From Natural Language to Drum Language
An Economical Encoding Procedure in Banda-Linda (Central African Republic)

France Cloarec-Heiss
CNRS-LLACAN, Paris

Introduction

As a means of communication, drum language is a unique system, as it *diverts* sound from a normal musical function into use for the transmission of information. This medium for sound transmission is in fact neither speech, nor an audible transposition of script (as in Morse code). What then is the nature of these signals, which the Banda produce by beating two wooden drums? This question cannot be answered without an understanding of one of the most widespread features of African languages, that is, *tone systems*. Indeed, Banda-Linda[1] is a tone language, in which every vowel must be spoken on a specific pitch: any change in pitch brings about a corresponding change in meaning. Since tone systems are not found in the languages which are more familiar to us, we may tend to underestimate how much information they can convey. These tones contribute to the 'mysteriousness' of drum language for those who are unacquainted with it; moreover, the role of tones in African languages has undoubtedly helped to give them an important place in colonial folklore and to make them the subject of an extensive bibliography[2].

The signal in a drum message is all that remains of speech after elimination of every audible feature provided by consonants and vowels, i.e., *the pitches assigned to each of the vowels*. It is surprising to find that communication is still possible after such a major amputation. This has nevertheless been proven to be the case by research[3] involving the recording of a set of fifteen different messages, each in two forms: one, the ordinary continuous transmission, and the other, in alternation with speech.[4] In the latter case, the drummer stopped to allow another informant who had no prior knowledge of the content of the message to decode it, step by step, into ordinary speech. The speed and

accuracy of the decoding observed in the course of this study provided spectacular proof of the effectiveness of this form of communication, which does not have an audience of specialists: any Banda-Linda speaker can understand drum messages and immediately transpose each statement into spoken Banda-Linda before the start of the next.

It will naturally be asked how communication can still be possible after such compression, and how all the problems arising from tonal homophony can be solved. We shall see that loss of information is compensated for by the sociolinguistic circumstances under which the system is used and by the structure of the message itself.

1. The principles of encoding

Encoding involves the reproduction of two parameters of ordinary speech: pitch and rhythm.

1.1 Pitch

In Banda-Linda, the tone of any vowel corresponds to one of three possible pitch levels. Tones have the same distinctive capacity as consonants and vowels. They suffice of themselves to produce both lexical and morphological contrasts. For example, in Banda-Linda as in many other languages in this part of Africa, most of the verb conjugation relies on differences in tone.

In all the messages we recorded, we found a one-to-one correspondence between the pitch of the drum beats and the pitch of the vowels in the matching spoken utterances.

1.2 Rhythm

Comparison of each drum message with the matching decoded statement in Banda-Linda shows the rhythmic curve of the message in perfect superposition to the spoken utterance.

- There are shorter intervals between the tones within a given lexical item than between two successive tones in different words, allowing the listener to distinguish the repetition of a disyllabic word from a single quadrisyllabic one with identical tones.
- The final tone of a noun or a phrase is struck more heavily and held slightly longer.

- Some tone sequences provide information on medial consonants: e.g., in a $C_1V_1C_2V_2$ structure, if V_1 and V_2 are separated by a liquid consonant in C_2 [l, r, ʋ][5], the drum beats on the vowels will be closer together than if the consonant were of another kind. This is how the drum expresses the loss of the first vowel often observed in spoken language.
- The tones of two immediately following vowels with the same tone are represented by a single lengthened beat of the drum.

Here we find, in a different form, the characteristic demarcative procedures of spoken language, whereby accentual and intonational features allow segmentation into units of meaning. This close relationship between spoken language and drum language proves that the drummer formulates his message in natural language as he transmits it. Drum language is thus derivative from speech, which it encodes by copying speech intonation and rhythm.

Our work also shows that the messages sent are not part of some preconcerted set, enabling the drummer to play a sort of 'score'. A comparison of the continuous and alternate forms of each message reveals important variations in both the form and the order of its melodic and rhythmic components.

The underlying linguistic content is, with the exception of very few features (see below), exactly what the general syntax of the language demands. The messages are composed of complete utterances which are grammatical in the spoken language. There are thus no abridgements such as the elimination of articles or grammatical markers characteristic of telegraphic language.

