

# John Benjamins Publishing Company



This is a contribution from *Diachronica* 22:1  
© 2005. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: [www.copyright.com](http://www.copyright.com)).

Please contact [rights@benjamins.nl](mailto:rights@benjamins.nl) or consult our website: [www.benjamins.com](http://www.benjamins.com)

Tables of Contents, abstracts and guidelines are available at [www.benjamins.com](http://www.benjamins.com)

# Phonological irregularities, reconstruction and cultural vocabulary

The names of fish in the Bantu languages of the Northwest (Gabon) \*

Patrick Mouguiama-Daouda

CNRS, Lyon

This study aims to distinguish irregularities due to borrowing from those due to lexical diffusion and those due to expressivity. The method adopted proposes the comparison of virtual reconstructions as the basis for reconstruction. Virtual reconstructions are obtained by applying in reverse the phonological rules set up for the fundamental vocabulary to the cultural vocabulary. From that point it becomes possible to establish chronological stages for roots or words and assign an order to them. The method is illustrated by a study of names of fish in the Bantu languages of Gabon. We show migration currents from the east towards the west, and the comparison of virtual reconstructions reveals that the ichthyological culture is relatively recent and on the whole does not go back to the Proto-Bantu period.

**Keywords:** Borrowing, ethnoichthyology, expressivity, irregularity, lexical diffusion, Proto-Bantu, root, virtual reconstruction

## Introduction

Irregularities in phonological reconstructions have always been a problem for those working in comparative linguistics. Indeed these linguists do not even always look at the process of change in the same way. The Neogrammarians held

---

\* We wish to thank Madame Louise Fontaney who translated this article, originally written in French, into English. Her queries and comments have helped us in the organisation of our arguments. Naturally, we alone are responsible for the proposals and conclusions presented.

a radical position: they allowed no exceptions other than those due to borrowing and analogy. As the effects of analogy on morphological paradigms can become regular, the chief cause of irregularities was held to be borrowing. Borrowing between dialects or closely related languages can be a stumbling block to systematic comparative reconstruction. At the beginning of the 20th century the question of borrowing again gave rise to much controversy, especially between Boas and Sapir (for a summary, see Darnell & Sherzer 1971). The former thought that, beyond a certain time depth, similarities due to borrowing could not be distinguished from those inherited from the parent language. The latter, on the other hand, held that it was possible, by a fine-grained analysis of the morphological features, to distinguish the two. Sapir, like most early comparativists, considered morphology to be very stable and the deciding factor in establishing genetic relationships. Other investigators, and above all Greenberg (1955, 1963), have maintained that basic vocabulary was equally important in establishing genetic relationships, whereas specialised vocabulary is less stable and more subject to borrowing.<sup>1</sup> Loanwords are thus the major stumbling block in comparative linguistics, as they are the source of phonological irregularities. More recently the problem of irregularities in phonetic change has been taken up again by various writers (Wang 1969; Chen & Wang 1975) who suggest lexical diffusion as another possible cause, but they recognise that it is not always easy to distinguish its effects from those of borrowing.

It is the purpose of this article to present a reliable method for detecting borrowings, distinguishing their effects from those of lexical diffusion and expressivity. We show that borrowings give rise to a chain of irregularities in a group of dialects or closely related languages, whereas the effects of lexical diffusion and expressivity are more local. The data come from the Bantu languages, their general distribution making it possible to test the method of comparison we envisage: they are divided into areas where words and grammatical features spread readily from one language or one dialect to another. This method was first used by Hombert (1988) for the names of mammals. Here, we give the results of its application to the ichthyological lexicon, where the data are equally good for all the languages. This is as true for the linguistic

---

1. The distinction between basic vocabulary (parts of the body, pronouns, verbs of perception, etc.) and cultural or specialized vocabulary (fauna, flora, economy, technology, etc.) is fundamental for our argument. Although our study deals with only a small part of the fauna, the distinction between the two classes of vocabulary coincides with the definitions of Morris Swadesh.

aspect as for the scientific identification of species. Moreover, the results obtained from the names of fish show the fecundity of the method outstandingly well; they also give rise to hypotheses for protohistory which are sometimes quite new.

This article consists of three parts. In the first, we examine the main causes of irregularity, which leads us to conclude that comparative linguistics has not yet found a satisfactory methodological framework for systematically distinguishing irregularities due to borrowing from those due to internal factors (lexical diffusion and expressivity). We will suggest strategies for characterizing the different effects of the three potential causes of irregularities. In the second, we introduce and develop the idea of ‘virtual reconstruction’; this concept will serve as the basis for setting up comparative series. A virtual reconstruction is the word, not as it exists in the language, but as it appears when the rules of historical phonology are applied to it backwards. This idea makes it possible to detect borrowed words for which an etymon has been provided. These words appear to be regular when looked at in a given dialect, but if they are examined in several dialects of the same language, the virtual reconstructions are seen to be different. In fact, these differences reveal diverse integration strategies. In the third, we present the results of that analysis. That is, the distribution patterns of inherited roots are examined and compared with those of the terms acquired by spreading. The general tendencies that emerge are illustrated by the conclusions arrived at from the study of the names of fish in the languages of Gabon.

The reader can consult the appendix to see how the method has been applied systematically to a fragment of cultural vocabulary, in a group of distinctly characterized languages. It is a study of a particular case illustrating a method that we believe is of much wider and more general interest.

## 1. The causes of phonological irregularities

### 1.1 Borrowing

As a start, let us take a new look at a number of problems encountered with the classical comparative method. Banal as they are, they nonetheless make it easier to define and situate our own procedure. Comparativists held borrowing and analogy to be the sole causes of irregularity in phonological change. As analogy involves morphophonological features, it gives rise to irregulari-

ties that fade gradually and its effects are less striking than those of borrowing. So phonological irregularity was for a long time attributed primarily to borrowing.

When the foreign origin of a loanword is still perceptible there is no problem. In standard French, for example, words ending in a velar nasal are mostly from English (*parking, standing, living*, etc.). That they are foreign words is obvious as the velar nasal does not have phonological status in French and there are few words containing one. On the other hand, when all the segments of a borrowed word do exist in the language, these words call for more careful examination. Let us take the case of the word *alcool* in French. The sounds *a, l, k, ɔ*, that constitute it are all phonemes of the language – it is the sequence *lk* that betrays its Semitic origin.<sup>2</sup> A fine-grained synchronic analysis makes it possible to detect words that are well yet incompletely integrated into the borrowing language.

In the French word *gnou* “gnu” (“*Connochaetes sp.*”), the sounds *g, n, u* are all phonemes of the language; it is the sequence *g-n* that betrays its Khoisan origin. So the first routine procedure for identifying a loanword is to look for segments or sequences of segments that do not exist in the borrowing language.

It is especially when loanwords disrupt the regular pattern of correspondences that they constitute a trap for comparativists. Let us compare for example a series of French/Spanish word pairs: *cheval/caballo, château/castillo, champ/campo, cheveu/cabello*. The initial correspondence *ʃ/k* is well established and explained by palatalisation, a characteristic rule in the evolution of French. But we find the pair *champagne/champaña*, suggesting a correspondence *ʃ/tʃ* for the initial consonant, contrary to the regular series. Fortunately, historical data are available to show that the Spanish word was borrowed from French after the process of palatalisation was complete. Without this information, the linguist would be tempted to reconstruct for the Spanish a double reflex for initial \*k: *k* and a less regular *tʃ*. In such a situation, which is after all com-

---

2. From the 13th century at least, in the passage from vulgar Latin to French, there is a rule changing *l* to *u* that blocks the sequence *lk* (*aucun* < *aliquunus*, 1209). The sequence in words like “*calcul*” (1484), “*calcaire*” (1751 – but “*chaux* < *calx, calcis*” 1135) are of later date; as part the technical lexicon, they do not belong to the evolution from vulgar Latin. There is certainly a possible sequence *lk* now at word juncture (e.g. “*belle cousine*”), but this is the result of the optional dropping of *a*, so there is a structural vowel, which is not the case with words of Semitic origin.

mon enough, the linguist chooses borrowing to explain the irregularity, even without historical proof, in conformity with the Neogrammarian principle.

Let us now take the case of the words for “rice” in two Gabonese languages,<sup>3</sup> *òrésì* in Mpongwe and *ùrèsì* in Punu. The segments of these words are regular phonemes in each language. A comparison of the second consonant (+*resi*, +*resi*) yields the correspondence *s/s*, whereas regular evolution, as attested by the basic vocabulary, would yield a *z* in Mpongwe as C2, corresponding to an *s* in Punu. For example, in the words for “twin” we find the pair Mpongwe/Punu *ìbázà / òdivásá*, for “raw”, *mbézò / mbísú*, etc. It is Proto-Bantu, as reconstituted from the comparison between Punu, Mpongwe and all the other Bantu languages, that establishes the regular correspondences: if we take \**pácá* “twin”, \**bécú* “raw” in the protolanguage, we see that the consonant \**c* as C2 calls for the alternation *z/s* for the Mpongwe/Punu word pairs. The conclusion is that the words showing the alternation *s/s* in the same context have not come down from Proto-Bantu. In fact, we know that *òrésì/ùrèsì* are borrowed from English “rice”, as are the words for “glass” (*èlásì/yílásì*) and “plate” (*èpélè/yìpélè*). These words were introduced by English explorers when the first trading stations were set up and during the Atlantic trade on the Gabon coast (16th to 19th centuries). When words are borrowed, they may disrupt the regularity of inherited correspondences, thus constituting a false series. When such a breaking of a diachronic rule is observed, it is necessary to look for the etymon that a segment in the language could go back to. In Mpongwe, one will never find an *s* as C2 going back to a Proto-Bantu etymon. So, the second routine procedure for identifying a loanword is to check for compatibility with the rules of sound change.

After what time lapse is there still a chance of detecting a loan? Some loanwords are very quickly assimilated; others continue to show traces of their foreign character long after they were borrowed. For example, the word *alcool* was borrowed into French from Arabic some five centuries ago, but traces of its Semitic origin remain. In fact, the main parameter is linguistic proximity: the greater the difference between the systems of the two languages concerned, the longer the period of integration is. Influences between dialects have been considered by the Neogrammarians to be the chief cause of disturbance in the regularity of correspondences.

---

3. The languages of Gabon referred to in this study are presented in the appendix.

## 1.2 Lexical diffusion

If we leave aside the effect of borrowing between dialects, we should expect phonetic changes to be absolutely regular. But might we not, in this case, risk excluding from the language we are reconstructing words that in fact belong to it? This is a legitimate question: between the 19th century and the first half of the 20th, dialectologists (Schuchardt 1885; Sturtevant 1907, 1940, 1947; Gauchat 1905) were already claiming that such a position was exaggerated. They questioned the absolute validity of this fundamental principle of comparative linguistics. They claimed that phonetic changes were not regular in the sense that they affected all the words of the vocabulary simultaneously; according to their formula “each word has its own history”, social motivation being the explanation of change. According to this principle, regularity is a progressive process since changes take place word by word, just as in the case of analogical changes. Empirical evidence for the sociolinguistic motivation of linguistic change began to accumulate, in particular with Sapir (1921). Sapir showed that in certain dialects of English the change from *ū* to *ǔ* took place progressively, giving rise to doublets. Nonetheless, the Neogrammarian model remained dominant until the end of the sixties. It was not until the beginning of the seventies that the ideas put forward by opponents of the Neogrammarians were integrated into a coherent framework and regarded as a theoretical model. Its conception is due to Weinreich, Labov & Herzog (1968), and Labov (1963, 1972). Thanks to the quantity of data and the effort to systematize, the work of these linguists can be considered to have shown that another view of phonetic change was possible. With his study of Martha’s Vineyard, Labov (1963, 1972) defined the modalities of setting up phonetic changes in terms that have numerous analogies to what we understand today by lexical diffusion. He showed that the centralisation of *a* ([a]>[ə] in the diphthongs [ay] and [aw]) started in a small number of words before spreading to others, according to various contexts, phonetic, morphological, morphophonological, semantic, and social. Between the beginning of the process and its generalisation, the change shows considerable irregularity, but this generalisation is nonetheless relatively quick. And in this there is a fundamental difference from lexical diffusion as defined by Wang. Indeed, Wang (1969), the most important theorist for the latter explanatory model of phonetic change, considers the period during which the rule change becomes general relatively long. In the course of time irregularities become blurred and disappear – unless there is a breakdown in the process and/or another rule intervenes.

Table 1. After Chen &amp; Wang (1975:258)

	A	B	C	D
	-vcd	-vcd	+vcd	+vcd
	-asp	+asp	-son	+son
1a	8	4	1	–
1b	1	–	5	10
2a	5	4	1	–
2b	14	–	56	69
3a	222	78	11	5
3b	6	3	61	46
4a	5	–	–	–

Since the time factor plays an important part in the evolution of sounds, phonemic splits are not necessarily explicable by imperceptible phonetic or morphological conditioning or by analogy. This may be explained by the gradual propagation in the lexicon of changes in progress. The study of the tones of Cháozhōu, a modern Chinese dialect where tone type 3 of Middle Chinese has split and is now represented in two tone types, apparently without any conditioning, is given as basic evidence for lexical diffusion. The table sums up this evolution.

In fact, in this language, the four tones of Middle Chinese (1, 2, 3, 4) split according to the initial consonant. When this consonant was voiceless, they have become 1a, 2a, 3a, 4a and when it was voiced they have become 1b, 2b, 3b, 4b, respectively. However, in the case of tone type 3 this evolution has not been symmetrical. For voiceless initial consonants (columns 1 and 2) the evolution is as predicted, as type 3a in Cháozhōu comes from Middle Chinese type 3 (222 for the unaspirated and 78 for aspirated). However, with voiced initials (columns 3 and 4), type 3b coming duly from Middle Chinese type 3 is under-represented (61 for voiceless, 46 for voiced), while type 2b is, on the contrary, overrepresented. The suggested explanation of this breakdown in the predicted system is that many words of Middle Chinese type 3 have developed into 2b in Cháozhōu. In precisely the same phonetic context, there seem to have been two different evolutions of one and the same original tone type. Having excluded borrowing and analogy as the cause of this double evolution, Chen & Wang have proposed lexical diffusion as an explanation: the overrepresentation of type 2b would be due to a gradual evolution of type 3b, and to their thinking this is a flagrant contradiction of the Neogrammarian hypothesis.



Likewise, many Middle Chinese homonyms have become differentiated in modern Chinese dialects, which would show that different evolutions are not necessarily due to phonetic conditioning (Aleshire & Streeter 1970, quoted by Chen & Wang 1975:260).<sup>4</sup>

All this might give the impression that lexical diffusion has been unanimously accepted and the Neogrammarian hypothesis of regularity thoroughly shaken. However, certain foundations of the theory have been challenged. For example, the Chinese dialects have been re-examined and it appears that the evolution of the tones is to be explained by dialect borrowing and the influence of literary Chinese (Labov 1994:451). Another study has shown that the evolution of homonyms is compatible with a Neogrammarian account. Lastly, Kiparsky (2003) has shown, in the framework of lexical phonology and radical underspecification, that lexical diffusion could be described in the same way as analogy, the only difference being that in this case the paradigms involved are not morphological but phonetic or phonological.

