

Afrikanistische Arbeitspapiere (1996) N° 45, Köln : pp 27-52

*Proceedings of the Sixth International Nilo-Saharan Linguistics Conference
Santa Monica, 1995, March 27-29*

Problems of Grouping and Subgrouping: the Question of Songhay

Robert Nicolai

GDRE 1172 of the CNRS

Researchers who use the resemblances method don't all work in the same way. Some hold that this approach doesn't permit the development of a valid hypothesis concerning genealogical relationships between compared languages. Rather they believe that it supplies a "presumptive hypothesis" if only because, in many cases, the evidence shows that genealogically-related languages demonstrate resemblances. Handled this way, the technique provides an initial count allowing one to define a set of languages requiring further analysis. However, this operation is not a method. As a procedure capable of describing reality, it has no scientific validity whatever. It is an exploratory practice. In "reality", there are numerous apparently "obvious" structures, which doesn't mean that the regular features they contain and which can be defined by mere observation enjoy any kind of necessary status. However, it is by showing just such a necessary status (even when recognized a posteriori) that one can validate a method. In addition, as Dalby (1966) has already pointed out, it can be dangerous to classify and group languages too hastily since any form of classification involves a limitation which will prestrain the organisation of the data and predetermine the results. This is not too serious when there exist strict procedural rules which allow one to check whether various items belong to a class. However, the situation is very different when these procedures are merely "**evaluative**" and non-falsifiable. Hence the idea that the method allows one to account for 80% of reality and can therefore be used to make a rough outline of a situation is probably a "bad good idea" since, for the eighty per cent, the method doesn't supply significantly different results from more traditional approaches, whereas for the remaining 20%, it provides a non-decidable result.

Yet another idea is often advanced whereby, given the absence of detailed studies concerning all African languages - and it is a fact that there are many languages for which we possess only limited information - the resemblance method is the only one which allows us to integrate them into into a genealogical classification as they require only a

Robert Nicolai

small number of items, a lexical list being sufficient. Obviously this kind of operation can be performed. However, in view of the limitations I have already mentioned, it is extremely difficult to draw sound conclusions using such scanty data. As a result, if these items do point to new directions for research - and this is essential because we never work on the basis of chance alone - so much the better. Nonetheless, this consideration does not entitle us to integrate them into a genealogical structure as though they constituted a "result". Moreover, whilst it is an established fact that a given little-known scarcely-spoken dialect can be useful in as far as it provides clues concerning the genealogical relationships within a given group, at the same time it is well-known that we don't necessarily need to have to hand every item of information concerning every language to be in a position to draw up serious hypotheses regarding relationships. Similarly, it is common knowledge that one needs sound, carefully analysed data rather than voluminous, unreliable information which has been gathered in the manner of an amateur entomologist pinning up his butterflies.

The other underlying idea is, of course, that there exists a "method" allowing one to account for language development and that one can always explain this development by means of refinements to this method combined with additional efforts of interpretation and collection. In fact, the actual practice of researchers shows that "approach methods" often show divergences depending on the family of languages they are handling. Thus, they are sensitive to the empirical nature of the objects to which they are applied. There is no one single method. Rather there exist:

- *some general principles* concerning the modalities of language development and language changes,
 - some empirically established and theorised principles which emerge as research into language development progresses,
 - some logical principles which need to be applied to ensure general coherence in one's reasoning,
- and, at the same time, a store of knowledge built up over time thanks to the implementation of these principles and guidelines and which is used as source material.

