

- In some languages, syllables are hard to discern; ex., in Gokana (Hyman 2009).

- In many language, there is a problem of interpretation of the sequences CVV, monosyllabic (CV:) or dissyllabic (CV-V).

- My suggestion: the interpretation depends on the monophonemic or biphonemic interpretation of the long vowels. Therefore, it is language-specific.

Problems concerning the moraic TDI

- A „standard" situation in a moraic language: 1 light syllable = 1 mora, 1 heavy syllable = 2 morae.

- There seems to be no unity on the issue of identification of a mora and its relevance in some languages. See below on the problem of segmentation of syllables into morae in Bambara.

- Sometimes, even in closely related languages, heavy syllables are interpreted differently with respect to the number of morae.

- There are languages where 3 or even 4 degrees of syllabic weight are postulated (e.g., some Pular-Fulfulde dialects).

The choice

- Cross-linguistically, segmentation into syllables seems less problematic than segmentation into morae.

- The syllabic TDI, even if its calculation is not devoid of difficulties, seems more universal and provides a better comparability.

- Therefore, we suggest the syllabic TDI as more universal (although the moraic TDI can be also calculated, if necessary).

A very preliminary hypothesis:
Three types of tonal languages

- „syllabic tonal languages": TDI close to 100%, each syllable has its distinctive tone.

- „tonemic languages": TDI > 50%, tonemes and tonal domains can be postulated.

- „languages with marked and default tones": TDI < 50%. At its lower end, this type comes close to a toneless language: if a language has regularly only one marked tone per word, such a language can be regarded as accentual, rather than tonal.

Case studies. 1. Eastern Dan

- 5 level tones, 1 to 3 contour tones (tones or combinations of tones?)
- replacive grammatical tones
- additive grammatical tones
- profusion of structures CVV, ClVV, CVŋ, CVVV etc. Syllables or sequences of syllables?

Status of the structures CVV, CVVV... etc.: bi- or threemoraic syllables or featural feet?

- If the "moraic" interpretation is accepted, we have to postulate long and extralong vocalic phonemes and several dozens of (phonemic) diphthongs and triphtongs (*oȃ, îȓ, îɛɛ, etc.),* many of them extremely rare (1, 2, 3 occurrences in the vocabulary).

- There are some instances of morphemic boundaries inside the sequences CVV (e.g., *yīʎ* 'otter' < *ýí* 'water' + *gʉ́* 'in').

- Vowels of the sequence CVV can be sometimes separated by an infix, e.g.: *fʉ́ʉ̋fʉ̀ʉ̏* 'light (not heavy)' → *fʉ̋kʉ̋fʉ̏kʉ̏* 'very light' (an intensive form).

- Therefore, these structures are dissyllabic or trisyllabic feet (CV-V, CV-V-V, etc.).

Tonal mapping

- By default, one tone (toneme) is carried by one syllable (= vowel).

- In trisyllabic feet, 1 or 2 different tones are allowed: *ɓúʌ̀ʌ̀* 'beard', *ɓã̄ã́ã́* 1sg. negative focalized pronoun (and not **ɓūʌ́ʌ̀* or **ɓáā̄ã́*). Therefore, all 3-syllabic feet can be considered as carrying 2 tones (rather than 3).

- Replacive grammatical tones always substitute lexical tones of the entire foot. E.g.: *dlã̄ã̂* 'teach' → *dlã̏ã̂* 'teach\ɴᴇᴜᴛ'. Such grammatical tones can be considered as 1 toneme.

- Additive grammatical tones are associated with the preceding syllable which carries, subsequently, 2 tonemes: *gó`* 'sell/ɪɴꜰ'.

TDI for Eastern Dan

- In a text containing 138 syllables, I have counted 146 tonemes.
- The TDI is 105,8.
- Eastern Dan is a syllabic tonal language.

Case study 2: Bambara

- Syllable, mora, featural foot:

- A foot realized as CCV can be always pronounced as CVCV in a slow speech (e.g., *flà* 'two' = *fìla*), therefore, it is considered dissyllabic.

- A contrastive long vowel is allowed only in the non-final position in a foot, e.g. *mí:ri* 'to think', *bá:ra* 'work'. The sequence CV: is considered as one syllable, rather than two (no morphological boundary can cut it into two vowels; there are no diphthongs…).

The problem of mora in Bambara

- Two or even three tones can be docked on one syllable, e.g.: [cě`] 'man/ART', [mùsô] 'woman/ART'. In this case, the vowel of the syllable is realized as long.

- However, the same vowel is realized as short in a different position in a word, ex.: *cèfarinya*` [cèfàrìnyâ:] 'courage', *mùsokɔrɔba*` [mùsòkɔ̀rɔ̀bâ:] 'old woman'.

- In Bambara, there is no difference in syllabic weight which would affect the ability of a syllable to bear more than one tone.

- Therefore, mora is irrelevant in Bambara.

Tonal system

- Two tonemes, high and low. Contour tones (rising and falling) on one syllable are allowed, but they are regarded as combinations of level tones.

- A toneme is mapped on a tonal domain. The size of a domain most often equals to a word, however, there are words which contain 2 tonal domains: *(kùnna)(bìla)* 'to reproach'.

- On the other hand, one tonal domain can include elements which can be considered sometimes as separate words, e.g.:

(à bɛ) (ń) (dɔ́n) 'He knows me.'

- A tonal domain can have zero segmental base (a floating low tone):

(mùso)(`) (kúlela) 'The woman yelled.'

Is Bambara a "type 3" language?

- It can be said that in a dissyllabic or plurisyllabic tonal domain, the tone is lexically attributed to the first syllable, and the other syllables are toneless (in fact, they bear surface tones, but these tones are mapped according to general rules of Bambara), ex.:

súruku [súrúkú] 'hyena', *bàlaka* [bàlàkà] 'to precipitate'

- Since (Creissels & Grégoire 1993), the established opinion is that in Bambara (and other Eastern Manding varieties), the low tone is marked (active), and the high tone is a tone by default: *bàlaka*, but *suruku*.

- However, it seems inappropriate to regard high-toned domains as „toneless": such interpretation would blur the distinction between syllables to which a high tone is attributed lexically, and „true toneless syllables", to which not tone is attributed lexically.

TDI for Bambara

- In a text containing 100 syllables, I have counted 70 tonemes.
- The TDI is 70.
- Bambara is a tonemic language.

Case 3: Navajo (analysis by Elena Perekhvalskaya)

- Complex verbal morphology, long polysynthetic verbal forms.

- Long vowels are frequent in texts, they are interpreted as biphonemic combinations. Two adjacent vowels can bear different tones.

- Two tones, high and low. High tone is marked, the low is a default tone.

- TDI for Navajo: 48 high tone syllables per 134 syllables; TDI = 35.8.

Some other languages, by types

- Type 1 (syllabic tonal languages"): Mwan (South Mande; the data from Elena Perekhvalskaya), Akebu (Southern Guang < Kwa, the data from Nadezhda Makeeva), Ikposo (< Northern Guang < Kwa, the data from Nadezhda Makeeva).

- Type 2 (tonemic languages): Kpelle, Mende (Southwestern Mande), most of Central Mande languages; Tibetan (data from Elena Perekhvalskaya).

- Type 3 (languages with marked and default tones): presumably, many Bantu languages.