Rhythm in West African tone languages: a study of Ibibio, Anyi and Ega

Ulrike Gut, Eno-Abasi Urua, Sandrine Adouakou & Dafydd Gibbon
Bielefeld University, University of Uyo
{gut, urua, adouakou, gibbon}@spectrum.uni-bielefeld.de

Abstract

The speech rhythm of three West African languages is measured with current methods which focus on durational relationships of various units. It is shown that predictions about the phonological structure and their classification in terms of speech rhythm can be confirmed with data from these languages.

1. Introduction

The concept of speech rhythm is a much discussed issue in phonetics and phonology. Impressionistic accounts agree that the languages of the world differ in their rhythm (syllable-timing, stress-timing, mora-timing) but attempts to capture these differences acoustically have so far been unsatisfactory. The aim of this paper is to describe three West African tone languages – Anyi, Ega and Ibibio – which have not been previously classified in terms of their speech rhythm. We started from the first working hypothesis that they are syllable-timed, as reflected in a low variation between syllable durations. Next, we tried to correlate phonological features such as absence/presence of phonemic vowel length, syllable structure, and tone system with the speech rhythm.

2. Rhythm measurements

The rhythm of the languages of the world have traditionally been divided into stress-timed and syllable-timed (Pike 1945, Abercrombie 1967). Rhythm is understood to be a periodic recurrence of events, which in stress-timed languages are stress beats and in syllable-timed languages syllables. Abercrombie (1967) sees speech rhythm as "essentially a muscular rhythm, and the muscles concerned are the breathing muscles" (p. 96). In syllable-timed languages "chest-pulses, and hence the syllables, recur at equal intervals of time – they are isochronous" (p. 97). Syllables are assumed to be equal in length (p. 98), stress-pulses, on the other hand, are unevenly spaced. Abercrombie cites Yoruba, a West African tone language, as an example for a syllable-timed language.

Stress-timed languages such as English, in contrast, are supposed to have regular recurring stress beats with the same time interval separating two beats of equal length. Since the number of syllables between two stress beats varies, their length is adjusted to fit into the stress interval – syllable length, hence is very variable in stress-timed languages.

Many researchers have tried and failed to find an acoustic basis for these claims. First, the interstress interval in stress-timed languages such as English is not of equal length (Classe 1939, O'Connor 1965, Uldall 1971, Hill et al. 1979, Fauré et al. 1980, Roach 1982, Dauer 1983), but varies from 488 to 566 ms. Roach (1982) divided the duration of a tone unit in three stress-timed (English, Russian, Arabic) and three syllable-timed languages (French, Telugu, Yoruba) by the number of feet, which gives a hypothetical ideal duration for a foot with complete isochrony assumed. Actual measurements of feet durations taken from speech in these languages were then compared with the predicted value and the percentage deviation was calculated. Roach showed that the variance of the percentage deviation in English is higher than in French, Telugu and Yoruba, which is contradictory to expectations.

Second, syllable length in stress-timed and syllable-timed languages does not differ significantly. Roach (1982) calculated the standard deviation of syllable durations in English, Russian, Arabic and French, Telugu and Yoruba. No significant difference was found: English 86ms, Russian 77ms, Arabic 76ms; French 75.5ms, Telugu 66ms and Yoruba 81ms.

Dauer (1983) suggested that “rhythmic differences […] between languages […] are more a result of phonological, phonetic, lexical, and syntactic facts about that language than any attempt on the part of the speaker to equalize interstress or intersyllable intervals” (p. 55). In Dauer’s view, speech rhythm reflects variety of syllable structures, phonological vowel length distinctions, absence/presence of vowel reduction and lexical stress. Whereas languages classified as stress-timed such as English show a variety of different syllable structures (CV (30% frequency), CVC (34%), VC (15%), V (8%), CVCC (6%)), languages classified as syllable-timed have a majority of CV syllables (58% for Spanish). Since syllables increase in length when segments are added and closed syllables are longer than open ones (Delattre 1966), speech rhythm measured in syllable duration differences reflects syllable structure distribution. Equally, differences in rhythm between languages reflect whether a language has vowel reduction or not; those classified as stress-timed usually do. In addition, syllable-timed languages either do not have lexical stress or accent is realized by variations in pitch contour. Conversely, stress-timed languages realize word level stress by a combination of length, pitch, loudness and quality changes, which result in clearly discernible beats.