When so much has been called into question, where does lexical diffusion stand? Another study carried out by Labov (1989, 1994:421–439) among a hundred English speakers in Philadelphia provides additional evidence, apparently irrefutable, of lexical diffusion. While *a*, like most vowels in the same context, is lax before voiced stops (*sad* for example), there are three words, *mad*, *bad*, *glad*, in which a tensing rule affects the vowel. As this process is very old and has never spread to other phonetic contexts, we have here strong evidence of a phonetic change taking place very gradually, over a long period of time.

It should be remembered that the various cases of lexical diffusion dealt with by different theorists cover several types of evolution, ranging from a simple change in the phonetic quality of a given phoneme (not affecting the phonemic systems), through shifts in phoneme distribution (affecting individual words), to a complete phonemic change (affecting the system).

Finally, above and beyond the conflict between different schools of thought, it appears that both mechanisms are needed for describing phonetic

---

4. Examples are reported from other language families: we will mention two by way of example. In Swedish, it has been observed that the dropping of final *d* affects far fewer words than fifty years ago. In the same context *d* may either appear as *d* or disappear (Janson 1973, quoted by Chen & Wang 1975:262–263). In English, it has been shown that the process involving the change of the place of the accent of a verb to produce a noun has spread very slowly: only 11% of such words have been affected by the process since it has been in operation (Sherman 1973, quoted by Chen & Wang 1975:261–262).

change. So we can accept that, in certain situations, the effects of lexical diffusion may give rise to irregularities. Fundamentally, the model does not rule out the reconstruction of roots based on series manifesting a certain degree of irregularity. How then are we to distinguish irregularities due to lexical diffusion from those due to borrowing? Wang conceded that in many cases this was impossible, since the ways in which phonetic changes come about are the same in both cases: “It is not always easy to determine if a given situation is the result of dialect borrowing or if it is due to lexical diffusion” (Wang 1969: 45).

This is true to a certain extent. To take an example from Gabon, for the Mpongwe *ntfínà* “blood”, which goes back to Proto-Bantu \*gìdá, a process of fricativization is involved. Now this type of evolution, attested only in this word in Mpongwe, is characteristic of the languages of South Gabon (B 40, B 50, B 60, B 70) and the greater part of zone H.<sup>5</sup> Several scenarios are possible to explain this highly marginal evolution in Mpongwe – one case out of nearly 400 etymologies established.

It could be a Mpongwe process in its initial phase. This would confirm that time is a parameter of fundamental importance in the origin of phonetic change: as Guthrie (1967–1971) already noted for this marginal case, one must recognize that the process is slow, or has stopped – in more than thirty years, the change is far from beginning to be generalised. It could be a case of borrowing from a language in which the process was regular – the classic borrowing scenario. Or it could be a borrowing from a language in which the process was irregular, that is, in its initial phase. This is the scenario of the conjunction of two causes: a change which is borrowed when the process is only beginning in the source language. Or it may be a borrowing from a language in which the irregularity is already due to borrowing. In this last scenario, we already have the situation of successive loans that lexical diffusion alone cannot explain.

To solve the problem, we can adopt a guiding principle, the heuristic value of which will become clear below. This position is based on a rational explanation of the cause of phonetic change. While lexical diffusion offers an explanation of the ‘mechanism’ of change, incorporating the time factor, it does not explain the ‘cause’. It is the sociolinguists who propose the hypothesis which is most appropriate to the case: it is the variability among speakers that is the driving force in linguistic change, at least as far as internal factors are concerned. In

---

5. This zone consists of most languages in the south of the Congo, Lower Congo (DRC) and the north of Angola.

fact, in all languages, every speaker produces and perceives an infinite variety of sounds, the frequency values of which are close to those of the conventional value of the distinctive sound they are associated with, but nonetheless different from it. Phonological changes become possible when certain divergent values are repeated to the point of being shared by several speakers. Socio-cultural factors (prestige, fashion, demographic conditions) will lead to the dominance of one of the divergent forms at the expense of the others. If this principle is accepted, it becomes possible to distinguish the effects of borrowing from those of lexical diffusion, if not systematically, at least in certain cases. If an item is irregular in several dialects of the same language, it is highly probable that it is due to successive borrowings. There is little likelihood that variability alone should be responsible for the same item being irregular in several languages.

It is not then a question of whether irregularity in a given item in a given language is due to lexical diffusion or to borrowing. When the word is part of the basic lexicon, as was the case with Mpongwe *ntfina* “blood”, it is very difficult, if not impossible, to say. The pertinent question is this: can inherited words in related languages show an irregularity recurrently? Lexical diffusion can be the starting point, and only the starting point, of a chain of irregularities appearing in dialects of the source language – but it cannot be the only cause.

If, on the other hand, there is a high proportion of irregularities in a particular lexical domain, this is probably due to borrowing, as there is nothing in the formulation of lexical diffusion to restrict its effects to a particular lexical field. For example, the words *alcool*, *alchimie* in French, derived from Arabic and irregular in French, have the same history and are restricted to a particular lexical domain. The words for “kitchen”, “glass”, “plate”, “rice” are irregular in Gabonese languages: restricted to a particular lexical field, they all have the same history, that of the first contacts of Gabonese populations with the West.

Lastly, irregularity arising from lexical diffusion is limited in three respects: first, because it characterizes only the initial stage of the process; second, because it only persists if there is a conflict between two phonetic changes; and third, in so far as the appeal to it as the cause of irregularity in an item in several dialects of the same language is incompatible with the principle of variability as a driving force of change.

### 1.3 Expressivity

Irregularities due to expressivity also arise from internal variability, but they can be motivated by socio-cultural considerations. We therefore distinguish

them from those occasioned by lexical diffusion which are determined by physiological, acoustic and perceptual constraints.

Coupez (1975) showed that expressivity was also responsible for lexical variation in Bantu. For instance, in Rwanda one finds *-mààr/tààr* “to gather nectar” *-bátik/mátik* “to stick (trans.)” *-mécùr/kécùr* “to munch” – etc. These are free variants, in so far as the speaker can use either form without difference of meaning. Coupez considers that it is legitimate to group together such forms as phonetic variants due to expressivity in the case of very small differences.

This phenomenon can include taboos and prohibitions, which are known to favour a linguistic strategy of avoidance. By the very nature of taboo, it is very difficult to classify the effects aimed at, and variants may be of totally different kinds, ranging from modifications of sounds to replacement of a term, and including syllable inversion or substitution. In Mpongwe there is a word *isiki* designating “a reincarnated person of very small stature”. Since this comes from Proto-Bantu \**kéci* (or \**kéti*) “spirit”, it is clear that a phenomenon of avoidance has led to the metathesis. Similarly, the term for the “electric silurid”, “*Malapterurus electricus*”, is highly irregular and could well be the result of avoidance strategies, since this fish is not eaten by men as it is thought to weaken the male member. In certain Bantu languages, irregularities found in the terms employed to designate the leopard can also be explained by avoidance strategies, as this animal plays an important role in certain initiation brotherhoods. Again, there are masculine rites in Gabon the name of which women must not pronounce, so they use paraphrases that can be translated as “the men’s affairs”.<sup>6</sup>

From these few examples, we see that it is not only in the lexical field of “mysteries” or of “sorcery” that avoidance strategies are encountered, although they are certainly much less frequent in the lexical field of fauna.

From a historical point of view, it is above all variants involving the substitution of sounds or syllable inversion that disrupt comparative series. Several scenarios for the evolution of such variants seem possible: (i) phonetic or phonological differences can increase while meaning remains stable; (ii) phonetic change may be accompanied by a semantic evolution, in which case the common origin of the variants may not be perceived; (iii) meaning evolves while the phonemes remain relatively stable. Phenomena of this kind are to

---

6. This avoidance process is attested among the Okande, an ethnic group of central Gabon. The women do not use the word *mwidi*, which designates a masculine rite, the social importance of which has been pointed out in other ethnic groups in Gabon.

be found at present in languages of the south of zone B 30. For example, the older generation preserves a distinction between *modyenge* “*Barbus batesii*”<sup>7</sup> and *modyɔngɔ* “*Barbus holotaenia*” whereas the younger generation tends to use both terms for one and the same species. In fact, the process can be cyclic. One may have a period when there is a single term A to designate a referent X. At a second stage, a term B appears, to designate a referent Y resembling X. At a third stage, A and B are used as variants to designate a single referent. Or it may be that only one of the terms (A or B) remains, to designate both X and Y – or simply X or Y, in which case a semantic distinction is lost.

Therefore, it would seem that expressivity is one of the possible sources of Guthrie’s ‘osculant series’ in Proto-Bantu: the variants \*kúdò/\*kúdù “tortoise”, \*cádàkù/\*cídàkò/\*tídàkù “ant sp.”, \*gùbó/\*gùbù “hippopotamus”, etc. may have undergone a type of evolution similar to that of the “barbus” terms.

The impact of lexical variation in the reconstruction of protolanguages could be more widespread and involve other language families, as Coupez (1975:201) writes:

The hypothesis of lexical variation should be tested in other proto-languages. Swadesh, in 1970, pointed out some troublesome alternations in the reconstructions of several proto-languages. In proto-Indo-European he attributes alternations between consonants differentiated by a single phonetic feature to paradigms having a semantic value, but there is much obscurity still. He reports also that in Altaic Ramstedt resorts to dialectical interference to elucidate otherwise inexplicable alternations.

How then are we to distinguish the effects of expressivity from those of lexical diffusion or of borrowing between dialects? We will consider that the phonetic variants due to expressivity are characterized by very small differences. We therefore agree with Coupez, who is reluctant to group together forms involving wide divergences. Meaning can also help in our decision: often when a language borrows terms from a neighbouring group, when it already has a term to designate the same entity, we have a different sort of doublet: words constituted of totally different segments, referring to the same entity, in short, synonyms.

Generally speaking, there is no reason why variants should be restricted to a particular lexical field or a specific grammatical category. It is true that certain animals may be particularly liable to be given avoidance names, but nothing

---

7. These two species are small freshwater fish called *goujon* in local French (English *gudgeon*).

justifies the presence of a high proportion of variants in a specific lexical field. In such a case, there is good reason to suspect borrowing. The high number of variants in Guthrie in animal names would seem to be more compatible with borrowing between dialects than with expressivity.

Lastly, one may have recourse to expressivity as a possible cause of irregularity if the word in question is highly localized. It is not likely that in several languages the same animals should be the subject of taboos, the designations of which are all irregular. In the lexical field of “mysteries” and “sorcery” avoidance terms are possible without being necessarily widespread.

In any case, one should always put forward a rational explanation to justify avoidance as a motivation, if necessary appealing to semantic and cultural universals.

## 2. Virtual reconstructions as a basis for reconstruction

### 2.1 Virtual reconstructions

We have examined the fundamental causes of phonological irregularity: borrowing, lexical diffusion and expressivity. We have tried to show that the effects of each could be different. Borrowing tends to be restricted to specific lexical fields; it can spread through different dialects of one language or through languages of the same family, and thus constitute a chain of irregularities. Lexical diffusion is not restricted to a particular lexical field and cannot account for recurrent irregularities in a network of dialects or languages. As expressivity characterizes words socio-culturally marked, before proposing it as the source of an irregularity it should first be shown that there is motivation for avoidance. Furthermore, variants due to expressivity are not restricted to a particular lexical field – this being, as already said, one of the characteristics of borrowing.

Distinguishing these three causes implies making different choices when setting up comparative series with their resulting reconstructions. If the irregularity is due to borrowing, the series is irregular and should not be used for proposing a root in the protolanguage in question. On the other hand, if the irregularity is due to lexical diffusion, it is legitimate to reconstruct a root on the basis of an irregular series.

These two routines are among the tools the comparativist uses to avoid setting up false series or overestimating the time depth of certain reconstructions. However, for some loanwords they are inadequate, the final stage of adaptation

of a foreign segment being the ‘divising’ of an etymon that respects the rules of evolution of the language, in appearance only. Taking an earlier example, it may well be that in a few years *òrésì* will become <sup>x</sup>*òrésì*<sup>8</sup> in Mpongwe. At that point, it would be possible to set up an etymon with *z*, as all the *z* as C2 come from Proto-Bantu \**c*. And in this precise linguistic area such an evolution is quite possible, as many speakers understand, and speak, different dialects; they are conscious of phonological correspondences.

This point is crucial from a historical point of view: if one admits that it is possible to construct false etymons, the status of certain reconstructions becomes problematic. Borrowing between closely related languages is hard to establish, precisely because the historical rules are alike and it is easier to construct etymons. This appears to be the case in Proto-Bantu, where numerous reconstructions may in fact be borrowings passed from one language to another. As the comparativist considers that at this level of integration there is no possibility of detecting loans, regularity is not necessarily proof of inheritance from the parent language.

It is possible then to reconstruct words and integrate them into the stock of the parent language on the grounds of their regularity, when in fact they are of foreign origin, since, with a long time span, it is possible to (mis-)construct an etymon for loanwords. Guthrie (1967–1971), for example, reconstructed the term \**kòndè* “banana” in Proto-Bantu. However, thanks to historical research we know that the plant was imported into Africa. Since a concept is generally borrowed together with the term designating it, it is not impossible that the Bantu cognates of this root are of foreign origin, especially as the root \**kòndè* has a doublet in \**kòndò*. Other reconstructions of Guthrie’s also have variants and refer to various kinds of knowledge, and thus could also be successive borrowings. Even certain single roots with irregular reflexes could be borrowings.

By applying the classical method of reconstruction, one runs the risk of attributing to roots a status they do not actually have. If the root \**kúdò*/\**kúdù* “tortoise” is a problem in Bantu, it is because Guthrie simply compared the forms found in various languages, accepting the divergences of certain ‘cognates’. Under those conditions, it is possible to assign to the protolanguage a word of foreign origin belonging in fact to a later chronological stage. So it

---

8. <sup>x</sup> indicates a form which does not actually occur in the language and is likewise not a root or a virtual root.

is the whole approach to setting up series for the reconstruction of cultural vocabulary that needs to be reconsidered.

The comparativist can detect words in the process of diffusion if he applies to each phoneme the historical rules appropriate to each language, if he simulates a potential origin, using the principles for evolution set up from the basic vocabulary. This is where the notion of 'virtual reconstruction' comes in; this is the word, not as it occurs in the language, but such as it appears when one has tried to find an origin for it by applying the rules established on the basis of the core vocabulary.