Such a system should potentially be capable of transmitting any utterance whatsoever. It nevertheless remains to be seen how the loss of information resulting from the elimination of all segmental phonemes can be made up for. An analysis of the content of the messages and the decoding process shows that, in practice, the theoretical capability of transposing any utterance whatsoever into drum language is unrealizable. This capability is severely limited by the social requirement of comprehensibility. While *anything can be said, not everything can be understood.* Drum language thus faces, on an even larger scale, the difficulty inherent in every communicative situation, whether relying on spoken language or not: the *dysymmetry* between source and receiver, i.e., things which are not ambiguous for the speaker may often be so for the hearer.

2. How decoding takes place

The decoding strategy, based on the *availability of clues*, i.e., the possibility of figuring out a full meaning when only part of its components is known, makes use of two given elements: 1) the extralinguistic situation and 2) the structure of the message.

2.1 The extralinguistic circumstances

The listener recovers the meaning of what he hears from information external to the message as much as from internal elements, i.e., the contents of the message itself.

2.1.1 Physical circumstances

The listeners (whether or not the message is intended for them) are able to situate its geographical origin. They can tell where the message is coming from and thereby eliminate a large number of possible addressees.

2.1.2 Sociolinguistic circumstances

Under ordinary conditions, drum language is used locally and involves a fairly limited number of participants. This is because the Banda live mostly in villages of no more than a few hundred inhabitants, and the actual social function of drum language restricts its use to a few emergency situations requiring one or more people to go to the place from which the message originates. These situations include traditional social events (such as birth, death, inauguration of a chief, inaugural ceremonies for new drums), accidents (e.g., on hunting expeditions or when someone gets lost in the bush), or any of a series of events having their origin in colonial times (such as the arrival of an administrative official, an egg collection for a subprefect, a cotton market, the arrival of a team of health visitors, or tax collection). The range of messages collected is proof in itself of the fact that there is no fixed set and that the system can *generate new information.*

2.2 The structure of the messages

Analysis of our set of fifteen messages shows them to be composed of *two types* of utterance whose order of occurrence is not codified, but left to the drummer's discretion.

- **Cue formulae** (CF), i.e., utterances or parts of utterances found in every message, around which the message is organized;

- **Informative units (IU),** which are peculiar to a given message.

The CF are the vectors for transmission of the IU, in that they make up most of the body of the message and are familiar, well-known melodies which have an immediate meaning for the listener. These formulae are *semantically*, though not formally, invariant: some of them can be realized in many different ways.

There are several features which are common to any utterance in the system, whether a CF or an IU:

- constant repetition: the time allotted to sending a message is fairly long (usually around five minutes), although the number of different utterances is small (around ten CFs and one or two IUs);
- further reinforcement of negative structures, which already have a reduplicated form in spoken language: in drum messages, the first syllable of the second occurrence of the verb is repeated;
- the informative nouns peculiar to the message are systematically reduplicated;
- the addressee(s) is/are particularized in the address formulae.

2.2.1 Cue formulae

There are some ten semantic types of CF, there may be more than one formula of each kind, and each formula has variants. The semantic types are as follows:

- The *opening* is always the same: a High [H] tone representing a spoken *wá*. Then comes the name of the addressee, e.g.,

wá à-màkònjì!	"Oh village chiefs!"
wá à-yī-kɨ̄ndɨ̄ kɨ̄ndɨ̄!	"Oh farmers!"

- There are two formulae which are used *to set off the IUs*, the first to mark the beginning:

ə́ pà də̀ yɛ̄	"and I say that..."

and the second to mark the end:

...də́ mə̄ nə́ pà ndə́ nə̀ kə̄ ʔɛ̄ kó	"...is the one I am telling you about"

- There are formulae which *confirm the message:*

mə̄ fórɔ̀ tə́mə̄ fófórɔ̀ nɛ̄	"I am not 'fooling myself', not wrong (about this)"
ávóró pɨ̄pɨ̄ də́ mə̄ sə́ pà kɛ̄	"it's the truth I'm telling you"

