RHYTHMIC STRUCTURE OF DISYLLABLES IN YORUBA

M.I. KAPLUN

COLLEGE OF ASIAN AND AFRICAN STUDIES, MOSCOW UNIVERSITY,
USSR

ABSTRACT

Studies of the tonal-rhythmic structure of Yoruba words (on the basis of disyllables) make it possible to suggest a general pattern of the rhythmic arrangement of tones and also to formulate the basic rules governing tones in this unit of speech.

The method of synthesis used in linguistic studies makes it possible to approach many linguistic problems at a qualitatively new level. In coping with these problems researchers nevertheless encounter considerable difficulties in view of the fact that, to meet the requirements of modelling speech, it is necessary to know the so called phonetic characteristics of speech which bring simulated speech closer to the linguistic prototype modelled in every particular case, alongside the features traditionally referred to as functional (phonological). Experiments in synthesis show, for example, that when modelling the syllabic tone (as well as the complex tonal-rhythmic mechanism of words), it is necessary to take into account all the multilevel functionally distinctive features.

In natural speech tones are known to be adjusted to each other and for this reason disyllables, the minimal lexical unit in which the basic tonal-rhythmic laws regulated by the rules of sandhi are manifest, have been chosen as the linguistic material in studying the tonal-rhythmic structure of Yoruba. In view of the fact that latest works on tones in Yoruba usually single out three tones, namely, medium, low and high (designated M for medium tone, H for high tone and L for low tone), nine tonal-rhythmic models of disyllables have been chosen for investigation: M+M, M+H, M+L, L+L, L+M, L+H, H+H, H+M, H+L.

In studying the tonal-rhythmic models of great interest are those phonation segments in which tones are joined together because it is precisely there that they are "adjusted" to each other, coordinating in a certain way the contour, register, time and amplitude characteristics. The junctures of tones manifest most graphically the parameters that organise the tonal-rhythmic model as a certain semantic unit of the given linguistic system. It should be stressed that for the tonal-rhythmic structure to be modelled successfully it is necessary to take into consideration not only the contour and register parameters but to equal measure amplitude, time and interval parameters which, interacting with each other, alone can ensure that the given simulated tonal-rhythmic model is fully or at least partially associated by native speakers with its natural prototype.

That is why we can hardly agree with J. M. Hombert[1] who concludes from his analysis of the perception of tonal-rhythmic structures in natural speech on the basis of bisyllabic nouns in Yoruba that acoustic parameters are informative in different degrees. He asserts that the native speakers of Yoruba use only two main acoustic characteristics which enable them to distinguish rhythmic structures in six tonal-rhythmic models. The first characteristic, according to Hombert, is connected primarily with F_0 in the vowel of the second syllable (V_2). The second characteristic is presented by him as a combination of three parameters -- the interval of the modification of F_0 in V_2, the medium value of F_0 in V_2 and the frequency breakage between the end of V_1 and the beginning of V_2. As is seen, Hombert includes only frequency characteristics in the set of functional features, treating amplitude and time characteristics as non-functional. It is apparently indisputable that the native speakers of Yoruba are capable of distinguishing one tone from another in the final position by the features mentioned by Hombert. However, they are inadequate for the synthesis of rhythmic models normative from the point of view of prosody. This is borne out by the analysis of simulated M+L and M+H tonal-rhythmic models, which strictly reproduced all the rhythmically important contour, interval and time parameters and deliberately distorted only the amplitude parameters. After listening to the sounding of these models, the auditors observed that they could hardly be considered as normative.

Analysis of tonal-rhythmic models with the medium tone (M+M, M+L, M+H) shows that it is most stable rhythmically in all the combinations with other tones and retains the even contour, its own duration and medium register (with respect to other tones within the three-level register scale). The Yoruba register scale for male voices close to baritone tenor can be conventionally divided into three ranges:

low	medium	high
90 + 115 Hz	120 + 160 Hz	165 + 200 Hz

The specificity of the M+M tonal-rhythmic model consists in the fact that its contour is formed by two even tones, with the second syllable tone invariably beginning at the frequency of the end of the first syllable. The absence of a frequency interval between the syllables in the medium tone model is important from the point of view of rhythm, while any deliberately made interval between the syllable tones results in the broken rhythm of the model. Depending on the nature of the frequency breakage between the syllables, the auditors described the sounding they heard as a combination of the medium and high tones or the medium and low tones. For the M+M rhythmic model to sound naturally and be perceived unambiguously, its both syllables should be actualised in the medium range exactly. When the register was deliberately changed (with all the other acoustic parameters of the combination of these tones remaining unchanged) the auditors identified them, for example, as a combination of two high tones. The determination of the time relationship between the syllables, one of which is characterised as a medium tone, is a key functional acoustic feature in simulating these tonal-rhythmic models. The presence of the medium tone in disyllables suggests a certain strategy of synthesis, i.e., the creation of an exact time relationship between the medium tone, on the one hand, and the high and the low tone, on the other. Our experimental material included cases of the synthesised medium tone being identified as high. Analysis of this fact brought to our attention primarily the time relationship between the syllables of the disyllable, in which the first syllable (medium tone) is equal in its duration to the second (low tone), the time relationship between syllables characteristic of the combination of the high and the low tone. Consequently, when modelling the rhythm of a disyllable with the medium tone its duration should be longer than that of the syllables characterised by other tones.

One of the key rhythmic characteristics of bisyllabic models with the medium tone in the first syllable is the correlation of the frequency of the end (medium tone) and of the beginning (low or rising high tone), with its indispensable prerequisite being the absence of any frequency contrast at the juncture of tones. Several variants of (M+L and M+H) disyllables with different frequency contrasts and without these at the juncture of syllables have been synthesised to verify this hypothesis. The auditors identified only those of the synthesised (M+L and M+H) rhythmic models which had no frequency rupture between the syllables and also those with the minimal frequency rupture (not longer than a second). The rest of the synthesised models were identified incorrectly. The second-long interval of the frequency rupture between the syllables can, apparently, be considered admissible, whereas a longer one places these rhythmic models outsides the normative rhythm of disyllables.

The peculiarity of the M+H and M+L tonal-rhythmic models consists in the fact that they are organised by two types of equivalent tonal contours, every one of them worked out in the Yoruba language. Different acoustic characteristics become functionally important in their rhythmic organisation. All the programmes of synthesising the M+H and M+L bisyllabic models were recognised as normative when their tonal contour was either formed of two even tones belonging to different register levels or consisted of a combination of the even and the rising contour. Each type of the M+H and M+L models has its own rhythmic peculiarities. In the first variant the functionally important factor is the register contrast between syllables equal to a minor third. The cases of the frequency relationship between the registers exceeding that interval were described by the auditors as "rhythmically pronounced", "exaggerated", or too "robot-like", whereas those with a smaller interval were perceived as a combination of two medium tones. Apparently, an interval of a minor third can be considered