The first systematic application of this method was by Hombert (1988), when he used it for the reconstruction of mammal names in the Bantu languages of zone B in Gabon. Although the principle of testing regular forms was not new, Hombert did not stop there since, as we have seen, this procedure does not make it possible to detect apparently etymological loanwords. For the first time, as far as we are aware, he introduced the concept of 'pseudo-reconstruction'. We have gone further with the idea, using ethnoichthyological names, and prefer the term 'virtual reconstruction', as 'pseudo' suggests that the forms produced by simulating the history of words are false roots or something like roots, which is by no means the case.

The method presupposes the existence of a protolanguage and of rules making it possible to understand how the descendant languages emerged. The system of the protolanguage and the rules governing its evolution are established as a result of the comparison of the core vocabulary of the descendant languages. As this vocabulary is least subject to borrowing, it is assumed that it approximates more closely the early stage of the language. As the cultural vocabulary is more readily renewed, thanks to borrowing in particular, its reconstruction calls for a particular procedure. The idea of testing the cultural vocabulary against a selected protolanguage makes it possible not only to detect terms that have appeared in the language relatively recently, but also to postulate another level of reconstruction, later than the protolanguage.

We shall now see, starting from Proto-Bantu, how we set about making virtual reconstructions. Linguists working on Bantu use the Proto-Bantu reconstructions of Guthrie (1971) and of Meeussen (1965), between which there is no significant difference as far as the phonological system is concerned. This system consists of seven vowels (\*i,\*e,\*ε,\*a,\*ɔ,\*o,\*u) and ten consonants (\*p,\*t,\*c,\*k,\*b,\*d,\*j,\*g,\*m,\*n), as well as a series of prenasalized consonants (\*mp,\*mb,\*nt,\*nd,\*nc,\*nj,\*nk,\*ng). Comparison of the Proto-Bantu lexicon and that of any particular Bantu language makes it possible to set up a table



of correspondences. For instance, if we take the Proto-Bantu/Mpongwe pairs for “breast”, “sand” and “raw”, we have \*b<sup>é</sup>édè>+bene (àmbénè), \*c<sup>é</sup>g<sup>é</sup>>+zeye (òzégè), \*b<sup>é</sup>cù>+bezo (mbézò). From this it is possible to make the following predictions: the Proto-Bantu segments \*b,\*d,\*c,\*g,\*ε,\*e,\*u correspond to b, n, z, γ, ε, e, o in Mpongwe. In this way, with a sufficient number of Proto-Bantu/Mpongwe terms, it is possible to establish the regular correspondences between Proto-Bantu and Mpongwe:

- \*i>i, \*e>e, \*ε>ε, \*a>a, \*ɔ>ɔ, \*o>o, \*u>u for the vowels;
- \*p>β, \*b>w, \*t>r, \*d>n, l, \*c>z, \*j>z, \*k>γ, \*g>γ, \*m>m, \*n>n for the consonants;
- \*mp>mp, \*mb>mb, \*nt>nt, \*nd>nd, \*nc>ntf, s, \*nj>nd<sub>3</sub>, s, \*ng>ng, nk, \*nk>nk, ng for the prenasalized consonants.

Any native Mpongwe word that derives from Proto-Bantu must conform to these rules. It should not be forgotten that at this stage the words tested consist of segments that are phonemes in the present-day language. Words containing segments which do not belong to the phonemic system of the language have already been eliminated by the first procedure. For instance, “kitchen” is *kíftn* in Mpongwe. The sound *f* is found only in this word, derived from the English – another word belonging to the same lexical field as “glass”, “plate” and “rice”. Similarly, *òrésì* “rice” cannot go back to Proto-Bantu as *s* does not occur as C2 in Mpongwe in regularly derived words; nor can *èpélè* “plate”, as *p* derived from Proto-Bantu is necessarily preceded by *m*. The more incompatible segments a word contains, the less it is likely to go back to Proto-Bantu – on the contrary, the more recent is its acquisition. The segments in *ikókò* “sugar cane” violate the evolutionary pattern of the language: there is no rule allowing *k* to be derived from Proto-Bantu, either in C1 or C2, unless preceded by a nasal – otherwise, Proto-Bantu \**k* becomes *γ*. So there are two segments breaking the rules of correspondence between Mpongwe and Proto-Bantu. In this case, it is the second routine procedure that applies: there is still no virtual reconstruction.

On the other hand, *ntféyè* “*Papio mandrillus sphinx*” could derive from Proto-Bantu, as *ntf* could come from \**nc*, *ε* from \**ε*, *γ* from \**g*. If so, the term should also occur in other Bantu languages, and be compatible with their rules. It is now necessary to apply the third routine procedure, which makes it possible to detect etymologically feasible borrowings in a reconstruction.

°*cege* is the virtual reconstruction of Mpongwe which justifies the form *ntféyè*. Compare this with reconstructions in other Gabon languages. In Sangu, for example, one finds *tséyi* for “*Papio mandrillus sphinx*”. In this language,

the Proto-Bantu front vowels have become *i* in word-final position (\**i*>*i*, \**e*>*i*, \**ε*>*i*). The evolution of the palatal and velar consonants is \**c*>*ts*, *s*, \**g*>*ʃ*, \**k*>*ʃ*. So there are several virtual reconstructions possible for +*tseyi*: °*cege*, °*cegi*, °*cege*, °*ceke*, °*ceki*, °*ceke*. At this stage in the procedure all virtual reconstructions are equally good, except those in which certain segments involve a marginal rule. This is the case in Sangu *kùdù* “tortoise”: *k* can come from \**k*, *u* from \**u*, *d* from \**d*, but the process \**k*>*k*/*\_*\**u* is marginal: in this context, there is regularly fricativization (\**k*>*f*/*\_*\**u*). To indicate that there is a problem with a virtual reconstruction in relation to the protolanguage we will use parentheses, so for the above example, we have (°*kudu*) in Sangu.

Ultimately, it is the comparison of the virtual reconstructions of various languages of the family that determines the choice of the best root for the status of a proto-form.

## 2.2 The comparison of virtual reconstructions

In classical comparative linguistics, a proto-form is arrived at by comparing items as they are found in different languages. Determining the etymon is not automatic; briefly, it may be said that the rules justifying it must be simple, phonetically plausible, compatible with the typology, and adequate to explain the form. In our case, the procedure for settling on the etymon is somewhat different: it consists in choosing from among the various virtual reconstructions the one that is common to all the languages; it is often directly evident.

We have seen that in Mpongwe *ntféyè* “*Papio mandrillus sphinx*” could come from °*cege*; in Sangu *tséyi* could from °*cege*, °*cegi*, °*cege*, °*ceke*, °*ceki*, °*ceke*. Taking into account that °*cege* is also one of the possible virtual reconstructions in a number of languages in Gabon having items analogous to those of Mpongwe and Sangu, one can say that \**cege* is the proto-form and that °*cege*, °*cegi*, °*ceke*, °*ceki*, °*ceke* are by-products of the procedure in Sangu. It is in this case, then, that the form selected is preceded by an asterisk: \**cege* is a reconstructed root, whereas °*cege* (form preceded by a small circle) is a virtual reconstruction, one virtual root among others.

When selecting a proto-form, it is important to examine the virtual reconstructions within language groups carefully, comparing the various dialects. The virtual reconstructions that raise a problem in several dialects of the same group should be assessed differently from those that are problematic in one language only.

When there is a great difference between several virtual reconstructions, it is probable that they do not go back to the reference protolanguage. If we remember that it is possible to construct an etymon for a foreign word, we should not be surprised if the ‘etymologizing’ strategies differ from one language to another. In other words, if some languages borrow a term from a given language, there is no guarantee that the proto-segments that each ‘invent’ to generate the segments for the new word will be identical. The more closely the languages are related, the more likely it is that the virtual reconstructions will resemble one another, but even when the differences are very slight it is possible to detect borrowings. For example, the virtual reconstruction for Mpongwe *kɔ̀pù* “cup” gives  $^{\circ}k\text{ɔ}pu$ , whereas Punu *kòpà* gives  $^{\circ}k\text{ɔ}pa$  and  $^{\circ}k\text{ɔ}p\text{ɔ}$ . Although slight, the difference is significant as it reveals two different strategies for integrating the English word “cup”.

We have said that everything that is irregular does not go back to Proto-Bantu, but the reverse is not necessarily true. Not every regular item, not every reconstruction from virtual roots is inherited from the reference protolanguage. There are several factors that can lead to a ‘false root’, even though it is based on virtual reconstructions.

In the first place, the length of the reconstruction matters. For example, in Punu the infinitive prefix goes back to a virtual  $^{\circ}to$ , like the English “to”. It is obvious that this is pure chance. The example is of course a caricature, but the fact is that there are many monosyllabic stems in Bantu languages, which can present problems. All in all, virtual disyllabic roots are surer candidates for Proto-Bantu status. The ‘etymological weight’ of a proto-form increases with its length.

The nature of the rules linking the reference language to the forms found is also to be taken into account. Let us take the example of the Proto-Bantu root  $*m\grave{a}n\text{-}a$  “finish”; the corresponding forms in Mpongwe and in Fang-Ntumu are both *man-*, for in these two languages, Proto-Bantu  $*m$ ,  $*n$  and  $*a$  have remained unchanged. On the other hand, the Proto-Bantu root  $*d\acute{e}m\grave{e}$  “tongue” has given *+nɛwɛ* (*ðnéwè*) in Mpongwe and *+yám* (*áyám*) in Fang-Ntumu. Proto-Bantu  $*d$  has given *n* in one case and *y* in the other. The  $*e$  has given *ɛ* in Mpongwe and *ə* in Fang, the  $*m$  has given *w* and *m* respectively. The English word *man* can be borrowed easily because it consists of segments for which etymons can easily be constructed. On the other hand, the French word *dame* can easily be identified as a borrowing: the same strategies do not apply for the construction of the virtual roots from which it could be generated.  $^{\circ}dam$  is possible in Mpongwe, but in Fang, as Proto-Bantu  $*d$  gives a *y*, one is

confronted with an item the first consonant of which is incompatible with the rules linking Proto-Bantu and Fang-Ntumu.

Given that certain segments of a proto-form can be stable from one language to another, in so far as the phonological rules linking them to the attested forms do not modify them, or do so in the same way, the 'etymological weight' of a proto-form is inversely proportional to the degree of stability of the segments. In fact, this corresponds to a principle of classical comparative linguistics according to which it is the regularly divergent analogical forms, rather than identical forms, that make it possible to suspect a relationship.

Lastly, geographic proximity is a factor to be taken into account: monosyllabic items that could be generated by 'non-modifying' rules can pass easily from one language to another, especially if the languages are geographically close together. So the 'etymological weight' of a proto-form increases with the geographic distance separating the languages in which the items going back to this proto-form are attested.

Thus a proto-form which raises no problem as far as reconstruction on the strictly phonological level goes must still be assessed with regard to these three parameters. They are important when, after reconstruction of proto-forms, one tries to group the various languages in order to find historical patterns. Take for example three languages – A, B and C – where B has features in common with both A and C, but A and C have nothing in common with one another. The concept of etymological weight can help to solve this apparent contradiction. We will consider B to have a stronger, older, link with one of the languages if the proto-forms they have in common are not restricted to a geographical area, if these proto-forms are not simply a few monosyllables and, finally, if the rules of derivation produce different features between the proto-forms of the reference language and the derived segments.

One often finds two simple cases when comparing virtual reconstructions: either the items have radically different virtual reconstructions, or the reconstructions are identical. But sometimes items have virtual reconstructions that are similar but not identical. These may be derived from the same proto-language but reflect variant forms in the reference language.

### 2.3 Chronology of phonetic changes and the status of the reconstructions

Some processes, either not found or marginal in the basic vocabulary, may, in the cultural vocabulary, show a certain regularity. It is conceivable that the changes have extended to new contexts and then to new words, as with lexical

diffusion. Thus, by carefully distinguishing the different irregularities, one can situate the roots better in time. There is a period X, the time of the reference protolanguage, to which certain words showing irregularities in reconstruction cannot be traced back. There are subsequent periods in the course of which certain processes become regular; the words in which these occur can then be reconstructed. When the word is irregular in several languages, one can distinguish between what is due to borrowing and lexical diffusion and what is due 'exclusively' to the latter.

Let us again take the example of the words *ikóndò* "banana" and *ikókò* "sugar cane", which we find in Mpongwe. They cannot come from Proto-Bantu, for, if we take the case of the basic vocabulary, there is no *k* from Proto-Bantu in the language unless it is preceded by *n*. To show the irregularity of words like this, it is necessary either to check through the rules (2nd routine procedure) or to compare the virtual reconstructions (3rd routine procedure). For this reason, we insist on the principle that a reconstruction of the cultural vocabulary can only be carried out in the light of the rules established for the basic vocabulary. This is essential if we are to avoid assigning roots to the wrong chronological stage. A marked divergence between virtual reconstructions is a sure indication of reconstructions to which incorrect time depth might be assigned.

Those processes which would be irregular in the basic vocabulary will be the clue for the virtual reconstructions in which they occur: the roots derived from them will belong to chronological stages later than the reference protolanguage.

We shall now look at a certain number of phonological processes which are either non-existent or very rare in the basic vocabulary, but much less so in the cultural vocabulary, specifically in ichthyological items.

*i – Extension of the context of fortis consonants*

In Mpongwe the distribution of fortis/lenis consonants is conditioned by the nasal consonant: fortis consonants occur only after a nasal ( $*p > p/*m\_$ ,  $*b > b/*m\_$ ,  $*t > t/*n\_$ ,  $*d > d/*n\_$ , etc.); lenis occur intervocalically ( $*p > \beta/*V\_ *V$ ,  $*b > w/*V\_ *V$ ,  $*t > r/*V\_ *V$ ,  $*d > l/*V\_ *V$ , etc.). However, fortis consonants are beginning to appear between vowels.<sup>9</sup>

9. In this paper, in accordance with the practice of Bantu scholars, we apply the term 'fortis' to consonants which have remained after a Proto-Bantu nasal; in Mpongwe this is

*ii – Preservation of stops before close vowels: no development of fricatives*

In many languages of the groups B 40, B 50, B 60, B 70 stops followed by a close vowel in Proto-Bantu become fricatives: (\*mp>mf/\_\*u, \*mb>mv/\_\*u, \*d>r/\_\*i,\*u, \*t>s/\_\*i). Comparing the data of these languages with the data of languages where this process of fricativization does not occur, we are sometimes led to set up sequences °du, °di, °mbu, °mpu, contrary to the rule of fricativization.

*iii – Appearance of prenasalized consonants as C2*

In the groups B 60 and B 70 there is a rule of simplification of Proto-Bantu prenasalized consonants as C2 (\*mb>m,\*nd>n,\*ng>ŋ). However, our corpus shows examples contradicting this rule: there are words with *mb*, *nd*, *ng* as C2.

*iv – Appearance of voiced labial prenasal consonant as C2*

In Sake, the Proto-Bantu prenasalized sequence \*mb becomes *mp*, but in our material we also find *mb* as C2.