- There are formulae which *solicit the listener's attention* (by questioning, threatening, or insulting him):

yē ngbárə̀ ə̀ndɨ̀ ngbə́ngbárə̀ nē	"don't 'play the fools'" (pretend you all don't understand)"
yē jí nɔ̂?	"have you understood all that?"
yē jí gbə̀lá gbə̀lá nɔ̂?	"have you really understood every bit of it?"
ə̀njē sə́ zà bə̀ gálə́ kâkángà nə́ kɔ̄mándá!	"they're going to throw you in jail"

- There are calls to assemble:

yē ngbùrù gù gá mɔ̄!	"all of you get together right away"

- There are calls to come:

yē gù gá mə̄ kèrə̀ kèrə̀ kèrə̀!	"come to me as fast as you can!"
ə̀rə̀ yēkɔ̀ də́ bə̀ sə́ kángbərə̀ nə̀?	"what are you all up to (that's delaying you)?"

- There is a notice of the approaching end:

pà bàlē cémà mə̄ də́ mə̄ nə́ pà kē	"this is the last time I'll tell you"

- There is a *closing formula*, a [LL MM] tone sequence which can be repeated any number of times. This formula can correspond in spoken language to onomatopoeic expressions (*gbə̀là gūlū gbə̀là gūlū*) or evocative descriptions (*gə̀vɔ̀ gūlū gə̀vɔ̀ gūlū*).[6]

- The drummer has his *signature*:

wàkángà kə̀ ʔé ʔē kó	"this is Wakanga speaking to you"

It is striking from the cognitive standpoint to see that these formulae can be identified and understood despite all the variations in their acoustic form. Identification of the paradigm to which a formula belongs is made possible by the existence of a *typical minimal sequence* (a "signature") for each variant, which constitutes a *prototype* whose presence suffices to ensure a correct interpretation. Thus, the "come to me" formula may have any of the following forms, among others:

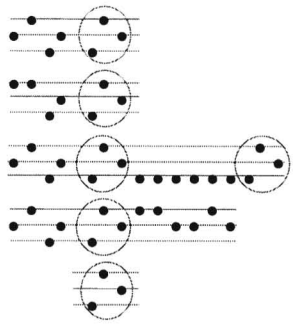
yē gûgū *gă mə̄*

ə́ gûgū *gă mə̄*

yē gûgū *gă mə̄* kèrə̀ kèrə̀ kèrə̀ *gă mə̄*

yē gûgū *gă mə̄* málə́ gōgō ná ʔē

gă mə̄



The prototype of the call to come is realized in the last example, the minimal form, by a [LHM] tone sequence, which can be played on the drum as part of any of a wide variety of sequences.

We shall see below that decoding into spoken language is carried out by one of *two different cognitive procedures,* according to whether the utterance involved is a CF or an IU. In the case of the CFs, which make up the greater part of the message, the listener, knowing the melody, thereby has immediate access to the meaning and can skip the transposition into spoken language.

2.2.2 Informative units

The indication of the addressee comes immediately after the opening formula. The message may be addressed to a group:

wá à-màkònjì! "oh village chiefs!"

The plural marker *à-*, represented by the first [L] tone after the opening, indicates that the message comes from a higher authority than the village chief and is addressed to the entire population administered through the system of village chiefs set up in colonial times. The hearer may therefore conclude that the message is coming from the seat of local government (subprefecture).

Wá à-yī-kīndī kīndī "oh farmers!" and *wá à-kōʃē kōʃē! wá à-yāʃē yāʃē!* "oh men! oh women!" indicate messages sent by the village chief to all people who are somewhere outside the village (working their farms, hunting, or gathering). The message emanates from the village.

The message may be addressed to an individual. The addressee's name will be followed by that of his father and, if necessary, that of his lineage.

1. yālénə́mə̄ māndá rə̀málè
"Yaleneme, *son of* Remale"

2. kòngbɔ̀ kòngbɔ̀ māndá léwā léwā
"Kongbo, *son of* Lewa"
3. yímàngā māndá kpádàkà málɔ́ gōgō làmbɔ̀là
"Yimanga, *son of* Kpadaka *of the lineage of* Lambela"
4. kíndɨ̀ngō nɔ́ cà?ɔ̀à málɔ́ gōgō nɔ́ ānéʃē gbàgɔ́rá
"Kindigo, *son of* Tshawa *of the lineage of* Gbagera"

When the name of the addressee has less than three syllables (as in example 2), it is systematically repeated.