Data complexity, rather than simplicity, is the general rule, since the determining features of language growth are not limited to the mechanics of structural system transformations. Many other variables come into play, the potential impact of which should not be neglected. One can well be dealing with situations which do not allow one to advance a serious hypothesis concerning development. One must then draw one's conclusions not only using the possible refinements of the model, or models, but also by

Problems of Grouping and Subgrouping: the Question of Songhay

considering the nature of language changes, the determinants which trigger them and the theories which "explain" them. It is not at all certain that one will always discover genealogical relationships and, in difficult cases, if the use of "weak" research procedures allows one to obtain results, it is not at all sure that these findings will be related to the initial description and that they constitute a mere approximation of what was being sought. In a given set of languages, a model can account for a type of change defined by reference to the conditions which permitted it to develop. If this same set of languages had been determined by another series of conditions, the changes would probably have been different.

From this point of view, the controversy concerning the attribution of Songhay to Nilo-Saharan is an interesting one and the problem becomes even more complex when, having postulated the relationship, one deals with its sub-grouping. Indeed, and on the one hand, this attribution is contested by some workers (Nicolai 1990) who do not adhere to Greenberg's methodology. And on the other, among the researchers who hold this opinion and who accept the attribution, there are major fluctuations concerning its place within the Nilo-Saharan family. M.L.Bender (Forth.1996), for instance, places Songhay in the category of "outliers" or isolates which implies that it broke away at a very early stage. Ehret, for his part, finds reasons to situate it more centrally within the family. His latest findings assert: "*In contrast to most other classifications of Nilosaharan since Greenberg's, Songhay is here not the result of a very early branching within NS. It developed out of a relatively late branch of NS, a branch which it shares with Eastern Sudanic and from which Songhay's geographical Nilo-Saharan neighbour, i.e. Saharan, is excluded*" (Rottland 1994 : 45). How can one come up with such diverging results and what does this all mean? The problem raised involves the empirical data as well as the principles and methods of analysis.

Logic

If one refers exclusively to the principle of continuous division, every classification project implies that as soon as the linguist integrates a given language, rightly or wrongly, into his comparison, he is *logically* postulating the existence of a "family". The hypothesis that the integrated language belongs to the same family has therefore been admitted *a priori*. The hypothesis, as such, *must* therefore be validated or invalidated. However, in the framework of the resemblances approach (which is based upon the evaluation of degrees), there exists no paradigm allowing one to do so since the "resemblance" is asserted or postulated, but not demonstrated. As a result, loose and even suspect resemblance-based relationships can be adopted as it will be supposed

Robert Nicolai

that they illustrate rather than prove more ancient relationships and their possible scarcity will even plead in their favour. Hence, both the failure to discover more resemblances and the greater approximateness of the relationships will justify diachronic depth. In this context, if language X has been "wrongly" integrated into the comparison...the most "normal" course of action, should the analysis of the relationship prove unsatisfactory, is not to exclude it. This is not "necessary" (v. Bender:1991b). Rather its place in the family should be redefined by locating it more or less on the fringes or by assuming a more ancient separation. In this way, it is relegated "asymptotically" to the outskirts of the family.

This is somewhat similar to Bender's position and his concept of "outliers". This approach which, by definition, reverses the classic method with its tendency to "climb back" up the branches of the tree, has a "perverse" effect since one is led - since any hypothesis is potentially the foundation of a taller structure - to establish hypotheses not on the bases of fact (even of "constructed" fact), but on hypotheses...with all the risks that this entails. Bender (1991a) is aware of this danger when he remarks that "*the investigator has to make judgments about plausibility and fitting of overall patterns which emerge as one sifts and resifts the data, while trying to be objective and not force the data to support preconceptions or present a neater picture than indeed exists*". However, the facts show that "awareness of danger" does not preclude "divergent results" as illustrated by the difference between the analyses of Ehret and Bender on the Songhay question. At all events, the constructed genealogical tree is more an expression of the results of this intuitive quest for resemblances and their appraisal by the worker than an account of the stages of development.