*v – Appearance of r as C2*

In B 50 and in Kanigi there is no *r* derived from Proto-Bantu, the consonants from which it might have been derived have given *d*, *l*, *t*. However, this consonant does occur in certain fish names.

*vi – Unconditioned nasalization of Proto-Bantu \*d*

Whereas the rules of the evolution from Proto-Bantu generate an *n* from \*n (\*n>n), and occasionally from \*d, in accordance with Meinhof's Law,<sup>10</sup> comparison of the correspondences between Mpongwe and the other languages examined implies an unconditioned nasalization of \*d.

These examples suggest that a good deal of the lexicon is later than the original Proto-Bantu (PB-X in Guthrie). There is no doubt that there were stages after the initial Proto-Bantu period; Guthrie suggested a splitting of the original Proto-Bantu into two major groups, Proto-Bantu A and Proto-Bantu B. This can be seen as a new chronological stage. The reconstructions for certain zones

---

the case with the stops and the voiceless alveolar fricative; all the other fricatives are 'lenis': they occur in intervocalic position.

10. "In a word, a sequence consisting of a nasal and a voiced oral consonant is represented as a double or long nasal if the following syllable contains a nasal" (quoted by Meeussen 1965:5).

can also be seen as indicative of stages subsequent to PB-A and PB-B. What counts is the chronology of the phonetic changes and the age of the intermediate levels. We see that the phonological system has undergone reorganisation at later stages, in the evolution of the protolanguage.

### 3. Distribution of roots, diffusion of terms and historical interpretation

Finally, a word is irregular if it is not possible to find potential parent-forms for each of its segments in several languages, or if the virtual reconstructions are different in several languages. When a word is irregular in one language only, or when the virtual reconstructions differ only slightly, lexical diffusion or expressivity are likely explanations. If repeated divergences are the sign of diffusion by borrowing, an irregular process, occurring in a single language, can mark the initial phase of a change.

The concept of area is well-suited to describe a space in which languages share similar words. These can originate in one language and spread to adjacent languages. It is possible to distinguish areas of old diffusion from those of recent diffusion. Languages which are no longer in contact but share certain words must have been contiguous at some earlier time. Five patterns can be detected as the result of a comparison of the cultural vocabulary of the dialects of a given language, following the procedures already presented.

(i) The first case is that in which words that are irregular in certain languages occur in languages geographically near, where they are regular. Such words are to be accounted for by relatively recent spreading. For example, almost all the irregular words in Evia (B 30)<sup>11</sup> are related to roots in B 40, especially Eshira.

(ii) The second case is that of languages in which there are irregular words that are found in geographically remote languages; in this case, the borrowing is older. Thus, many irregular words in B 10 are related to roots found only in B 30. While the names of freshwater fish are irregular in Mpongwe and regular in B 30, and even in B 40, the names of saltwater fish tend to be regular in B 10 and irregular in the other groups: borrowing has gone both ways.

---

11. For the list of language groups and zones, see appendix.

(iii) The third case is that in which terms that are irregular in certain languages are very widespread. They can be found in several languages of the whole territory that the group in question belongs to. A distribution of this kind is the result of a very old diffusion. There are words which are irregular in many language groups in Gabon which are also found in other Bantu areas. Guthrie's reconstructions for "tortoise", "banana" and many other roots denoting cultural entities belong to this category.

(iv) The fourth case is that of regular words that coincide with the boundaries of the language groups. Here, the study of the specialised vocabulary can confirm the classifications based on traditional criteria (phonetic laws, lexicostatistics). Groups B 30 (except for Evia) and B 40 are particularly homogeneous and clearly delimited by roots found only there. Group B 20, despite the scatter of the languages constituting it, likewise has roots peculiar to it. On the contrary, A 75, B 10 and especially B 50 and B 60 have very few terms peculiar to them. This is not surprising: they are made up of languages in which many terms are irregular and borrowed from the neighbouring languages.

(v) The fifth case is that of terms that suggest certain regroupings that are interpreted as areas over which words have spread, and are still spreading, rather than as genetic units. The distribution is not always regular, the patterns can change according to the roots, analogies in the cultural vocabulary being determined often by geographic proximity. For example, certain roots make it possible to distinguish the languages of B zone in which they are found from Fang (zone A), which has different words. Within zone B, there are affinities between B 10 and B 30, which are in turn distinguished from the block B 20, B 40, B 50, B 60, B 70. Albeit, the languages of the groups B 20 and B 50 have some terms in common with B 30.

Finally, one distinguishes what is inherited from the reference protolanguage, what goes back to the great ensembles formed after the break-up of the original body, and lastly what is much more recent. In this way, it becomes possible to see more clearly certain aspects of the history of the language and populations involved. And we can then establish correlations with the findings of other disciplines with more confidence.

Given the variety and the great number of irregularities in ichthyology, it would appear that the greater part of the knowledge about freshwater species is not inherited from PB-X. Admittedly, some roots could be quite old as they



are to be found in other Bantu zones as well (chiefly C, D and H), but the majority show very restricted distributions. The most likely hypothesis is that of foreign origin, combined with phenomena of internal renewal. The Bantu peoples in Gabon must have encountered other populations who transmitted this knowledge. As the Pygmies are not heavily involved in fishing, the origin of this knowledge is to be sought elsewhere. The data clearly indicate two great occupation areas within Gabon: the north and the southeast. In the north, group B 30 appears to have settled before the A 75 groups, and especially before B 10, which owes the majority of its names for freshwater fish to B 30. In the southeast, B 40, B 50, B 60 and B 70 seem to constitute a single area, reflecting a shared migratory tradition. Group B 20 is influenced by the traditions of both north and south: closer in the present day to the southern groups, B 20 must have had early contacts with the northern groups.

Knowledge of saltwater fish seems to be late also, which would imply a recent settling of the B 10 on the coast. At any rate, the linguistic data do not offer proof of a migration by sea, at least for the populations we have studied. Most ethnic groups settled along the coast would have stayed at one stage in the region of the Ivindo River, to judge by the B 10 borrowings from B 30. So the migrations would have taken place from east to west, essentially along the inland rivers. The Ogooue basin has therefore been an important factor in the implantation of the present populations within Gabon.

It is a fact that the ichthyological fauna is more varied than that of mammals, and this may in part explain why there has been such extensive renewal in this lexical field. As the very few pan-African ichthyological species that exist are stable, it may be that the diversity of names is the consequence of this biological fact.

It is noteworthy that the study of the names of fish on the whole confirms the findings of other branches of study; sometimes it sheds light on shady points. Lexicostatistics (Bastin et al. 1999) and the study of grammatical innovations (Nurse & Philippson 2003) have already shown the affinities between B 10 and B 30, and our data confirm this. But is it really a case of genetic relationship or is it the result of a long period of cohabitation? Van der Veen (1991), studying the grammatical morphemes of the two groups, favours the hypothesis of prolonged contact. The spread of freshwater fish names from B 30 fits the contact scenario better. Our data also confirm the cohesion of groups B 40, B 50, B 60, B 70. Bastin & Piron (1999), using lexicostatistical trees, have also shown that B 20 is a 'floating group', sometimes going along with the B 40, B 60, B 70 groups and a number of H languages, and sometimes with the languages

of zone A, B 10 and B 30. Our data confirm this tendency to fluctuate, but they also confirm the unity of the group, the languages of which share quite specific roots, in spite of the fact that they are widely scattered.

Is it possible to establish a chronological frame for the appearance or the development of this ichthyological culture? It is clear that most fish terms do not go back 2,000 years. Actually, the age of West Proto-Bantu is generally assigned to a period between 3,500 and 3,000 years ago, the earliest subdivisions occurring between 3,000 and 2,000 years ago.<sup>12</sup> As there are virtually no Proto-Bantu terms for fish and as very few of those could go back to these first splits, the bulk of the current ichthyological culture would not reach back 2,000 years.

It is striking, furthermore, to see that the spread of fish names follows the same trajectory as that of iron; there is evidence of iron in the centre and south-east of Gabon long before it reaches the coast (Clist 1995). It would seem that there are also points in common with anthropology, as it has been shown that B 10 has borrowed the totality of its traditional rites from B 30. Ichthyological knowledge is thus an element in an ensemble of cultural features that have spread from this group.

## Conclusion

The aim of this study has been to show that the examination of a cultural vocabulary necessitates distinguishing between different causes of phonological irregularity. It is possible to show that the recurrence of irregularities in several dialects or languages is the result of successive borrowings. Lexical diffusion can thus be called on to account for an irregularity in a given language, but not for a string of irregularities. The mechanism can explain the starting point of a change and the resulting irregularity, before the phenomenon becomes generalised. On the other hand, it cannot explain the spread of an irregularity through a network of dialects or languages. Expressivity can also produce effects that diverge from the regular phonological pattern, but this is quite restricted.

The concept of 'virtual reconstruction' was introduced to detect borrowings which are in the process of being etymologised. The basic idea is that lan-

---

12. The Tervuren classification (1999) implies the following four primary branchings: the Mbam-Bubi; Northwest Bantu (the rest of zone A, B 10, B 30 and perhaps B 20); Centre West Bantu (B 40, B 50, B 60, B 70, B 80, C, H, K, R); East and South Bantu.

guages, even when they are very closely related, do not apply exactly the same rules for constructing etymons for borrowed words. The virtual reconstructions for any given word may differ, a marked divergence being a clue to borrowing. To attempt to retrace the history of cultural items by applying to them the rules established for the basic vocabulary amounts to examining their time depth. It is thus possible to establish the order of phonetic changes and assign the roots to chronological stages in a satisfactory way.

The limits of the method are the same as those of comparative linguistics in general. The vocabulary of any language includes borrowings going back to very remote periods and these are no longer perceived as borrowings. By extrapolation, it can be said that in a language it can never be proved that words have not been borrowed; which amounts to saying that the demonstration of borrowings, like that of linguistic relationship indeed, is always positive.

However, by showing the processes of diffusion, the identification of words being integrated is of real interest for historical reconstruction in general. We have shown that a large part of the ichthyological vocabulary was later than Proto-Bantu. As the lexicostatistical classifications and those established on the basis of phonetic laws are on the whole confirmed by our data, the method is of real utility, all the more so as certain points support archaeological and anthropological hypotheses.

Finally, even if endogenous processes can also account for the recent character of the ichthyological vocabulary, the percentage of irregularities is too great not to think of borrowing. So the situation differs from that of mammal names, which are on the whole of unquestionable Proto-Bantu origin. Could it be that outside the Bantu area fish names show the same characteristics? The comparison would only be possible if there was a comparable study, which might make it possible to improve on certain suggestions made here.

## References

- Bastin, Yvonne. 1994. "Reconstruction formelle et sémantique de la dénomination de quelques mammifères en bantou". *Afrikanistische Arbeitspapiere* 38.5–132.
- Bastin, Yvonne, André Coupez & Michael Mann. 1999. *Continuity and Divergence in the Bantu Languages: Perspectives from a lexicostatistic study*. (= *Annales Sciences Humaines*, 162). Tervuren: Musée Royal d'Afrique Centrale.
- Bastin, Yvonne & Pascale Piron. 1999. "Classifications lexicostatistiques: Bantou, bantou et bantoïde. De l'intérêt des 'groupes flottants'". *Bantu Historical Linguistics* ed. by Jean-Marie Hombert & Larry Hyman, 149–163. Stanford: CSLI Publications.

- Chen, Matthew Y. & William S.-Y. Wang. 1975. "Sound Change: Actuation and implementation". *Language* 51:2.255–281.
- Clist, Bernard. 1995. *Gabon: 100000 ans d'Histoire*. Libreville: Centre Culturel Français Saint-Exupéry & Sépia.
- Coupez, André. 1975. "La variabilité lexicale en bantou". *African Languages/Langues Africaines* 1.164–203.
- Darnell, Regna & Joel Sherzer. 1971. "Areal Linguistic Studies in North America: A historical perspective". *International Journal of American Linguistics* 37.20–28.
- Gauchat, Louis. 1905. "L'unité phonétique dans le patois d'une commune". *Aus Romanischen Sprachen und Literaturen: Festschrift Heinrich Morf*, 175–232. Halle: Niemeyer.
- Greenberg, Joseph H. 1955. *Studies in African Linguistic Classification*. New Haven: Compass Publishing Company.
- Greenberg, Joseph H. 1963. *The Languages of Africa*. The Hague: Mouton.
- Guthrie, Malcolm. 1967–1971. *Comparative Bantu*, 4 volumes. Farnborough: Gregg.
- Hombert, Jean-Marie. 1988. "Mammals in the Languages of Gabon: A step towards the reconstruction of Proto-Bantu fauna". 18th Colloquium on African Languages and Linguistics, Boston.
- Kiparsky, Paul. 2003. "The Phonological Basis of Sound Change". *The Handbook of Historical Linguistics* ed. by Brian D. Joseph & Richard D. Janda, 313–342. Oxford: Blackwell Publishing.
- Janssens, Baudoin. 1993. *Doubles réflexes consonantiques: Quatre études sur le Bantu de Zone A (Bubi, Nen, Bafia, Ewondo)*. Ph.D. Dissertation, Université Libre de Bruxelles.
- Labov, William. 1963. "The Social Motivation of a Sound Change". *Word* 19.273–309.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 1989. "The Exact Description of the Speech Community: Short a in Philadelphia". *Language Change and Variation* ed. by Ralph Fasold & Deborah Schiffrin, 1–57. Washington, D.C.: Georgetown University Press.
- Labov, William. 1994. *Principles of Linguistic Change: Internal factors*. Oxford & Cambridge: Blackwell.
- Meeussen, Achille E. 1965. *Bantu Lexical Reconstructions*. (= *Archives d'anthropologie*, 27.) Tervuren: MRAC.
- Mouguiama-Daouda, Patrick. 1995. *Les dénominations ethnoichthyologiques chez les Bantu du Gabon: Une étude de linguistique historique*. Ph.D. Dissertation, Université Lumière-Lyon 2.
- Mouguiama-Daouda, Patrick. 1999. "Les noms de poissons dans les langues bantoues du Gabon: Une étude de la motivation étymologique". *Afrikanistische Arbeitspapiere* 58.69–96.
- Nurse, Derek & Gérard Philippson. 2003. "Towards a Historical Classification of Bantu Languages". *Bantu Languages* ed. by Derek Nurse & Gérard Philippson, 164–179. London & New York: Routledge.
- Sapir, Edward. 1921. *Language: An introduction to the study of speech*. New York: Harcourt, Brace & Co.