The *paradigmatic set* of all IUs shows that the amount of new information transmitted in any message is very small. There are hardly more than *two or three semantically distinct units* in a given message, although each of these may have more than one form.

...sándákà sándákà màdāyè màdāyè nɔ́ wàyéwò wàyéwò...
ceremony x 2/installment/of/Wayewo x 2
[This is about] a ceremony to install Wayewo

ɨ̄pɨ̄ kùzū kùzū gōngérɔ̄ gōngérɔ̄
subject/death x 2/Gongere x 2
[This is] about the death of Gongere...

ɔ̀njē zú āméyā āméyā málɔ́ gōgō nɔ́ ?ē
they/*real*-give birth/*pl.*-twins x 2/ in/village/of/you
Twins have been born in your village

These units are usually bracketed by the CFs *and I say that* preceding and *that is what I am telling you* following.

Using these cues, the hearer can *break the message down* and identify the part whose meaning he has to recover, within the boundaries provided by recognizable preconcerted signals. Thus, in the address formulae, the name of an individual addressee is preceded by an invariant opening formula and followed by the name of his father and, if necessary, that of his lineage. These names are introduced in a codified way:

yímàngā **māndá** kpádàkà **málɔ́ gōgō** làmbɔ̀là
kíndɨ̀gō **nɔ́** cà?ɔ̀à **málɔ́ gōgō nɔ́ ānéʃē** gbàgɔ́rá

The words *māndá* "behind, after, heir" and *nɔ́* "connective element" introduce the father's name, while *málɔ́ gōgō* and *málɔ́ gōgō nɔ́ ānéʃē* (literally, "in/village" and "in/village/of/children") introduce the name of the lineage. These cues are used to delimit the beginning and end of the tone sequences corresponding to an "unlisted" unit and therefore requiring interpretation.

Unlike the CFs, the IUs cannot be decoded without being transposed into natural language; the melody does not give direct access to the meaning. With the help of the opening and closing formulae and all the other cues provided by

the message, the new information can be restored to spoken form. Understanding is facilitated by reference to all the contributing extralinguistic factors and by insistent repetition. The difficulties arising from the linear nature of language, aggravated here by the strictly acoustic nature of the transmission, can thus be obviated by "rerunning" the message often enough.

3. Recovery of the message as a cognitive process

Three operations are involved in the cognitive process of recovering the linguistic message: segmentation, selection, and reconstitution.

3.1 Segmentation

Rhythm, with its demarcative properties, plays a major role in allowing accurate decoding within this system. If the beats were produced at identical intervals and at a constant rate, they could only be perceived as a succession of individual entities, and it would be impossible to identify the segments containing elements requiring reconstitution. This would prevent the meaning from being recovered. Rhythm allows segmentation through its use of pauses, stresses, and holds, so that the segments requiring interpretation (words, phrases, or clauses) can be identified. The shape of the melody of each one of these segments, though the object of linear perception, is then decoded as a whole by means of processes of selection and reconstitution.

3.2 Paradigmatic selection

Having identified the relevant segments, the hearer can now decode different sets of tone sequences as wholes. Information theory, as developed by Shannon and Weaver (1975), accounts for all the observed features of this system by defining information as a measure of freedom of choice in selecting a message. Polysemy varies inversely with the number of theoretically possible tone combinations for a given sequence.

This number depends on two factors: one, the number of distinctive tone levels (which varies from language to language), is paradigmatic; the other, the number of syllables in the message, is syntagmatic, or sequential. Thus, the number of possible tone sequences in a given rhythmic unit is equal to the n^{th} power of the number of distinctive tone levels in the language, where n is the number of syllabic tones in the unit. When the number of syllables on the syntagmatic axis is increased, the number of possible combinations increases

exponentially and thereby reduces the number of interpretations for any sequence.

Thus, in a two-tone language, a disyllabic CVCV item can have four different tone patterns, LL, HH, LH, or HL, for the drum to reproduce. Whichever one is produced eliminates the other three possibilities, i.e., three-quarters of all the possible sequences.