as a sort of a crucial point, beyond which these models disintegrate.
Another important characteristic of the M+H rhythmic model is growing amplitude in the second syllable marked by a high tone with an even contour. Analysis of this type of sounding showed peak amplitude invariably at the second syllable. A sizeable increase in amplitude in the second syllable (compared to the first one) apparently complements the even contour and a similar correlation of these acoustic parameters will produce the high tone effect in the given tonal-rhythmical model for the native speakers of Yoruba. The natural question arises about the functional importance of each of these parameters for the high tone to be perceived unambiguously. With this aim in view several models were synthesised, with only one parameter changing in every one of them and the rest remaining intact. For example, in one case the first and the second syllable of a disyllable had equal amplitudes, in another the amplitude of the second syllable was, on the contrary, increased but there was no register contrast between the medium and the high tone and in still another the first syllable had a bigger amplitude compared to the second syllable. The sounding was repeatedly recorded and offered for auditing at random. The results of the auditing analysis show that none of the acoustic parameters can be singled out as a factor determining the rhythm of a given model. It is rather the combination of these features that accounts for the certain stability of the rhythmic model in any contextual conditions.
Another type of tonal-rhythmic models (M+H and M+L) is formed by an even (in the first syllable) and rising M+H or falling M+L (in the second syllable) tonal contour. The factor important from the point of view of rhythm in this type of model is the rising or falling intervals, which begin at the frequency level of the end of the medium tone. In all the rhythmically normative models of this type the interval between the medium, the high and the low tone was within the range from a minor third to a fourth. It should be pointed out that the register frequency rupture in the first variant of the M+H and M+L models (_ - and - _) the rising and the falling tonal contour in the second syllable of the second variant of the same models (-/ and -\ ) has one and the same set of intervals, which should be not less than a minor third and not more than a fourth because the sounding with an interval less than a minor third does not give stable identification results, whereas that with an interval exceeding a fourth is described as "rhythmically accentuated", for this reason a strictly prescribed interval between the syllables of a bisyllabic structure can be viewed as a key rhythmic characteristic of this type of M+H and M+L models.
Analysis of the M+H and M+L tonal-rhythmic models makes it possible to see that the very term "high tone" and "low tone" accepted in Yoruba tradition and unambiguously defining their register does not exactly correspond to the real acoustic nature of these tones or at best corresponds only to one of the possible variants. This supposition is borne out by the results of the analysis and synthesis of the rhythm of bisyllabic structures, in which the high and the low tone was in the first syllable. The H+L and L+L tonal-rhythmic models, like any other with the high and the low tone in the first syllable, can have a double tonal contour, namely, rising and even (high tone) or falling and even (low tone), which is explained by the complex acoustic mechanism of these tones, which presupposes in each case a certain combination and correlation of register, time, interval and amplitude values.
In one of the programmes of the H+H tonal-rhythmic model recognised by the auditors as "unsuccessful" the tonal contour was formed by two rising tones. The auditors' judgment was, presumably, influenced by the fact that the rising contour of the high tone in both syllables of the disyllable was sounded in the high register, whereas the rhythmic arrangement of the high tone requires either the medium register (rising contour) or the high one (even contour). Each of these parameters in combination with the necessary set of other parameters is rhythmically important and functional in producing a normative sounding of one of the high tone variants. The synthesis of realisations with pronounced register and contour characterstics leads to a rhythmic disharmony, which violates the rhythmic stereotypes worked out in the Yoruba linguistic system and traditional among native speakers.
The low tone in the final position is of special interest in the H+L, M+L and L+L tonal-rhythmic models. To say that the low tone in the final position always has a falling contour will in no way be enough to cover all the peculiarities of that tone in the final position nor to show its effect on the rhythmic organisation of the entire disyllable, as the frequency interval of that tone and the speed of its formation remain outside the scope of research. Meanwhile, as is seen from the synthesis, these acoustic characteristics are functionally important not only for the rhythm of the model itself but also for contrasting other disyllables with the low tone in the final position. Orientation only to the falling tonal contour of the low tone in the final position is justified when this tone is contrasted with a different one, say, medium. Nevertheless, this criterion no longer works when two rhythmic models — H+L and M+L — with the low tone in the final position are set against each other. In this situation the interval and the speed of its formation rather than the falling nature of the low tone (it remains the same) become rhythmically significant, alongside other acoustic features involved in the differentiation of these models.
Studies of the tonal-rhythmic structure of Yoruba words (analysis and synthesis on the basis of disyllables) make it possible to suggest a general pattern of the rhythmic arrangement of tones and to formulate the main rules of sandhi governing tones in this unit of speech.
The interaction of tones in disyllables is based on three major rules. Two of them — the rule of register and register-contour oppositions -- operated in all the rhythmic models and their possible variants, while the third rule regulates the rhythm within only those models that fall under the rule of register-contour oppositions.
The specificity of the rhythmic organisation of bisyllabic Yoruba words consists in the fact that in one and the same model tones interact differently, depending on the context, for which reason one of the variants of the model can fall under the rule of register-contour oppositions, while another variant under that of register oppositions. For example, a variant of the "M+H" model with an even tonal contour in both syllables of a disyllable sounded in different registers — medium and high (_ -) — is governed by the rule of register oppositions, while its other variant falls under the rule of register-contour oppositions because the medium tone has an even contour in the medium register and the tone going after it has a rising contour (-/).
The rule of register oppositions regulates the interaction of tones in those tonal-rhythmic models which have tones with similar even tonal contours produced in different frequency bands, i.e., in different registers, which in their turn account for the opposition of tones in disyllables.
The rule of register-contour oppositions regulates the tonal-rhythmic models in which the tones combined have opposite contours and registers.
The third rule can conventionally be formulated as the rule of frequency correspondence or the equi-frequency correspondence of tones, which covers and regulates all the register-contour changes in the tones at all the segments of disyllables. The meaning of the main requirement of this rule by definition boils down to the equi-frequency correspondence between the beginning of the consequent tone in a disyllable and the end of the preceding one. Sometimes this equi-frequency correspondence can take place in the so called frequency correspondence zone, which allows for an insignificant frequency breakage (that is not perceived as a register contrast in Yoruba).
The phonological principle underlying the rule of the frequency correspondence of tones elucidates both the general mechanism of the interaction of tones in rhythmic models covered by the rule of register-contour oppositions and any particular manifestation of this general regularity. This ensures the necessary stability of tones and prevents their confusion in any contextual situation.

[I] J.M. Hombert. Perception of Bisyllabic Nouns in Yoruba. Studies in African Linguistics, 1976, Sup. 6.