- Schuchardt, Hugo. 1885. *Über die Lautgesetze gegen die Junggrammatiker*. Berlin: Oppenheim. (Repr. & Engl. transl. 1972, Theo Vennemann & Terence H. Wilbur).
- Sturtevant, Edgar. 1907. *Linguistic Change*. Chicago: University of Chicago Press.
- Sturtevant, Edgar. 1940. *The Pronunciation of Greek and Latin*. Philadelphia: LSA.
- Sturtevant, Edgar. 1947. *An Introduction to Linguistic Science*. New Haven: Yale University Press.
- Van der Veen, Lolke. 1991. *Etude comparée des parlers du groupe okani (B 30)*. Ph.D. Dissertation, Université Lumière-Lyon 2.
- Wang, William S.-Y. 1969. "Competing Changes as Cause of Residue". *Language* 45:1.9–25.
- Weinreich, Uriel, William Labov & Marvin Herzog. 1968. "Empirical Foundations for a Theory of Language Change". *Directions for Historical Linguistics* ed. by Winfred P. Lehmann & Yakov Malkiel, 97–195. Austin: University of Texas.

## Résumé

Cette étude propose de distinguer les irrégularités dues à l'emprunt de celles dues à la diffusion lexicale et aussi de celles dues à l'expressivité. La méthode suggère la comparaison des reconstructions virtuelles comme base de la reconstruction. Les reconstructions virtuelles résultent de l'application à rebours au lexique culturel des règles phonologiques établies à partir du lexique de base. Nous parvenons ensuite à mieux sérier les racines ou les items sur différents paliers chronologiques. Nous illustrons la méthode en étudiant les noms des poissons dans les langues bantu du Gabon. Nous parvenons à mettre en évidence des courants migratoires orientés d'est en ouest. De même, la comparaison des reconstructions virtuelles fait ressortir le caractère récent de la culture ichtyologique dont l'essentiel ne remonte pas au proto-bantu.

## Zusammenfassung

Ziel dieser Studie ist es, die Unregelmässigkeiten, die aus Entlehnungen entstehen, von denen zu unterscheiden, die in der lexikalischen Diffusion oder in der Expressivität ihren Ursprung haben. Die Methode schlägt einen Vergleich von virtuellen Rekonstruktionen als Basis für die Rekonstruktion vor. Virtuelle Konstruktionen entstehen dadurch, dass die für den Basiswortschatz etablierten phonologischen Regeln in umgekehrter Reihenfolge auf den kulturellen Wortschatz angewendet werden. So kann man die Wurzeln oder die lexikalischen Einträge verschiedenen chronologischen Stufen besser zuordnen. Wir illustrieren die Anwendung dieser Methode, indem wir die Namen der Fische in den Bantu-Sprachen Gabons erforschen. Somit können wir in Anbetracht der Verteilung der Wurzeln die Migrationsbewegungen von Osten nach Westen besser erklären. In ähnlicher Weise verdeutlicht der

Vergleich der virtuellen Rekonstruktionen, dass die ichthyologische Kultur relativ jung ist und im wesentlichen nicht aus der Proto-Bantu-Zeit stammt.

*Author's address:*

Patrick Mouguiama-Daouda  
Laboratoire Dynamique du Langage (UMR 5596, CNRS, Lyon)  
14 Avenue Berthelot  
Institut des Sciences de l'Homme  
69363 LYON Cedex 07  
France  
E-mail: [patrick.mouguiama-daouda@ish-lyon.cnrs.fr](mailto:patrick.mouguiama-daouda@ish-lyon.cnrs.fr)

## Appendix A

Table of the roots for each group<sup>13</sup>

	A 75	B 10	B 20	B 30	B 40	B 50	B 60
<i>Alestes macrophthalmus</i>	°pepakɔ						
<i>Barbus batesii</i>			*j̀̀n̄j̄i	*j̀̀ng̀̀è			
<i>Barbus compinei</i>					*g̀̀n̄i		
<i>Barbus holotaenia</i>				*j̀̀ng̀̀ò			
<i>Bricynus kingsleyae</i>			*g̀̀ng̀̀à		*p̀̀nz̄i		
<i>Brycinus longipinnis</i>		°jagayamba		*panga			
<i>Brycinus schoutedeni</i>			*k̀̀m̀̀á			*b̀̀à	
<i>Bryconaethiops microstoma</i>			*b̀̀eǹ̀a	*d̀̀udu	*d̀̀àng̀̀á		
<i>Caecomastacembelis sclateri</i>			*ɔ̀̀dng̀̀é		*t̀̀nd̀̀ò		
<i>Channallabes apus</i>			*t̀̀ú̀̀ba				
<i>Chrysichthys nigrodigitatus</i>			*ɔ̀̀mb̀̀ɔ̀̀da				*k̀̀eyi
<i>Clarias gariepinus</i>					*ng̀̀ɔ̀̀t̄ɔ̀̀		
<i>Distichodus fasciolatus</i>			*c̀̀ánd̀̀ó	*c̀̀òk̀̀ò	*k̀̀ù		
<i>Distichodus hypostomatus</i>				*k̀̀òc̀̀è			
<i>Distichodus notospilus</i>		°judu		*k̀̀òc̀̀ò	*k̀̀àc̀̀a		
<i>Elops lacerta</i>	°b̀̀òd̀̀é	°anga					
<i>Hemichromis elongatus</i>	°c̀̀òC̀̀á						*k̀̀eg̀̀e
<i>Hepstus odoë</i>	°c̀̀òd̀̀ó						
<i>Labeo sp.</i>			*b̀̀úng̀̀ú			*ng̀̀oada	
<i>Mormyrops sp.</i>			*t̀̀ɔ̀̀d̄i				
<i>Parauchenoglanis pantherinus</i>			*mb̀̀òng̀̀í				
<i>Pellonula vorax</i>		°cenjede					
<i>Raiamas buchholzi</i>					*p̀̀èní		*piyepiye
<i>Schilbe grenfelli</i>	°c̀̀òp̀̀v̄		*b̀̀én̄j̄é				
<i>Synodontis obesus</i>			*t̀̀ùk̀̀á		*g̀̀àng̀̀á		
<i>Synodontis polyodon</i>		°domba	*g̀̀ánỳ̀à	*p̀̀òng̀̀è			
<i>Tilapia cabrae</i>							
<i>Tilapia heudelotii</i>			*b̀̀únd̀̀ú				

13. We give here the roots which are discussed in Appendix B and also those which pose no particular problem and have therefore not been analysed there.

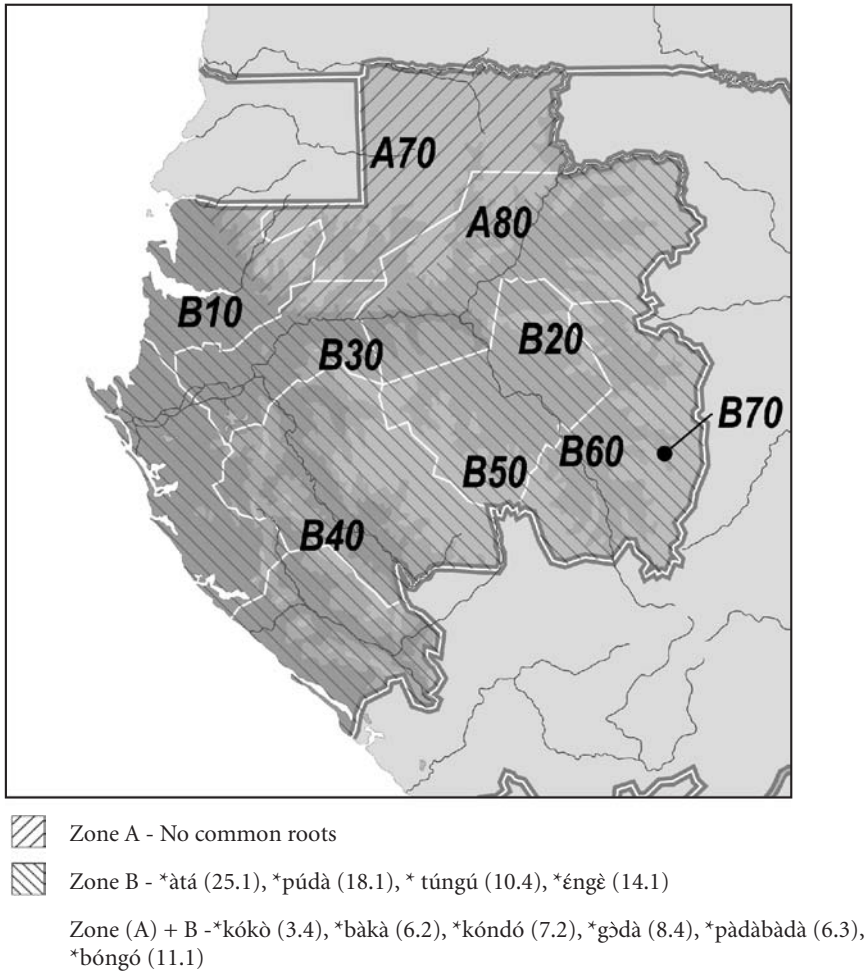
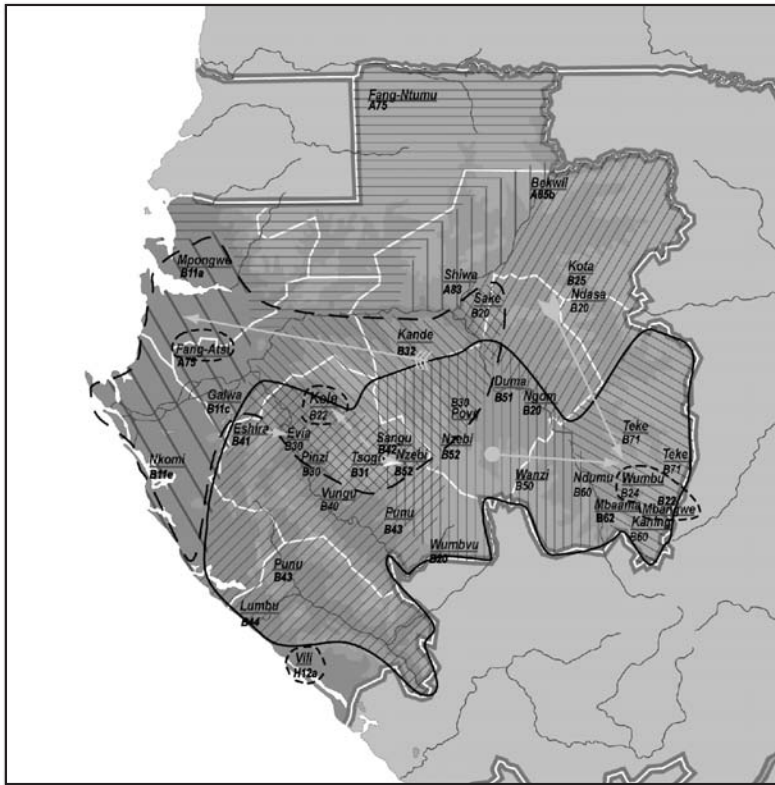


Figure 1. Guthrie's zones A and B in Gabon<sup>14</sup>

14. This map shows the roots of zone B and those which are common to the two zones (A+B).



- +tòrè (2.2), \*dómé (8.3), \*bòkà (10.2), \*còkò (13.1), \*jèngè (10.1)
  - \*kònà (3.2), \*cómí (8.3), \*jɔ̀nɔ̀ (10.1)
  - \*kémbé (3.2), \*còkò (13.1), \*tíndí (7.1), \*kóndó (7.2)
  - \*dàngá (6.5), \*gɔ̀tɔ̀ (8.2), \*gèni (10.2), °aβara (6.3)
  - \*gɔ̀tɔ̀ (8.2), \*gèni (10.2), \*kembé (3.2), \*còkò (13.1)
  - \*kómbò (3.2), \*ángá (6.1), ɛ̀ɔ̀ɔ̀ (13.1), \*tɔ̀ti (16.1)
- NB: On the map, the arrow indicates the direction of borrowing.



- ▨ A70   ▨ A80   ▨ B10   ▨ B20   ▨ B30   ▨ B40   ▨ B50   ▨ B60/B70
- B10 + B30 \*combo (3.1), \*yoci (8.4), \*kóndó (7.2), \*tíndí (7.1), \*yoyo (10.5), \*gòtè (20.1)
- B40 + B50 + B60 + B70 \*pété (1.1), \*kòkò (3.4), gɔ̀dà (8.4), \*báci (8.5), \*kòtò (13.2), \*jɔ̀mbó (24.1)

Figure 2. Distribution of roots<sup>15</sup>

15. Each type of hatching corresponds to one of Guthrie's groups, the languages are those for which we have data. The isoglosses show the two major areas which divide zone B. Group B 20 is divided between the two areas.

## Appendix B

### *Conventions and Explanations*

The following conventions should be noted, including some explanations for non-Bantuists:

◌<sup>o</sup>: Virtual reconstruction (e.g. ◌<sup>o</sup>baka). This is the reconstructed form for one language.

\*: Reconstructed root. This is the result of the comparison of the reconstructed forms (virtual reconstructions) of more than one language, implying a certain time depth. Such a form corresponds to a 'stage' in the phases of historical change and language grouping which may be more or less remote but is post-Proto-Bantu.

\*: The same sign is used to mark a Proto-Bantu root: Proto-Bantu is the earliest of such stages. The contexts always make it clear when a Proto-Bantu (PB) root is concerned.

PB-X is the abbreviation used by Guthrie (1967–1971) in Bantu Historical Reconstructions for his most remote reconstructions.

+ : Before an italicised form (e.g. +*βaya*) indicates that this is the stem of the actual word, minus the class prefix. In Bantu, the noun lexicon falls into classes, according to the prefix. The classes go in singular-plural pairs (class pairing). A word in italics without any preceding sign is the full word, with its prefix, usually the singular (e.g. *mukoyu*, *mu-*: prefix; *+koyu*: stem).

Tones are indicated as follows: H = High, L = Low, F = Falling.

Each syllable of a word has its tone, and one speaks of a tone sequence for the word. (LH) is the sequence Low-High, (HH) is the sequence High-High, etc. Most nouns are disyllabic. (LHL) is the sequence Low-High-Low of a trisyllabic word.

### List of languages

A75a: Fang-Ntumu, **Fang-Atsi**

A 80: **Bekwil, Shiwa**

B 10: Galwa, Nkomi, Mpongwe<sup>16</sup>

B 20: Kele, Kota of Mekambo, **Kota of Okondja**, Mbangwe, Ndasu, Ngom, Sake, Wumbu, **Wumbvu**

B 30: Evia, Kande, Pinzi, Pove, Tsogo

B 40: Eshira, **Lumbu**, Punu, Sangu, **Vungu**

B 50: Duma, Nzebi of Mbigou, **Nzebi of Koulamoutou**, Wanzi

B 60: Kanigi, Mbaama, Ndimu

B 70: Teke

H 12: Vili

We present in this appendix the reconstructions for the names of some fifty ichthyological species found in the languages of Gabon listed above. These languages are classified in zones (A, B, H) and, within the zones, into groups (A 75a, A 80, B 10, B 20, B 30, B 40, B 50,

---

16. The fish names in the three languages of B 10 are the same, so unless there is an indication to the contrary, we give examples only in Mpongwe, as these are valid for the whole group.