Sub-classification

The wish to establish a sub-classification, like the attribution of a language to a family, is also a logical outcome of the general principle of continuous division. One is aware that this operation rests on the identification of the development of innovations which, as they are assumed to have occurred locally and then to have spread further afield (since, by definition one finds the reflexes), allow one to pinpoint the different stages in the separation of languages. It is by means of these innovations that one can perform classifications. Correlatively, it is just as important to identify the forms which show proof of "retention" and which, when found in sufficient numbers, justify the coherence of the entire upper level. These transformations concern both the changes operated on one lexeme by another and the modification of their form and meaning. But, by referring exclusively in this way to the principle of continuous division to analyse the internal

Problems of Grouping and Subgrouping: the Question of Songhay

structure of language groupings, the establishment of a subclassification corresponds to a very strong hypothesis concerning the modalities of the assumed development in the "family". Nothing proves, a priori, that one is necessarily dealing with a language development pattern which occurred by continuous division and without interaction and "catastrophic" reorganisation, which is the only instance which would correctly validate a method which requires or allows one to verify, as the construction is built up, whether the uniformising principle of development is permanently present.

Data

The question of data, for its part, is just as crucial and, in the work of various researchers, one can discover the more or less explicit signs of their vigilance. Bender, for example, noted (1983): "*Numerous attempts have been made at finding Nilo-Saharan isoglosses [...] or those aimed at subunits within Nilo-Saharan [...] . All have been subject to severe criticisms [...] . for reasons of inadequate control of data, questionable interpretations, etc.*" and he went on to emphasize his own commitment to high-quality data : "*My own search for isoglosses (1981) is [...] based on a carefully compiled corpus of the data available, limited to the same choice of items for each language, and by setting up preliminary proto-forms involving the "mesh principle" of mass comparison*". It is clear that Bender and Ehret, whilst reaching divergent results, have developed their hypotheses by paying close attention to the character and validity of the data. And this clearly shows that although the problem is initially rooted in this question, it is later transformed as follows: data, yes, but what type of data, and what are we to do with them?

Indeed, in the present state of the development of Africa-based research and given the resources currently available, what is really essential is not to gather data - in a sense we are already submerged by information - it is what we intend to do with them that matters. What is crucial is not to catalogue them but to assess their usefulness and to ensure that the questions we are asking are in fact the right ones.

Global Hypothesis

Neither all the data nor all the results of analysis are relevant "in the absolute". Their irrelevance depends upon the hypothesis one makes about what is in the process of being demonstrated. Let's call this hypothesis "the global hypothesis". The example of Songhay is a particularly useful illustration as it is the subject of different hypotheses. Thus :

- if one adopts a filiation hypothesis favouring a normal tree-diagram development

Robert Nicolai

in the Nilo-Saharan context; then the identification of data regarding neighbouring languages such as Mande, Berber, etc. will "logically" be related to potential borrowings.

-if one adopts a "mixed language" hypothesis to explain phenomena in Songhay, the data from Berber and other data localised in the same geographic area assume greater importance.

-if, having adopted the "combined language" hypothesis, one decides to process the "borrowings from Arabic", one's analysis will necessarily be different and the form these borrowings take will be potentially crucial regarding considerations of relative chronology. This would not be the case in the traditional approach. Similarly, the relationships established with Saharan languages will have a different value according to whether one has given priority to one or another of the preceding hypotheses.

The pregnancy of the "global hypothesis" will also play a part in the random attribution or non-attribution of certain resemblances. If the global hypothesis one adopts, for example, is the "Songhay-as-a-Nilo-Saharan-language" hypothesis, then the comparison between Songhay and Berber 'pupille de l'oeil / oeil' *emma//mo*, 'bouche' *imi//me*, 'demonstrative' *wo//wa*, monosyllabic lexical units, cannot be entirely random. On the other hand, in the hypothesis claiming that Songhay is the outcome of a "mixed language with a Berber lexical base", the same connection would indicate the presence of a genealogical relationship. The self-evident nature of the "global hypothesis" thus leads one to undertake a discriminating selection of the relevant data and to predetermine one's analysis. Different data will be sought according to whether one subscribes to one or the other of these hypotheses and the same validity will not be attributed to them.