This approach is partly reflected in recent measurements of the acoustic correlates of speech rhythm. Ramus, Nespor & Mehler (1999) segment speech into vocalic and consonantal parts and compute the proportion of the vocalic intervals of a sentence, the standard deviation of these intervals, and the standard deviation of the consonantal intervals. By comparing carefully selected read sentences by four speakers each of 7 different languages along the axes of the percentage of vowels and the standard deviation of the consonantal intervals Ramus et al. succeed in grouping some languages similarly to the originally suggested groups of stress-timing and syllable-timing. English, Polish and Dutch, all presumed stress-timed languages, group together with a relatively low vocalic proportion (around 39%) and a relatively high standard deviation of consonantal intervals. French, Italian, Spanish and Catalan group together at a higher %V (about 44%) and lower delta C. Japanese, finally, differs from those two groups by having an even higher vocalic proportion (53%) and even lower consonantal standard deviation.

Grabe & Low (2001) measure the difference in duration between successive vowel durations and between successive consonantal intervals. Both approaches succeed in classifying languages that show mixed phonological properties as suggested by Dauer (1983), e.g. vowel reduction but small variation in syllable structure.

With the exception of Roach (1982), the speech rhythm of West African languages has not been investigated yet. In this paper, we selected three of these languages, Anyi (Kwa), Ega (putative Kwa) and Ibibio (Benue Congo: Lower Cross). Our hypothesis is that they have a syllable-timed rather than stress-timed rhythm and that their phonological features such as absence/presence of phonemic vowel length, syllable structure, and tone system correlate with their speech rhythm.

3. Languages

3.1 Anyi

Anyi is spoken in the Eastern part of Ivory Coast where according to an inventory by Burmeister (in prep.) 10 varieties exist. In Ghana, two additional varieties of Anyi are spoken. According to the classification by Stewart (1989), Anyi belongs to the Kwa languages. The consonantal inventory of Anyi includes a series of voiced and voiceless stops: bilabial, dental, palatal, velar and labio-velar. Voiceless labio-dental, dental and glottal fricatives occur. In addition, there are liquids, nasals and glides. In general, apart from the labio-velars, which are absent in the Ndenye variety of Anyi, all the consonants are permissible in the consonant structure of Anyi. Anyi has 14 vowels, 9 oral and 5 nasal. The Sanvi variety has more than 14 vowels, including the central vowel /□/ and the corresponding nasal vowel /□~ /. Phonetic long vowels occur but their analysis as a sequence of two vowels has been suggested, which is corroborated by the fact that they almost always have a complex tone. Words with contrasting vowel length therefore always also have a tonal contrast. Only the vowels of the third and fourth degree of aperture do not allow corresponding nasals. The following syllable structures occur in Anyi: V, CV, N, CVV, C_1C_2V where C_2 is a semivowel, and CLV where L is a liquid. Anyi has four phonological tones: Two level tones, H and L, and two contour tones, rising LH and falling HL. The appearance of a mid tone is due to the effect of tone sandhi rules.

3.2 Ega

Ega is an endangered isolate within a Kru speaking area of South Central Ivory Coast (Dida to the West, North, East; Godie to the South), with around 1000 full speakers. Ega has been classified as a Kwa langauge. However, much of the information about the language is of uncertain status, and Ega is currently the subject of ongoing linguistic and sociolinguistic documentation research by Connell, Ahoua, and Gibbon (2002). A published sketch of the language by Rémy Bole-Richard is included in Hérault & al. (1982), and a detailed phonetic study has been made by Dago (1999). Although the language has been classified as Kwa, Ega has phonological features, such as the full implosive series, and morphological features, such as a complex nominal classification system, which distinguish it from the geographically nearest Kwa languages. Syllables are V, CV, CCV. The consonant system contains a full series of unvoiced, voiced (voiced fortis) and implosive (voiced lenis) stops: labial, dental, palatal, velar, labiovelar. There are 9 vowels, with ATR harmony, but no nasal or length contrast (except in isolated words such as /'fe:~/ "all"). The tone system of Ega has a three-way contrast: high (H), mid (M) and low (L). Initial observations indicate that Ega has discrete-level tone patterning in context, with abrupt final lowering.