In a three-tone language like Banda, any disyllabic CVCV sequence may have one of nine possible tone patterns (LL, LM, LH, ML, MM, MH, HL, HM, HH). Any sequence of two tones in this language will thus eliminate, not just three-quarters, but eight-ninths of the possibilities. The same acoustic material will thus transmit nearly 14% more information (8/9 – 3/4) in this language than in a two-tone language. Adding a single tone at a given point in a rhythmic unit eliminates two-thirds of the remaining possible interpretations in a three-tone language, but only half in a two-tone language.

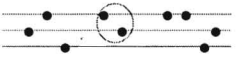
According to the number of distinctive pitch levels, tone languages thus have a given paradigmatic "depth", i.e., a relative extent of choice and level of polysemy. This may be why some drum languages deriving from spoken languages with two tones compensate for their paradigmatic shallowness by making use of *periphrasis* on the syntagmatic axis. Only by lengthening the sequences can the sender reduce the level of indeterminacy. As Arom and Cloarec-Heiss (1976) have already shown, this disambiguating procedure, which was formerly thought to be universal in African drum languages,[7] is entirely absent from the Banda system.

3.3 Reconstitution

As the message proceeds, the listener narrows his choice from the paradigm of possible alternatives by seeking the semantic compatibility of each new tone with the preceding ones.

Reconstitution is the process by which the linguistic message is ultimately recovered. The mental process which accompanies the perception of a tone sequence is not an additive one; rather, it works retroactively, with each new tone reorganizing the message as a whole, cf. the following two nearly homophonous tone formulae:

mə́ fórɔ̀ tə́-mə́ fófórɔ̀ nē



I/*real*+fool/myself/*neg.*+fool/not
I am not fooling myself

yē ngbárə̀ ə̀ndɨ̀ ngbə́ngbárə̀ nē

you all/*real*+seem/fool/*neg.*+seem/not
don't play the fools

These examples show how perception of the formulae is nonlinear with reconstitution taking place by "retroidentification". Indeed, they differ only in their fourth and fifth tones. The HM pattern in one and the LL in the other allow the listener not only to interpret the two tones as either the term for "myself" or the one for "fool", but at the same time to identify *retroactively* the words corresponding to the second and third tones, which are the same in both formulae, as "to fool" in one case and "to seem" in the other, according to considerations of semantic incompatibility.

Since the name of an addressee can, as we have seen, be particularized by adding the name of his father and, when necessary, that of his lineage, there will be a sequence of not just two to four tones, but six (at the very least) to nine, when the three names are given as described above. A sequence of six tones (the smallest possible number of information-bearing tones, interspersed with cue sequences) will then be one of 729 (3^6) possibilities, and nine, one of 19,683 (3^9). The message thereby contains ample information to identify any individual in a small community.

We thus find, not sequential perception of pitch, but retroactive identification of segments by reconstitution of the whole through a process of elimination. This process assumes storage of data in memory until a threshold is reached where understanding occurs through reconstitution. This proves that the melodic pattern, though linear for perception, is decoded *holistically*.

The process of decoding of IUs, unlike CFs, is thus symmetrical to encoding: the hearer simply reverses the sender's procedure. He seeks the natural language equivalent of the tones he hears, i.e., he recovers the words underlying the audible message.

Conclusion

While drum communication uses a peculiar form of audible material, it nevertheless has features which can be found in any kind of linguistic output. One is the central role of segmentation, where the rhythm of drum language replaces prosodic phenomena such as stress and intonation in spoken language. Others are the importance of extralinguistic factors in the process of disambiguation and an organizational structure paralleling the topic/comment structure of speech. The listener operates with three pairs of *poles of reference*:

- the linguistic information contained in the message *vs.* the extralinguistic information provided by the circumstances of reception;
- new information to be recovered (from the IUs) *vs.* preconcerted material (from the CFs);
- within the IUs, new material to be recovered (the information-bearing parts) *vs.* the accompanying predefined cues.

Recovery of the linguistic message by retroidentification and reconstitution are sufficient proof of the holistic nature of perception.