In addition, the pregnancy of the "global hypothesis" concerns "strong" methods such as internal reconstruction just as much as the straightforward selection of data we have just seen. In fact, one always reconstructs from existing data, from a body of knowledge concerning the dynamics of change and by using a priori inference processes. Therefore, the result obtained - which merely constitutes a "hypothesis-based projection" - only corresponds necessarily to a former state of the language when it is self-evident that there has been no "catastrophic" effect which has interfered with the development between present time ($t=0$) and the time ($t=1$) of the reconstructed stage and only in as far as one can justify the "global hypothesis" which is defined by the assumption that one is effectively dealing with a continuous division-type development. In theory, a dialectological analysis might provide an instance of this or when a genealogical relationship has been proven... However, there is no guarantee that this is the case so long as the evidence pointing to this relationship remains doubtful. By definition, the

Problems of Grouping and Subgrouping: the Question of Songhay

function of internal reconstruction is not to establish a relationship. It builds on the existence of the latter in order to advance proposals concerning the former situation.

"Weak" theories and "strong" theories are thus both conditioned by the "global hypothesis". However, in the framework of a "weak" theory, and one backed by a hazy method to boot, this "global hypothesis" is never placed in question.

Non-proof

This question, related to certain standpoints adopted by Ehret (1989ms. p.3), who seems to use these characteristics (the a priori "global hypothesis" and the weakness of data), to justify his results methodologically, is of particularly crucial importance. Thus, on the subject of empirical data likely to testify to "innovations" relevant for language subgrouping, he rightly observes that: *"To wit, a linguistic item or usage [...] can be identified as an innovation only if we can give probable identification to the item or usage it displaced or supplemented-- to the item or usage it was an innovation in place of or supplementary to. The non-presence of an item or usage in a language is not a negative evidence ; it is no evidence one way or the other. There are two possible reasons for an element to be lacking in a language :*

- 1) it may never have been present at any time in the language's ancestry, or*
- 2) it may once have been in the language but have dropped from use at some point of the evolution of that language. The presence of an element in two related languages and its lack in a third thus does not by itself allow us to choose between the two possibilities. It tells us nothing whatsoever about the subgrouping among the three".*

This statement, which basically points out that one cannot prove a non-relationship, is not really very new. A. Meillet (1958), in his criticism of Schuchardt, suggested something quite similar: *"Dans les langues où il n'existe pas de système morphologique comportant des formes pourvues de caractéristiques singulières, la preuve d'une parenté peut être très difficile à administrer. D'autre part, s'il s'agit de langues qui ont beaucoup divergé, soit par suite du long temps depuis lequel elles se sont séparées, soit par suite de la rapidité avec laquelle elles se sont transformées, une parenté réelle peut*

*être devenue indémontrable parce que tous les faits morphologiques communs se sont effacés."*¹ Nonetheless, when referred to a context in which the method is more justified in "confirming" rather than in "proving" postulated relationships, this introduces other effects. For what Ehret is suggesting does not relativize the use of a method by showing

Robert Nicolai

its limitations in the face of data. (One cannot prove everything since there are cases in which language development has "erased" the evidence).

Consequently, and more modestly, we should say that the genealogical relationships between certain languages will never be established. "Realistically", one must observe that they will remain, until such time as we are better informed, and perhaps for ever, beyond the pale of identifiable relationships. Rather, it is the opposite argument whereby, when the relationship or degree of relationship between languages is given a priori, and when the sub-grouping has been completed, it is no longer a matter of merely adding a new language to a set but, on the contrary, of confirming the status of an already classified language. (In this case, the statement is changed as follows: since it is obvious that there are cases in which language development has "erased" the evidence, we must not expect to find empirical justifications for everything: when such justification is lacking, the "global hypothesis" is sufficient to guarantee the relationship. In an "idealistic" assessment, one could claim that the absence of data does not place the construction in question provided that everything fits in harmoniously vis à vis this global hypothesis).