3.3 Ibibio

Ibibio has been classified as a Lower Cross language ((New) Benue-Congo) spoken in the south-eastern part of Nigeria (Faraclas 1989, Williamson 1989) by about four million people (Essien 1991). The Ibibio syllable structure is (V/N), CV, CVV, CVC, CVVC, CGV, C□V. The V/N is the syllabic prefix, which may be either a vowel or a syllabic nasal consonant, usually homorganic to the following consonant. Consonant clusters are hardly attested and where they occur, are usually restricted to only C□ or CG. The G is either a palatal or labial-velar glide (Cy or Cw) which arise from deletion or other phonological processes in the language. The □ is an alveolar tap arising from two phonological processes of /d/ weakening to □ and V_1 deletion in a CV_1dV_2 sequence. These structures may be modified through suffixation. Owing to overlapping and neutralisation in Ibibio consonants, there are differences in the number of phonemic and even phonetic consonants proposed for Ibibio. Urua (1990, 2000) proposes thirteen phonemic consonants, comprising five oral and four nasal stops, two fricatives, one palatal and one labial-velar glide. Although voiced and voiceless stops are attested, contrast between them is significant mainly in word-initial position. In word-final position the voicing contrast is neutralised and only unreleased and voiceless stops occur. In intervocalic position, the stops are weakened to homorganic continuants/taps. Consonants may be lengthened for morphophonemic reasons. The number of phonemic vowels proposed ranges from six to ten. The six which are common to all the researchers include /i, e, a, □, o, u/. These six vowels occur in all positions in the word, albeit with allophonic variations. Crucially, all six short vowels contrast with their long counterparts in a C-C environment, e.g. de□p/de□e□p 'buy/scratch'. Vowels undergo assimilation, lengthening, shortening and deletion but not reduction in the sense in which English vowels undergo reduction in unstressed syllables. The tone system fits into a terrace level pattern since it has two level pitches, High and Low plus a contrastive downstepped High tone in addition to two contour tones, High-Low and Low-High. The contour tones are combinations of the level pitches. Both High and Low tones in Ibibio manifest downdrift (Urua 1996/1997, 2000, Gibbon, Urua & Gut 2000).

4. Method

4.1 Subjects

One Anyi speaker, one Ibibio speaker and one Ega speaker were recorded. The Anyi speaker is male and has been living in the Ivory Coast all his life. The Ibibio speaker is female and has lived all of her life in Ibibio land, with occasional sojourns of not more than two years at a time in Scotland and Germany. The Ega speaker is male and has lived in the Ivory Coast all his life.

4.2 Data

The Anyi speaker told a story of 1.42 minutes length. The Ega speaker spoke an address to the elders of his home village lasting about 3 minutes. The Ibibio speaker read a formal address to the elders of her village of about two minutes length. For all three speakers 12 sentences were selected which had been spoken without hesitation or repairs and restarts and which had a minimum length of 8 syllables.

4.3 Analysis

The sentences were transcribed using Transcriber 1.4 (Ega and Anyi) and Praat (Ibibio). The length of each syllable was measured, a phonetic transcription was made in SAMPA, and the syllable structure was transcribed.

The Rhythm Index (RI) (Low & Grabe 1995) for each sentence was then calculated using the following formula:

$$RI = 100 \sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right| / (m-1)$$

where *m* stands for the maximal number of units, *d* stands for duration and $d_i=d_k$ and $d_j=d_{k+1}$. That is, for a sequence of units (either syllables or vowels) of length d, the average difference of the absolute differences between adjacent units is calculated. The differences are normalised by dividing each difference by the average duration of the syllables in the pair. If the units are very similar in duration, the RI will be close to 1, whereas for maximal differences the RI will approach 200. A value of 1 will be interpreted as perfect syllable-timing, higher values reflect a tendency for stress-timing.

In addition, we calculated the Rhythm Ratio (Gibbon & Gut, 2001) for each sentence. This measurement is based on the following formula:

$$RR = 100 \sum_{k=1}^{m-1} \frac{d_i}{d_j} / (m-1)$$

where $d_i=d_k$ and $d_j=d_{k+1}$ if d_i is smaller than d_j and $d_j=d_k$ and $d_i=d_{k+1}$ if d_i is not smaller than d_j. In other words, for each pair of adjacent syllables, the shorter is divided by the longer. The average of all these ratios is calculated and multiplied by 100. Thus, if the RR equals 100 we have perfect syllable-timing. The lower the degree of syllable-timing the lower the RR value. Unlike the RI, the RR does not calculate absolute differences in length between adjacent units but computes their ratio. Also unlike the RI, the RR measurement does not normalise for duration.

With the help of a wide-band spectrogram, vocalic and consonantal parts of the speech signal were annotated. In order to ensure comparability, the annotation technique used by Ramus et al. (1999) was adopted. This means that pre-vocalic glides were treated as consonants whereas post-vocalic glides were treated as vowels. Thus, vowels were coded as V and stops, fricatives, liquids, nasals, glides, implosives and approximants were coded as C. The beginning and end of a vocalic interval was determined using standard phonetic criteria.