B 60, B 70, H 12), according to the classification of Guthrie (1967–1971).<sup>17</sup> For each of the twenty-six languages not given in boldface, historical studies make it possible to establish a table of correspondences with Proto-Bantu. For the nine languages printed in bold above, no such study is available, but data from them help to make the area of distribution of a root clearer. For the same reason, we sometimes refer to roots, or simply names, found in Bantu languages in neighbouring countries.

Among our twenty-six languages, the tonal systems of some have not yet been studied, so there are roots reconstructed without tones.<sup>18</sup> The roots reconstructed with tones are presented as follows: for example, for “*Clarias pachynema*”, the common silurid, we give \*ngɔlə (LL); this means that the tone-pattern for this root is \*Low-Low. The three other tone-patterns reconstructed are \*HH (\*High-High), \*HL (\*High-Low), \*LH (\*Low-High).

It is not only the forms that are reconstructed, but also the meaning. It is important to remember that the first condition for obtaining meaningful results from the historical point of view, when reconstructing ethnobiological vocabulary, is the rigorous identification of the biological specimens. It is no use looking for likely historical hypotheses if the identification of the specimens is doubtful. Ideally, specimens should be collected and shown to a naturalist for identification.

The semantic reconstruction normally consists of studying all occurrences of a word in order to arrive at the historical meaning, if necessary showing the different mechanisms involved in the evolution of the meaning. Here we have studied essentially the meaning of the word in the ethnoichthyological classification, that is, the biological referent. We have limited ourselves to generic names as the specific names are few and excessively varied. However, as the taxa of the intermediate level<sup>19</sup> are sometimes designated by names identical with those of the genus, certain reconstructions in fact refer to two categorical ranks. For example, in Punu *ngɔlə* designates the common silurid “*Clarias pachynema*”, but also the family consisting of all fish ‘without scales, without bony fins and long in shape’; each member of this family of course has a generic name (*ndúmi* “*Clarias lazera*”, *mbási* “*Heterobranchis longifilis*”, *ngótù* “*Clarias gariepinus*”).

Synchronically, items analogous in form almost always designate the same biological referent, except within a minority of biological families, where they may refer to different, though related, genera. In principle, when there is no absolute identity of biological referent, we set up different series. Likewise, when certain simple items have a relation to complex items, the link is pointed out, but we avoid considering them as a single series. Indeed, if the evolution of the vocabulary involves, as well as borrowing, phenomena of composition,

---

17. The estimated number of languages spoken in Gabon is about fifty, divided into ten linguistic groups.

18. The historical data for each language, the virtual reconstructions and the corpus of fish names can be found in Mouguiama-Daouda (1995) and on the DDL website ([www.ddl@ish-lyon.cnrs.fr](http://www.ddl@ish-lyon.cnrs.fr)).

19. This level of popular classification corresponds on the whole to that of “family” in Linnean classification.

extension of meaning, etc., the difference in the use of these processes is relevant for the understanding of certain genetic relationships or possible contact phenomena.

Many fish names have a formal and semantic relationship with the names of other animals (birds, mammals, snakes). The existence of such semantic fields is mentioned in so far as they may make it possible to establish the origin of certain items.

The presentation of the roots studied follows the alphabetic order of the families to which the species designated by the root belongs. The Latin name of the family is in bold, the binomial is in italics, the root, tones and language groups enabling the reconstruction are in bold. When a group is placed between brackets, this means that, in at least one of the languages, the cognate is irregular and/or that the distribution is geographically restricted. We systematically examine the roots found in several groups where there are problems for some of them.

## 1. **Amphilidae**

### 1.1 *Doumea typica*

Wumbu °tétà, Pinzi °tete, Wanzi and Duma °tèná, B 60, B 70 °tenga are too different from one another to go back to the same root.

## 2. **Anabantidae**

### 2.1 *Ctenopoma kingsleyae*

\***pete** (HH), B 20, (B 30), B 40, B 50, B 60, B 70

This root comes from the virtual reconstructions for all the languages of these six groups: Teke °pede, Ndasá °pede, Kota °peta are variants. As there are no cognates in northern B 30, Tsogo and Evia may have borrowed the corresponding term from neighbouring languages.

## 3. **Bagridae**

### 3.1 *Chrysichthys nigrodigitatus*

\***combo**, B 10, B 30

This root comes from the virtual reconstructions for Mpongwe and Tsogo; the virtual reconstructions differ slightly, °combo in B 10, °comba in Tsogo. As this term is part of a complex lexeme in Mpongwe (*òzòmbò nkémbè*), the second element of which designates the following species, it is not unreasonable to think that the same has been the case in Tsogo. This could explain the difference in vowel quality, which might be the result of a coalescence of the vowel of the first term and that of the connective.

+*bungusu* is attested in Sake and in Shiwa; in Sake, *b* is irregular (\*b>bv/\_\*u).

### 3.2 *Chrysichthys ogoensis*

\***kembe** (HH), B 10, B 30, (B 40), (B 50)

This root is derived from the virtual reconstructions for almost all the languages of groups B 10, B 30 and B 40. Although Nzebi has a virtual reconstruction identical with those of these three groups, this is probably a case of borrowing, as the term is unknown in the other languages of B 50. On the other hand, in Kele the tones are incompatible with the language's path of evolution. In Sake also there is a problem, as the corresponding term has a prenasalized consonant in C2: in this language the Proto-Bantu voiced prenasalized consonant has become devoiced (\*mb>mp); this may be a borrowing from Kande, all the

more so as no other B 20 language has a corresponding form. As Lumbu and Vili<sup>20</sup> have a term deriving from another series (+*mbile*), it is not impossible that B 40 has borrowed the term from B 10 or from B 30.

\**kombɔ* (HL), B 30, (B 60)

This root comes from the virtual reconstructions for Pove and Ndumu; the term found in Kanigi (+*kɔmɔ*) has no etymon in Proto-Bantu (\*ɔ-ɔ>o-o).

\**kumba* (HL), B 20

This root comes from the virtual reconstructions for Mbangwe and Wumbu.

\**kona* (LL), (B 20), B 60

This root comes from the virtual reconstructions for Mbaama, Kota and Ndasā. However, in the last two languages the vowel co-occurrences are not unquestionably regular.

Clearly, we have here variants that differ not only in vowel quality, but also in tone pattern. There are terms in Pove corresponding to the variants \**kombɔ* and \**kembɛ*. In this same group B 30, there are similar phenomena for other species (e.g., *modyenge* “*Barbus batesii*” as against *modyngɔ* “*Barbus holotaenia*” in Tsogo. As the change \**ε-ε>ɔ-ɔ* has been shown for this group (Van der Veen 1991), it is reasonable to think that it is productive in the ethnobiological vocabulary and has served to overdifferentiate the species.

The form \**kombɔ* spread to the neighbouring language Nzebi, then to Ndumu and Kanigi, at a quite early period, as the rule of prenasal simplification was still active in these languages; later, the term spread to Wumbu, where it underwent a modification in vowel quality. Apart from the considerable difference between variants, the existence of another form in Duma, Wanzi and B 20, which are not in direct contact with Pove, strongly favours the hypothesis of borrowing.

### 3.3 *Chrysichthys thysi*

The reconstructions °*tudi* for Tsogo, °*tudu* for Pinzi, °*tute* for Punu are notably different. The origin of the items that led to these forms is obscure, the more so as the corresponding term in B 10 is irregular (+*tore*).

°*baka*, (B 20)

This virtual root is that of Wumbu, but the relation of *w* to \**b* (+*waka*<°*baka*) is not entirely regular. +*ɓaya* in Kanigi is close, but *ɓ* in this language does not go back to any Proto-Bantu form.

### 3.4 *Parauchenoglanis pantherinus*

\**kɔkɔ* (LH), B 40, B 50, B 60

This root comes from the virtual reconstructions for B 40 (minus Sangu), B 50 (Duma, Nzebi) and B 60 (Ndumu).

\**koko* (HH), B 20, B 30, B 50, B 60

This root comes from the virtual reconstructions for Wumbu, Mbangwe, Sake, Tsogo, Pove, Evia, Duma, Kanigi. A stem corresponding to this series is found in B 10 (+*koyɔ*), but the presence of *k* in C1 goes against the principle of strong/weak distribution and suggests that this form is post-Proto-Bantu. The form in Fang-Ntumu (°*ngongo*, LL) is different.

---

20. Vili belongs to group H 10, a group whose affinities with B 40 are well-known.

Even if these two variants are different, given their distribution it must be admitted that they are relatively old.

Another element supports this hypothesis: in various languages, a very similar term means “cat”. We find *mukoyu/mikoyu* (Sangu), *mukɔyɔ/mikɔyɔ* (Duma), *okoyɔ/ekoyɔ* (Ndumu), *bokoyɔ* (Pove, Simba). If the class pairing is different (classes 3/4 for “cat”, except in Pove and Simba, where we find class 14, classes 9/2 for the fish, everywhere), the segments are identical. It is possible that, starting from the earlier name for “cat”, languages developed different fish names, by varying vowel qualities. Subsequent borrowings between dialects have helped to obscure the direction of the derivation chosen by each language.

#### 4. Carangidae

##### 4.1 *Caranx hippos*

\*kaba (LL), B 10, B 30

This root comes from the virtual reconstructions for B 10 and Evia; an item in this series was found in Vili. This could be connected with Portuguese *kabala* “mackerel”; the borrowing would have been made by the coastal languages, Vili and Mpongwe. The dropping of a syllable in words of foreign origin is well-known in Bantu languages. The word would be all the more easily adopted as the segments (*k, a, b, a*) are generated by ‘non-modifying’ rules. The evolution \*b>w took place in Mpongwe after the 17th century as the semi-vowel *w* does not appear in the lists collected at that time.

#### 5. Channidae

##### 5.1 *Parachanna obscura*

\*pidi? (HL), B 20, B 30

This root comes from virtual reconstructions for Ngom, Sake and Pove; the difference between the virtual reconstructions (°pido in Sake, °pino (HH) in Pove) is not great. The colouring of the fish is like that of a viper; and indeed, in Ngom the name is a compound, +peyu a manga, literally “viper-of-water”. Differences occurring in V2 could therefore be explained as the result of contact between the vowels of the two parts of an earlier compound. Comparable items occur in A 83 (Shiwa) and B 50 (Nzebi of Koulamoutou). The centre of gravity for this form is situated between the Centre and the Northeast. The relation of this reconstruction to \*pidi, Puff adder “*Bitis arietans*” is clear. If we assume that the name for the reptile was extended to the fish (Mouguiama-Daouda 1999), it is probable that the process took place independently in each of the languages. We would have, then, a particular form of convergent development. The etymological weight of this reconstruction is thus slight, all the more so as some of the languages in the series are in the same geographic area: Pove, Nzebi and Ngom are all spoken in Koulamoutou or nearby.

\*toba (HH), B 10, B 40

This root comes from the virtual reconstructions for all the languages of B 10 and for Punu, Eshira and Sangu. The corresponding term in Nzebi (*múrùbà*) is incompatible with the rules of evolution of the language: *r* does not go back to a Proto-Bantu phoneme; this is a case of recent spreading from B 40.

## 6. Characidae

### 6.1 *Alestes macrophthalmus*

\***anga** (HH), B 20, B 30, B 50, (B 60)

This root comes from the virtual reconstructions for the languages of the four groups, with the exception of Ndasá, Kota, Wumbu, Evia, Wanzi and Nzebi, which do not have terms corresponding to this root. It goes back to a post-Proto-Bantu period, as medial *ng* is irregular in Ndumu and Mbaama, where \*ng>ŋ. This period must have been fairly recent, as the other languages of B 20 and B 40 have different terms. The absence of corresponding terms in zone A confirms that we have here terms that are not very old.

\***kundu** (LH), B10, (B 20), B 30

This root comes from the virtual reconstructions for the languages of B 10 and of Evia; there is a corresponding term in Vungu (*muyundu*). The virtual reconstruction of Kele is °kunda (>+unda), perhaps a variant of \*kundu; however, the °k must have dropped before fricativisation (\*k>kf/\_\*u).

### 6.2 *Brycinus kingsleyae*

\***baka** (LL), A 75, B 20, B 30, B 50, (B 60), (B 70)

This root comes from the virtual reconstruction for Fang-Ntumu and all groups of zone B, except B 10 and B 40. No items correspond to this root in Wumbu, Mbangwe, Tsogo, Evia, Wanzi, Nzebi, Kanigi and Ndumu. Sangu, the only language in B 40 not having a form corresponding to the root \*panzi (LH) which characterizes this group, may have borrowed its term (*múbakà*) from B 30, especially since it has divergent virtual tonal reconstructions (°baka, °HL, °HH). Mpongwe has a form with a *b* not preceded by a nasal in C1, and a *k* in C2, which does not conform to the language's expected development from Proto-Bantu. The items in Mbaama and Teke have a nasal integrated into the stem, in classes 3/4; this may indicate a borrowing from B 20, where the item is in classes 9/2 (with nasal).

\***gondo**, (B 60)

This root comes from the virtual reconstructions for Ndumu and Kanigi; in B 60, *nd* coming from Proto-Bantu in C2 is not regular (\*nd>n). Mbaama has a form going back to \*baka, so the +*gundu* in Ndumu and Kanigi could be a recent borrowing. This term is found in B 30 and B 10 designating *Alestes macrophthalmus*.

### 6.3 *Brycinus longipinnis*

\***padabada** (LL), B 10, B 20, B 30, B 40

This root comes from the virtual reconstructions for most of the languages of the four groups; Ndasá, Sake, Kele, Tsogo and Sangu are the only languages without a term corresponding to this series. Evia has a term with an *r* in C2 which cannot be of Proto-Bantu origin: it is a borrowing from Eshira; so Evia has replaced the stem occurring in the other B 30 languages (+*βale*) by that found in Eshira (+*aβara*); the °bate of Fang-Ntumu is a rather different variant.

### 6.4 *Brycinus schoutedeni*

\***ba** (L), (B 40), B 50, (B 60)

This root comes from the virtual reconstructions for B 50, Punu and Kanigi. There is no corresponding form in the other languages of B 40 and B 60; Punu and Kanigi could therefore have borrowed it from B 50.

**\*kema (HH), B 20, (B 60)**

This root comes from the virtual reconstructions for Sake, Mbangwe and Mbaama. The corresponding term in Fang-Ntumu (*nk'émé*) has an incompatible segment in C1 (*nk'*); in Ndasa, the vowel co-occurrence is irregular (*mukema*). Nonetheless, \*kema has its centre of gravity in the North with corresponding forms in Bekwil, Fang-Atsi and Kota spoken in Okondja.

**\*badi (LL), B 20, B 30, (B 60), (B 70)**

This root comes from the virtual reconstructions for B 30, Wumbu, Ndumu and Teke. There are traces of the root in a complex word, °kembadi (>+*kemǎle*), in Kele, as also in Shiwa (+*kemele*); this supports the hypothesis of a northern centre of gravity.