The path *from perception to meaning* is thus more or less complex accordingly as the material played by the drums is familiar (the CFs) or new (the IUs). In the former case, meaning is *directly* accessible from a predefined output; in the latter, a *recovery procedure* is required. Even when the audible material is the same, the *two cognitive processes* involved are different in kind. Indeed, the frequency of the CFs has given rise to a sort of 'drum vocabulary'. When a given tone sequence is sent out at a specific rhythm, the meaning is immediately apparent. This means that an inventory of signs has been constituted. In the case of the IUs on the other hand, the hearer is obliged to recover the meaning part by part as the message is sent. To do this, he must compensate for the informational shallowness of the message by making adequate use of strategies relying on whatever linguistic and extralinguistic cues are available.

This system of communication thus retains close ties to the spoken language, even after loosening them by eliminating the greater part of the audible components of speech. A study of the strategies for compensating for this loss of information reveals a clear-cut view of the cognitive processes at work in any communicative situation, which can be taken as characteristic of human perception and understanding in general.

Notes

1. According to Greenberg's classification, Banda-Linda is a Niger-Congo language (Adamawa-Eastern branch, now called ubangian: IA6b2. It has been described in Cloarec-Heiss (1986)
2. Sebeok and Umiker-Sebeok's 1976 synthetic study lists all the essential works up to that time.
3. This research was conducted in 1973 by Simha Arom (LACITO-CNRS) and the author in Ippy, Central African Republic.
4. Recordings of each version can be heard on the cassette accompanying Arom and Cloarec-Heiss (1976).
5. ʊ represents a sound called a "labial flap" which is unknown outside Central Africa.

6. gə̀ʊɔ̀ denotes a wreath placed on the roof of a traditional house when it is completed; gūlū is onomatopoeic.
7. See Alexandre (1969), Betz (1898), Carrington (1949), Heepe (1920), Nekes (1912), Rialland (1974), Thilenius et *al.* (1916), Verbeken (1920).

References

Alexandre, Pierre. 1969. "Langages tambourinés, une écriture sonore?". *Semiotica*, I: 3, 272-281.

Arom, Simha and Cloarec-Heiss, France. 1976. "Le langage tambouriné des Banda-Linda (RCA)". In L. Bouquiaux (ed.), *Théories et méthodes en linguistique africaine*. Paris, SELAF 54-55, pp 113-169, 1 cassette.

Betz, R. 1898. "Die Trommelsprache der Duala". *Mitteilungen von Forschungs-reisenden und Gelehrten aus den Deutschen Schutzgebieten*. XI, 1-86.

Carrington; John F. 1949. *The Talking Drums of Africa*. London: The Carey Kingstate Press.

Cloarec-Heiss, France. 1986. *Dynamique et équilibre d'une syntaxe: le banda-linda de Centrafrique*. Paris / Cambridge: Peeters-SELAF and C.U.P.

Heepe, Marius. 1920. "Die Trommelsprache der Jaunde in Kamerun". *Zeitschrift für Kolonial Sprachen*, X, 43-60.

Nekes, Hermann. 1912. "Trommelsprache und Fernruf bei den Jaunde und Duala in SüdKamerun". *Mitteilungen des Seminar für orientalische Sprachen*, XV: 3, 69-83.

Rialland, Annie. 1974. "Les langages instrumentaux sifflés ou criés en Afrique". *La Linguistique*, X: 2, 105-121.

Sebeok, Thomas A. and Umiker-Sebeok, Donna J. (eds). 1976. *Speech Surrogates: Drum and Whistle Systems* (2 volumes). The Hague / Paris: Mouton.

Shannon, Claude. and Weaver, Warren. 1975, *Théorie mathématique de la communication*. Paris: Retz - CEPL, Les Classiques des Sciences Humaines. Traduction de *The Mathematical Theory of communication*. 1949, The Board of Trustees of the University of Illinois.

Thilenius, G. *et al.* 1916. "Die Trommelsprache in Afrika und in der Südsee". *Vox*, 26, 179-208.

Verbeken, A. 1920. "Le tambour téléphoné des indigènes de l'Afrique centrale". In T. A. Sebeok and D. J. Umiker-Sebeok (eds), *Speech Surrogates: Drum and Whistle Systems* 1, La Haye / Paris: Mouton, 336-365.