5. Results

Table 1 lists the average RI and RR across all twelve sentences for Anyi, Ega and Ibibio. All three languages exhibit a clear tendency towards syllable-timing, i.e. the difference in duration between adjacent syllables is relatively small. Both the rhythm ratio (RR) being much closer to 100 than to 0 and the rhythm index (RI) being closer to 0 than to 200 indicate this.

	RR	RI
Anyi	65.8	43.9
Ega	70.1	37.3
Ibibio	66.3	42.4

Table 1. *Average RR and RI (syllables) across the 12 sentences in Anyi, Ega and Ibibio.*

Figure 1 shows the speech rhythm of Anyi, Ega and Ibibio, measured in the method proposed by Ramus et al. (1999) compared to the other languages investigated by them. Anyi and Ibibio group closer to the syllable-timed than the stress-timed languages, although they have a higher delta C value. Ega is even below Japanese with a very high vocalic proportion and a very low delta C value.

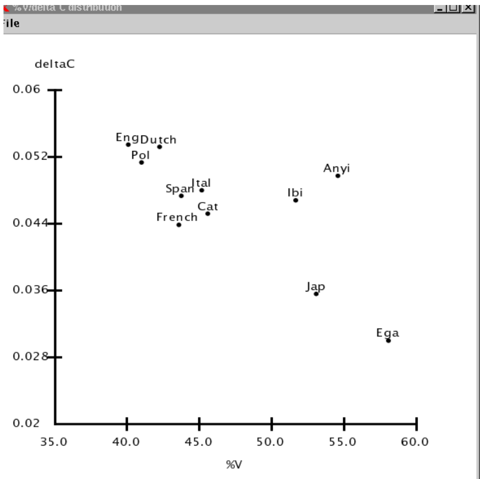

Figure 1. *%V and delta C of English, Dutch, Polish, Spanish, Italian, French, Catalan and Japanese classified by Ramus et al. (1999) and Anyi, Ega and Ibibio.*

Table 2 presents the relative frequency of different syllable types in Anyi, Ega and Ibibio. Differences in both the occurrence of syllable types and their relative frequency appear. In the Anyi speech, only six different syllable types; CV, V, CCV where the second C is either a liquid or an approximant, syllabic N, CVV and CCCV with a liquid and an approximant as the second and the third consonant) occur with CV being the most frequent (61.5%) and V being the second most frequent (21.9%). In Ega, we observed nine different syllable types; CV, V, CCV with /r/ or /j/ as the second consonant, CVV, CCCV with a liquid and an approximant as the second and the third consonant, VN, CVN, and CCVN. The most frequent syllable type is CV (70.1%), the second most frequent type is CVV (14.9%) and the third most frequent V (9.9%). In Ibibio, seven different syllable types were produced; CV, V, CCV with /n/ and /r/ as second consonant, N, CVV, CVC with either a nasal or a stop in the coda position and CVVN. The most frequent syllable type is CV (47%), the second most frequent is CVV (22.9%) and the third most frequent is CVN (14.7%).

	CV	V	CCV	N	CVV	CCVV	CCCV	VN	CVC	CCVN	Total
Anyi	61.5	21.9	9.1	3	3	-	0.6	-	-	-	164
Ega	70.1	9.9	2.5	-	14.9	1	0.35	0.35	0.35	0.35	281
Ibibio	47	5	1.7	7.8	22	-	-	-	15.1	0.9	231

Table 2. *Percentage of different syllable types occurring in Anyi, Ega and Ibibio.*

In Anyi, 3% of all syllables consist of syllabic nasals, but no closed syllables occur. Ega shows a very high percentage of open syllables (99%) and a complete absence of syllabic nasals. Only nasals occur in the coda of closed syllables. Ibibio has more than 16% closed syllables, all with a nasal in the coda position, and more than 7% syllabic nasals.

6. Discussion

Our data show that the degree of syllable-timing, as measured both in the RR and the RI, as well as with the method proposed by Ramus et al. (1999) is very high for all three West African languages under investigation in this paper. As we had predicted in our first hypothesis, durational differences between adjacent syllables are not very pronounced. However, there is no acoustic evidence for equal length of syllables as originally proposed by Abercrombie (1967).