There is a striking difference in terms for *Brycinus schoutedeni* in B 60, which is to be explained by spreading: Kanigi has borrowed \*ba from B 50, Mbaama \*kema from B 20, and Ndumu \*badi from B 20. The origin of the word in B 70 (+*bayi*) could be explained in the same way.

6.5 *Bryconaethiops microstoma***\*cɛmbeke (LHL), B 20**

This root comes from the virtual reconstructions for Mbangwe and Wumbu. It is unclear how we might link it to Sake +*kamba*; the vowel co-occurrence and *mb* are both irregular.

**\*danga (LH), B 40**

This root comes from the virtual reconstructions for Eshira and Sangu; the corresponding term in Evia is irregular in the virtual reconstruction of its tones (°HH).

**°coda, (B 50)**

This virtual reconstruction is that for Duma. The consonants have a certain resemblance to those of Fang-Atsi *kirà*, Shiwa *tsirà*, and Kele *kítà*, with a *t* in C2 which is incompatible (\*t>l, \*d>d).

## 7. Cichlidae

7.1 *Hemichromis elongatus***\*tindi (HH), B 10, B 30**

This root comes from the virtual reconstructions for B 10, Pinzi, Evia, Kande and Tsogo. There is a corresponding term in Kele, but in this language there is no *t* from Proto-Bantu in C1 (\*t>ts/\_\*i); the virtual tone reconstruction is also different (°LH). It is, then, a recent borrowing from the adjacent languages of B 30. There are corresponding terms in B 40 (Punu, Sangu, Vungu, Eshira) but the virtual reconstructions are different (°tende, °tendi). Taking the terms found in B 40 back to \*tindi would mean violating the rule of spirantisation of Proto-Bantu \*t, which is operative in this group. So the term is of recent date, when this rule was no longer productive.

**\*kaka (LL), B 40, B 20**

This root comes from the virtual reconstructions for Punu and Ndasa. The existence of °kaka in Punu supports the hypothesis of borrowing of the synonymous root °tende from the neighbouring languages (B 30). Also, °kaka is attested in both Lumbu and Vili, while +*tindi* is not known there.



**\*keɛ (LH), B 60**

This root comes from the virtual reconstructions for Ndumu and Kanigi; one might consider it as a variant of \*kaka but as well as the vowel quality the tone pattern is different.

7.2 *Tilapia heudelotii***\*kondo (HH), B 10, B 30, (B 40), B 50**

This root comes from the virtual reconstructions for B 10, B 30, Punu, Sangu, Wanzi and Nzebi. The item in Kele (*ákúndù*) has a tone pattern HL which is incompatible with a high tone in the prefix; in Fang-Ntumu (*ékóndé*) there is no rule that can explain the sequence *o-e*. Although it is regular as far as segments are concerned, the status of \*kondo in B 40 raises problems. The reconstruction gives a doublet with another term that is well represented in the other languages of B 40 and H 12 (*+yale*). So Punu as well as Kele will have borrowed the term from B 30 or B 50. The Fang term may come from the same centre of diffusion, but there is another possible route, as \*kundu (and also \*bundu, reconstructed for the same species in B 20) have correspondences in Duala (*ikokondo* and *dibundu*).

**\*bɔɔ (LH), B 60, B 70**

This root comes from the virtual reconstructions for Mbaama, Ndumu and Teke. There are two variants in B 20: \*bɔta for Ndasá and Ngom, °podɔ for Mbangwe: as they are rather markedly divergent and as the other languages of B 20 (not located in the Southeast) have different terms, diffusion from the groups B 60 and B 70 is likely. An irregularity in the vowel sequence and the tones in Kanigi could be accidental. In any case, this reconstruction is clearly restricted to the Southeast.

**°kaka (HL), B 50**

This is the virtual reconstruction for Duma. The corresponding stem in Wumbu (*+kaya*) is incompatible: the virtual tone reconstructions are different and there is no etymological *y* in C2 (\*k>k, \*g>k).

°pada for Eshira and °ada for Punu could have a common origin, but the tones do not correspond to the same virtual reconstructions (°LL for Eshira, °HL, °HH for Punu). Evia has the same form as Eshira, but the virtual tone reconstruction (°LH) is different.

## 8. Claridae

8.1 *Channallabes apus***\*tombi (HH), B 10, (B 30), B 50**

This root comes from the virtual reconstructions for B 10, Wanzi and Nzebi; similar terms are found in Vungu and Wumbu. The Sangu term has an *mf* as C2, which cannot be explained by any rule of derivation from Proto-Bantu (\*mb>mb/\_\*i, \*mb>mf/\_\*u). Or could there perhaps have been a period when fricativisation was generalised? Whatever the origin of this word in Sangu, it appears after the period of the reference protolanguage. The virtual reconstruction for Pove is identical with those for B 10 and B 50, but as the group B 30 is generally fairly homogeneous, the fact that there is no corresponding word in the other languages suggests that Pove has borrowed its form from Nzebi. It should be noted that this root refers to a species that is not common in Gabon. It may be old, as is suggested by the existence of a corresponding term (*mulombi*) in Duala, for the “moray”, a species much like the *Channallabes apus*.

8.2 *Clarias gariepinus*

\*ngɔɔ (LH), B 40, (B 30)

This root comes from the data of B 40 (minus Sangu, Eshira), Evia, Pinzi and Tsogo. Terms corresponding to the series are also found in Wumbvu and Vili, whereas they are absent in Kande, Wanzi and Duma, languages which are not in contact with B 40. It is therefore a root peculiar to one zone, in fact to one group, B 40; it has spread to B 30 and B 50. The virtual tone reconstructions for Nzebi (°LL) and for Pove (°FL), which are different from Punu, like an irregular *t* in Kele, support this hypothesis.

8.3 *Clarias lazera*

\*dome (HH), B 30, B 40, (B 50)

This root comes from the virtual reconstructions for B 40, B 30 and Nzebi of Mbigou; in B 10, the extension of Meinhof's rule requires an *n* as stem-initial consonant rather than *nd* (+*ndumi*); for instance, Proto-Bantu, \*dómé (ps 193, "male, husband") gives +*nowe* (ðnówè).

\*comi (HH), B 20, B 50, (B 60)

This root comes from the virtual reconstructions for B 20 minus Ngom and Kele, from Wanzi and from Mbaama, where the vowel sequence *o-i* is irregular in the stem +*somi* (\*o-i>u-i).

\*como (HH), (B 20), (B 60)

This root comes from the virtual reconstructions for Kele and Ndumu; the tone-pattern LH for Kanigi does not correspond and the virtual reconstructions for Ngom are rather different (°cimo, °cimɔ, °cimu). Items linked with this series are found in Fang-Atsi, Shiwa and Bekwil.

So for *Clarias lazera* there are three variants identical as far as tone goes, but with differences in vowel quality. They correspond to zones: \*comi for the East, \*como for the North and Northeast, \*dome for the Northwest, South and Centre. The original vowel sequence could be \*o-i, a vowel shift favouring the change to °o-o and °o-e. We have no explanation for the divergence in the initial consonant. In any case, the species in question is regarded as the largest freshwater fish, and some informants claim that it can reach the size of a man. Given this, it is not surprising that its names are subject to expressivity. There are corresponding terms in neighbouring groups showing further divergences in the segments (*mbumi* in Kunyi (H 13)).

8.4 *Clarias pachynema*

\*gɔda (LL), (B 20), B 40, B 50, B 60, B 70

The distribution is fairly general in the five groups, except for B 20, where Kele and Kota of Mekambo do not have terms derived from this root; nor do Kande, Pinzi and Tsogo in B 30. Its centre of gravity is rather to the South and Southeast. In the North and Northwest there are different terms. But for Fang, which is generally classed with the languages of the North, there are virtual reconstructions °ngono, °godi, °gɔngɔ, °gonoɔ, and for Evia and Pove, generally classed with the Centre languages, we have °gonoɔ, °gono, with tone pattern °LH. It may, then, be supposed that there has been recent spreading in these languages, Evia having borrowed its term from Eshira, Pove from B 50. The B 20 languages of Haut-Ogooue and Ogooue-Lolo have likewise borrowed their terms from B 50, B 60 and B 70.

+*nzangwe* occurs in Sake, but the vowel sequence (*a-ε*) is irregular; this term is close to Bekwil +*nfangə*.

### 8.5 *Heterobranchis longifilis*

\***baci** (HH), B 20, B 40, B 50, B 60

This root comes from the virtual reconstructions for almost all the languages of the four groups. Only Sake and Nzebi do not have items corresponding to this series. The distribution of \**baci* is almost identical with that of \**goda* (LL), with a centre of gravity more to the South and the Southeast. It can easily be shown that the corresponding terms in Pove and Evia, the only B 30 languages where such terms occur, come from B 40 or B 50, as is implied by the class pairing 3n/4n. Van der Veen (1991) showed that the words with this pairing come from the neighbouring groups. As for Fang, the virtual reconstructions for the segments are very different (°n+pembe, °n+binda, etc.). As there is no etymological *z* in Teke, the origin of its term (*ombazi*) is uncertain.

## 9. Clupeidae

### 9.1 *Ethmalosa fimbriata*

\***bete** (HH), B 10, B 40

This root comes from the virtual reconstructions for B 10, Eshira and Punu. Similar forms are found in various groups, but the virtual reconstructions are rather different. In Fang-Ntumu there is the variant °bede, in Kota °beda, in Sake °peta, °pete. The *r* found in Evia cannot go back to Proto-Bantu. The number of variants and the differences they show on the one hand, the incompatibilities of the Evia form (*mbèrè*) on the other, strongly suggest the hypothesis of borrowing. This is, in fact, a sea species; it might seem surprising that populations living a long way from the coast should have a name for it. Is it possible that they were in contact with the sea at some point in their history? Only a few oral stories in B 30 claim a coastal origin. However that may be, there is a much simpler explanation, based on the economic importance of this fish. Being the cheapest species, caught and smoked on the coast, the sardine is marketed even in the remotest villages. The name, probably originating on the coast, must have followed the trading routes of the fish.

## 10. Cyprinidae

### 10.1 *Barbus batesii*

\***dodo** (LH), B 40, B 50, B 60

This root comes from the virtual reconstructions for Kanigi, Ndumu, Punu, Wanzi and Nzebi. The analogous terms in Fang-Ntumu and Wumbu are irregular, their tones going back to °HL. In Ngom, we have °yoyoko, implying the addition of a syllable and the loss of occlusion, with palatalisation: it would seem difficult to link °yoyoko with \**dodo*.

\***ɲɲji** (LH), B 20

This root comes from the virtual reconstructions for all the languages of B 20 except Sake. A similar term is found in Galwa (+*ndzundzu*), but its virtual reconstruction is considerably different (°junju); the Mbaama term (+*dzuoni*) is irregular because of the vowels.

**\*jenge (LL, LH), B 30**

This root comes from the virtual reconstructions for Tsogo, Pove and Evia. The corresponding term in B 10 (+*dyenge*, classes 3/4) is irregular, as initial *dy* is justified only in classes 10a or 15.

10.2 *Barbus compinei***\*boka (LH), (B20), B 30**

This root comes from the virtual reconstructions for Kande, Pinzi, Pove and Kele. The corresponding term in B 10 (+*mboka*) is incompatible, as no Proto-Bantu phoneme gives a *k* in C2 (\**k*>ɣ, \**g*>ɣ): borrowing must be assumed. A corresponding term is found in Shiwa and in Bekwil, with the same meaning, but one in Vili is the name for *Polynemus quadrifilis* (21.1).

**\*geni (LL), B 40, (B 50)**

This root comes from the virtual reconstructions for Punu and Eshira; there is a similar term in Lumbu. Unknown in the other B 30 languages, the corresponding term in Evia (*mòyèni*) is incompatible, as the only Proto-Bantu vowel giving *i* is \**i*, but in C2 the Proto-Bantu sequence \**ni* gives *ny* (\**n*>*ny*/*\_i*); the virtual tone root (°LH) is also different. The corresponding term in Nzebi of Mbigou does go back to °*geni*, but it could also be a borrowing from B 40, as it does not exist in the other B 50 languages.

**\*koda (HH), (B 20), B 60**

This root comes from the virtual reconstructions for Ngom, Kanigi and Ndumu. There are corresponding terms in Ndasa, Duma and Wumbu, but the vowel sequence *o-a* cannot be accounted for by any rule (\**o-a*>*u-a*). In Teke there is an *r* as C2 which is irregular and no variant °*kuda* would be possible, as \**k*>*pf*/*\_u*. The area of distribution of this root is restricted to the East.

**\*ngongo (LH), B 30, (B 40)**

This root comes from the virtual reconstructions for Kande, Pove and Eshira; a corresponding term was found in Vungu. The southernmost languages of the group B 40 (Punu, Lumbu) do not have this root. This may be a case of shared innovations favoured by a semantic extension. And in fact, in zones B and C, Bastin (1994) reports similar stems for the “lion”, a voracious carnivore like *Barbus compinei*, one of the biggest freshwater fish.

10.3 *Barbus holotaenia***\*jonji (LH), B 20, (B 40)**

This root comes from the virtual reconstructions for Ngom, Wumbu and Sangu. That this is an old term in B 20 is confirmed by the fact that it occurs in other languages of the group to designate a closely related species, *Barbus batesii*. The term spread recently to Sangu, as can be seen from the difference in its virtual reconstruction (°*conji*).

10.4 *Labeo sp.***\*tungu (HH), B 10, B 20, B 30, B 40, (B 50), (B 60)**

This root comes from the virtual reconstructions for all the languages of groups B 10, B 30 and B 40, but also for Mbangwe and Wumbu of B 20, for Kanigi and Ndumu of B 60 and for Wanzi and Duma of B 50. However, in B 50, *r* deriving from a Proto-Bantu phoneme is rare (Wanzi). In B 60, *r* is irregular and comes from \**d*, in Kanigi, moreover, *ng* is a problem in C2 (+*rungu*), as in this context the rules require a simplification (\**ng*>η).

10.5. *Raiamas buchholzi*

## °pɛni (LH), B 40

This virtual reconstruction is for Punu; °bɛdi (LH) is the Eshira variant, in which the consonant in C2 is different. In the whole of B 40, only Sangu is without a corresponding term in this series; +pengi in Nzebi of Koulamoutou is similar, but the difference in C2 and vowel quality once again points to contact phenomena between the two groups.

## \*yoyo (B 10), (B 30)

This root comes from the virtual reconstructions for Galwa, Pinzi and Pove; in Pove, a class 19 prefix (*pi-*), diminutive, is added. °dodo in Mpongwe, for the same fish, is close but the classes are different (5/6 for Mpongwe, 3/4 for the other languages), as well as the consonants. If °dodo goes back to \*yoyo, we have to accept palatalisation in three languages, and a change of class for Mpongwe. All this seems very complicated. On the other hand, we know that the Mpongwe virtual reconstruction exists in other languages, where it denotes *Barbus batesii* (\*dodo, LH, B 40, B 50, B 60). So, it may be that in Mpongwe a name which originally designated *Barbus batesii* has been transferred to *Raiamas buchholzi*.