The syllable structures found in Anyi, Ega and Ibibio speech show differences between the first two and the third language. Closed syllables do not exist in Anyi and are very rare (1%) in Ega, where only a nasal can occur in the coda position. In Ibibio, conversely, 16% of all syllables are closed and both nasals and stops occur in the coda position. In addition, Ibibio is the only language with phonemic vowel length contrast. However, this difference in syllable structure and phonemic vowel length contrast is not reflected in a difference in speech timing between those three languages. We therefore did not find any support for Dauer's (1983) claim that speech rhythm reflects the variety of syllable structure and presence or absence of phonemic vowel length contrast and must consequently reject our second hypothesis. Instead, we propose that, in Ibibio, a compensation for these durational factors occurs and that speech rhythm is the dominant frame requiring this.

References

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Allen, G. (1975). Speech rhythm: its relation to performance universals and articulatory timing. *Journal of Phonetics* 3, 75-86.

Bond, Z. & Fokes, J. (1985). Non-native patterns of English syllable timing. *Journal of Phonetics* 13, 407-420.

Burmeister, J. (in prep.) On Anyi.

Classe, A. (1939). *The Rhythm of English Prose*. Oxford: Blackwell.

Connell, B., Ahoua, F. & Gibbon, D. (2002). Ega. *Journal of the International Phonetic Association* 32(1), 99-104.

Dago, G. (1999). *Étude phonétique et phonologique de l'éga*. M.A. thesis, Université de Cocody, Abidjan.

Dauer, R. (1983). Stress-timing and syllable-timing reanalysed. *Journal of Phonetics* 11, 51-62.

Delattre, R. (1966). A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics* 4, 183-198.

Essien, O. E. (1991). The nature of tenses in African languages: a case study of the morphemes and their variants. *Archiv Orientalni* 59:1-11.

Faraclas, N. G. 1989. Cross River. *Niger-Congo languages*. Bendor-Samuel, J. (ed.), Lanham: University Press of America, 377-399.

Faure, G., Hirst, D. & Chafcouloff, M. (1980). Rhythm in English: Isochronism, Pitch, and Perceived Stress. In: L. Waugh & C. van Schooneveld, *The melody of language*, Baltimore: University Park Press, pp. 71-79.

Gibbon, D. & Gut, U. (2001). Measuring speech rhythm. Proceedings of Eurospeech, Aalborg, Denmark, pp. 91-94.

Gibbon, D., E. E. Urua & U. Gut (2000). How low is the floating Low tone in Ibibio. Paper presented at the 30th Colloquium on African Languages and Linguistics, Leiden, August 2000.

Grabe, E. & Low, E.-L. (2001). Durational Variability in Speech and the Rhythm Class Hypothesis. In: Gussenhoven, C. & Warner, N., (eds.), *Papers in Laboratory Phonology 7*, Berlin: Mouton, 515-546.

Hérault, G. (1982). *Atlas des langues kwa de Côte d'Ivoire*. Tome 1. Abidjan: Institut de Linguistique Appliquée & Agence de Cooperation Culturelle et Technique.

Hill, D., Jassem, W. & Witten, I. (1979). A statistical approach to the problem of isochrony in spoken British English. In: pp. 285-294.

Hoequist, C. (1983). Durational Correlates of Linguistic Rhythm Categories. *Phonetica* 40, 19-31.

Hoequist, C. (1983). Syllable Duration in Stress-, Syllable- and Mora-Timed Languages. *Phonetica* 40, 203-237.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics* 5, 253-263.

Low, E.-L. & Grabe, E. (1995). Prosodic patterns in Singapore English. *Proceedings of the International Congress of Phonetic Sciences*, Stockholm, 3, 636-639.

O'Connor, J. (1965). The Perception of Time Intervals. *Progress Report 2, Phonetics Laboratory, UCL*, 11-15.

Pike, K. (1945). *The Intonation of American English*. Ann Arbor: University of Michigan Press.

Ramus, F., Nespor, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition,* 73, 3: 265-292.

Roach, P. (1982). On the distinction between 'stress-timed' and 'syllable-timed' languages. In: D. Crystal (ed.), *Linguistic controversies, Essays in linguistic theory and practice*, London: Edward Arnold, pp. 73-79.

Stewart, (1989). "Kwa". In Benda, S., *The Niger-Congo languages. A classification and description of Africa's largest language family*. Ianhan: University Press of America.

Uldall, E. (1971). Isochronous Stresses in R.P.. In: L. Hammerich, R. Jacobson & E. Zwirner, *Form and substance*, Copenhagen: Akademisk Forlag, pp.205-210.

Urua, E. E. (2000). *Ibibio phonetics and phonology*. Cape Town, South Africa: Centre for Advanced Studies of African Society.

Williamson, K. (1989). Benue-Congo overview. *Niger-Congo languages*. Bendor-Samuel, J. (ed.), Lanham: University Press of America, 247-274.