It is tempting to associate °yoyoko *Barbus batesii* with \*yoyo *Raimas buchholzi*, considering *ko* to be a later development. °yoyoko is the root found in Ngom for *Barbus batesii* where other languages have terms related to \*dodo; we considered that there were two different roots (cf. species 10.1). If, on the other hand, we link °yoyoko *Barbus batesii* with \*yoyo *Raimas buchholzi*, the implication is that °yoyoko, which originally denoted *Raimas buchholzi* has been transferred to *Barbus batesii*. This is the reverse of the Mpongwe situation, where the term for *Barbus batesii* would have been adopted for *Raimas buchholzi*.

In the end, the safest hypothesis would be to consider that there are two different roots for *Raimas buchholzi*: \*yoyo (Galwa and B 30) and °dodo for Mpongwe. In this case, we assume that the Mpongwe root originally designated *Barbus batesii* and that °yoyoko *Barbus batesii* in Ngom has no connection with \*yoyo *Barbus batesii* in Galwa and B 30.

## \*ganga (LL), (B 20), (B 50), (B 60)

This root comes from the virtual reconstructions for Mbangwe and Wanzi, and perhaps Ndumu – but in this language, *ng* coming from Proto-Bantu is rare in C2. In Wumbu (*ngàngà*) and Duma (*ngààngà*), one would expect *nk* in C1 (\*ng>nk); the virtual root for Ndasu is slightly different (°kanga). Finally, this root is restricted to the Centre and South-east, not found in the less easterly languages of B 20 (Sake, Kele, Ngom) and B 50 (Nzebi).

## 11. Cyprinodontidae

11.1 *Epiplatys sexfasciatus*

## \*bongo (HH), B 20, B 30, B 40, B 50, B 60

This root comes from the virtual reconstructions for Kota of Mekambo, Pinzi, Evia, Punu, B 50 and B 60. In Mpongwe, the *b* of +*bongo* does not fit, as for C1 the evolution of Proto-Bantu requires a *w*. Likewise, in Fang-Ntumu the rules require *mv* (\*mb>mv, \*mp>mv), whereas the word is +*mbonj*. Given its wide distribution, this term is probably very old. Although the terms in Fang and Mpongwe are irregular they would also be old, without going back to PB-X; it should be noted that corresponding terms are recorded in the neighbouring zones.

## 12. Dasyatidae

12.1. *Dasyatis ukpam*

°n+yuma, B 10

This virtual reconstruction is that for Mpongwe; +*nowa* in Kande has irregularities: in this language, there is no *w* derived from Proto-Bantu. Furthermore, *ny* in Mpongwe (+*nyuŋwa*) goes back to \**ny*, \**n+i*, or \**n+y*, whereas in Kande *n* goes back to \**n* in C1. As comparable terms occur in Duala (*duba*, *ebodume*), these could well be old names.

## 13. Distichodontidae

13.1 *Distichodus fasciolatus*

°tɔɔ (HL), B 50

This virtual reconstruction is that for Duma; corresponding terms are found in Ndumu and Kanigi, but in these languages there is no *p* derived from Proto-Bantu in C2 (\**p>b*). The virtual tone root for Kanigi (°LL) is also different.

\*coko (LL), (B 20), B 30, (B 50)

This root comes from the virtual reconstruction for B 30 minus Pove, and for Kele. There is a similar term in B 10 (+*soɣo*) but it cannot go back to Proto-Bantu, as \**c>z* in C1. The term found in Nzebi of Mbigou is also irregular – at least, this is what is implied by the difference of the virtual reconstructions, °joko for the segments and °HH, °HL for the tones. In Wanzi, the virtual reconstruction °koco (LL) suggests metathesis.

13.2 *Distichodus hypostomatus*

\*kɔɔ (LH), (B 20), B 40, (B 50), (B 60)

This root comes from the virtual reconstructions for Eshira, Punu and perhaps Sake. There are corresponding terms in Duma (*mùγɔɔ*) and Kanigi (*mòγɔɔ*), but there is an *r* in C2 which is irregular, as it goes back to \**d*, and the *γ* of C1 has nothing corresponding in Proto-Bantu. In Sake (*mungoru*), there is no virtual reconstruction for the vowel sequence but this could be accidental. Similar terms occur in Shiwa, Vili, Lumbu and Vungu. The distribution of the term and the fact that it is found in languages as far apart as Punu and Shiwa suggest that it is relatively old. As the same term exists in many languages to designate any animal “with curved beak”, it is legitimate to think that there have been convergent developments from one root (Mouguiama-Daouda 1999).

13.3 *Distichodus notospilus*

\*beta koco (LLLL), B 20

This root comes from the virtual reconstructions for Mbangwe and Wumbu. The virtual reconstruction for Ngom °betakuwe may be regarded as a variant, given that compound words are not always as regular as simple ones; moreover, the loss of C2 and the generalisation of *ε* as a final vowel are on-going processes in this language. On the other hand, °bilakwetsi in Kanigi is altogether incompatible.

## 14. Hepsetidae

14.1 *Hepsetus odoë*

\*enge (HL), B 10, (B 20), B 30, B 40, B 50

This root comes from the virtual reconstructions for the languages of all the groups except A 75, B 60 and B 70. An *ndz* in Kota is unusual in medial position, but as the evolution

\*ng>ndz is found in Wumbu, Ndasas and Shiwa it is reasonable to consider that fricativisation of prenasals was a feature common to the whole of B 20 at a given point of time. In Gabon, there are examples also in Shiwa, Wumbu, Lumbu, Vili and Nzebi of Koulamoutou. This species is particularly interesting: it and *Malapterurus electricus* are the only monotypical species, and absolutely pan-African. The chances of semantic change and of diffusion are consequently small.

**\*cumba (LH), B 20, (B 60)**

This root comes from the virtual reconstructions for Ndasas, Mbaama, Ndumu, Kanigi and Teke: In the last three languages, there is an *mb* in C2 which is irregular (\*mb>m).

**15. Malapteruridae**

**15.1 *Malapterurus electricus***

**\*niki (LL), B 30, B 50**

This root comes from the virtual reconstructions for B 30 minus Kande and for B 50 minus Nzebi. °nyikɔ, °nyigɔ are the variants in Ndumu and Kanigi. In Ndumu, there is the difference in the final vowel, but also \*k>ts/\_\*i; it is difficult to take the reconstruction back to \*niki.

**\*nici (LL), B 40, (B 50)**

This root comes from the virtual reconstructions for Sangu, Eshira and Nzebi of Mbigou. As in Eshira there is an evolution \*k>ts, °nici could come from \*niki. So there would have been a development \*niki>°nitsi>°nici (after all, the difference between *ts* and *c* can be very slight in some languages). Once again, we would have a case of Nzebi borrowing from B 40.

**\*nyingi (LL), B 20, B 30**

This root comes from the virtual reconstructions for Kande and for B 20, where there is a variety of forms: °nyinge for Wumbu, °nyenge for Mbangwe, °nyinga for Ngom, °indi for Kota, °inda for Kele and °nyendi for Ndasas. It could indeed be a case of terms coming from two different roots. However, as we have already seen with the terms for *Hepsetus odoë*, an evolution °ng>ndz is likely in some B 20 languages, after a first stage without change (\*ng>ng). It should be noted that this species is taboo: believed to weaken the male organ, it is not eaten in many ethnic groups. So the variety of forms found may be the result of avoidance strategies.

**\*ninda, B 10, B 20**

This root comes from the virtual reconstructions for B 10, Kele, Ngom and Kota. It is considered here that the terms with *nd* as C2, which were a problem with regard to the root \*nyingi above, constitute in fact a different series, to which the term found in Mpongwe would also belong. In B 10 the root \*nyingi also occurs but it designates *Protopterus dolloi* (24.1). The term found in Shiwa would belong to the \*nyingi series, whereas that found in Fang-Atsi (*anya*) goes with the \*ninda series.

## 16. Mastacembelidae

### 16.1 *Caecomastacembelus sclateri*

\*tɔndɔ (HH), B 40

This root comes from the virtual reconstructions for Punu, Sangu and Eshira. The term found in Nzebi, which is not known in the other B 50 languages, has an *r* as C2 which is irregular.

°tɔti, B 50

This is the virtual reconstruction for Wanzi. It is close to the form +tɔɔ in Ndumu and Kanigi, but in these, *r* is irregular, going back to \*d (°tɔdɔ); and V2 is assimilated to V1.

\*kɔngɔ (HH), (A 75), B10, (B 20), B 30

This root comes from the virtual reconstructions for B 10, B 30, Kele and Sake; a variant °kɔnga in Fang-Ntumu is generated by marginal rules; the virtual root for Ngom, °kɔko is rather different. The other B 20 languages have a root that may be regarded as characterising the group and marginalizing Kele and Ngom. Found also in Shiwa, °kɔngɔ is a Northern reconstruction.

## 17. Mochokidae

### 17.1 *Synodontis polyodon*

\*bɔnge? (LH?), (LL?), A 75, B 20, B 30

There are several virtual reconstructions and it is not really possible to choose between them: °bɔnge for Fang-Ntumu, °bɔngi for Mbangwe, °bɔnge for Tsogo and °ponge for Pinzi and Pove. The forms with voiced initial consonant have a virtual tone pattern °LH, whereas those with voiceless initial consonant have °LL. As there is a corresponding form in Bekwil and as \*bɔnge also designates *Parauchenoglanis Pantherinus* in B 20, this is a root which has its centre of gravity in the North.

## 18. Mormyridae

### 18.1 *Mormyrops sp.*

\*puda (?) (HL), (B 10), (B 30), B 40, B 50

This roots comes from variants of virtual reconstructions for almost all the languages of the four groups; only Wanzi of B 50 has no term related to this root. In Kande and Pinzi, the virtual reconstruction is °funa, which might be explained by a recent and non-systematic evolution of *p* and *d*. An *n* as C2 in Mpongwe comes from \*d, but it is a sporadic evolution. In B 40, there is a slightly different variant, °buda (HL). On the other hand, all the virtual tone-roots go back to °HL. This could be an old root, without going back to PB-X.

## 19. Mugilidae

### 19.1 *Mugil sp.*

\*mɔnɔ (LH?), B 10, (B 30)

This root comes from the virtual reconstructions for B 10, Kande, Eshira and Punu, but the segment difference in Evia (°ɔno) and the difference between the tone pattern for Punu and Evia (°LH) and that for Eshira (°LL) suggest that this is a case of diffusion. There is a comparable item in Lumbu and in Fang-Atsi. The term could come from Portuguese *molo* which is also the name for “mullet” (*Mugil sp.*).



**20. Notopteridae**20.1 *Xenomystus nigri*

°gote (LL), B 10, (B 30)

This is the virtual reconstruction for B 10; Evia has a variant °koe.

**21. Polynemidae**21.1 *Polynemus quadrifilis*

\*cena (HH), B 10, (B 30)

This root comes from the data for B 10 and Kande; Evia has a form which would go back to °jena and resembles B 40 forms (+nzʸenə in Punu, +nzʸenə in Eshira) but in these languages the initial consonant (nzʸ) is not traceable to Proto-Bantu.

**22. Pomadasyidae**22.1 *Pomadasys jubelini*

°koete, B 10, (B 30)

The term found in Evia (+kwerε) is analogous, but in this language, *r* as C2 is not of Proto-Bantu origin.

**23. Pristidae**23.1 *Pristis microdon*

\*bago, B 10, (B 30), (B 40)

This root comes from the virtual reconstructions for B 10, Evia and Eshira. The tone-reconstruction for Evia (°LL) differs from that for Eshira (°LH). In Vili, we find +*mbafu*: the final vowel must necessarily go back to \*u, *f* being due to the fricativisation of \*k or \*g before \*u. The Fang-Atsi form implies a vowel sequence \*a-a. Therefore, a certain number of elements favour the hypothesis of successive borrowings. In Evia, there is even a decisive argument: the class pairing 3n/4n is clearly from Myene. The root could come from Portuguese *tubarao*. In Mpongwe, the initial consonants of Proto-Bantu prefixes dropped (\*do>o for example in class 11), so the same rule could explain that the first syllable of the Portuguese word, interpreted as a class prefix, became *o-*; as the difference between the Portuguese *r* and the Mpongwe *ɣ* is not great, the development of *ombayo* from *tubarao* is not implausible.

**24. Protopteridae**24.1 *Protopterus dolloi*

\*bonga (LH), (B 20), B 30, B 40, B 50

This root comes from the virtual reconstructions for Evia, Sangu, Nzebi and Wanzi. In Sake there is a vowel sequence (*u-a*) which cannot be of Proto-Bantu origin (\*u-a>u-o, u-ε). There are comparable items in Fang-Atsi and Shiwa.

\*jɔmbɔ (LH), (B 20), (B 50), B 60, B 70

This root comes from the virtual reconstructions for B 60, B 70, B 20 minus Kele and Kota of Mekambo, and for B 50 minus Wanzi and Nzebi. The virtual reconstructions for Wumbu and Ndasa (°jɔmɔ) suggest that the term may have appeared recently in B 20 and B 50.

## 25. Schilbeidae

25.1. *Schilbe grenfelli*

\*ata (LH), B 10, (B20), B 30, B 40, (B 50), (B 70)

This root comes from the virtual reconstructions for B 10, B 30 minus Pove, B 40 minus Sangu and B 50 minus Duma. Teke has a term in this series with an *r* as C2, which is irregular. In Kele, there is an initial *dy*, which is a reinterpretation of the class prefix found in some B 40 languages for terms designating the same species. In Wanzi (*mutšara*) there is a *ts* which is hard to explain, unless one considers that it is the trace of an earlier class prefix, so there would have been a reclassification. This could be an old root, without going back to PB-X.

\*benje (HH), B 20

This root comes from the virtual reconstructions for Ngom and Ndasa. Comparable terms are found in Kota of Mekambo, Sake and Teke, but a *z* in C2 in these languages does not come from any Proto-Bantu phoneme. If that could conceivably happen in the B 20 languages, as the corpus studied does not make it possible to test the evolution of \*nj and \*j which might generate a *z*, in Mbaama there is certainly an incompatibility, as *z* in C2 goes back to \*j. Mbaama has taken this term from B 20, which explains its absence from the two other languages of the group.

25.2. *Pareutropius debauwi*

+βifi found in Nzebi of Koulamoutou, +wili in kota and +libi in Vili could be connected, if we accept a metathesis for the last form and a \*d giving rise in C2 to *l* in B 10 and Kota and to *f* in Nzebi (fricativisation before a close vowel); finally, one would have to posit a \*b in C1. On the other hand, +byolo in B 10 has an incompatible form in C1, so that one cannot be sure that it is of the same origin. It is clear that the processes involved are not very regular.