To my mind, this approach constitutes a potential perversion of the research method when it comes to drawing up hypotheses on language development using a weak method - I am not claiming that Ehret succumbed to such - since in that case, given that the relationship was established at the outset, instead of proving it, one "theorizes" the fact that one's inability to find data justifying the hypothesis ...doesn't allow one to invalidate it!

Semantic Chains and "Family Resemblances".

Appraisal of the resemblances and hypotheses concerning the modalities and directions of language change, which enable one to justify possible comparisons, remains the main question involved in the method from both the phonetic and semantic point of view. According to Ehret, this occurs in the three following areas: 1) ordered sound shift histories, 2) Lexical-replacement histories in basic vocabulary and 3) cross-familial semantic histories of reconstructed roots. Concerning this last point, he adds that (1989ms. p.7-8): *“the approach of cross-familial semantic histories requires the identification of a specifiable set of Nilo-Saharan roots. The members of this set have the following characteristic : The reflexes of such a root, as it appears in languages of two or more genetic subgroups, show a shared semantic shift, or evince a common earlier, underlying meaning shift, not found in or inferrable for the root as it appears in one or more of the other genetic subgroups of the family. The existence of a root with the particular shared*

Problems of Grouping and Subgrouping: the Question of Songhay

innovation is a datum for the languages having derived from a common line of descent within the family, meaning shift might possibly reflect separate, parallel innovations, a body of such cases consistent in their mutual implications is needed to sustain the argument”.

Here I will deal more specifically with the semantic level and with the inherent problems which characterize it when determining semantic chains. To illustrate this methodologically crucial issue, I will take the result of a series of comparisons which demonstrate the presence in Nilo-Saharan of a root with the *bVr* form signifying 'branche, lance, bâton' and, less frequently, 'arc' or 'flèche, tige, racine'. Bender (Forth. 1996) has found it in ten out of the twelve Nilo-Saharan families. Ehret places it under reference 45.

Quite independently, and using another methodology, one can group this Songhay entry in another set of relationships, the function of which is to establish a link (other than borrowings) between Songhay and Berber, and more precisely with Touareg (v. Nicolai 1994a). This brings us back to the question of the "global hypothesis". How can we explain the fact that these same items can be used to justify classifications with two different language families when hypothetically they are not supposed to be genetically connected? What does this mean? How can we deal with this problem?

Should the semantic item 'arc' in Songhay be related to 'lance, bâton, racine' as is suggested for Nilo-Saharan, in other words with a continuum stretching from 'plantes' to 'armes faites avec des branches' and corresponding to a well-defined semiotic sphere? Or should it be related to the set including 'flèche, os, esquille' as laid down by internal lexical comparison in Songhay which, moreover, bases it at the verbal level on the apparently fundamental notion of 'piquer, crever, transpercer'?

Methodological principles

Ehret (1989ms. p.9) points out the importance of semantic innovations in the structuring of sub-groupings and postulates the following "dynamic chains":

- “(1) A concrete meaning is historically prior to an abstract, metaphorical, symbolic, euphemistic, descriptive, or attributive meaning for a root;*
- (2) A non-technical meaning for a root is prior to a culturally, economically, or technologically specific meaning;*
- (3) A general or broad meaning for a root is usually prior to a specialized or more narrowly focussed meaning.”*

Robert Nicolai

One must, however, ask oneself to what extent the "logic" of conceptual structures - which is possibly valid for a structuring and synchronic representation of language - has a place in the rules of its diachronic transformation. Are the logics of development and its constraints the same as the classificational and organisational logics of the human mind and is there only one type of "classificational" logic?

Several problems are involved here. The first is conceptual and concerns the need one feels to lay down rules to justify hypotheses regarding semantic changes. The second is formal and concerns the difficulties involved in establishing these rules or at least of granting them a universal or categorical character. Rottland (1994), for instance, notes that *"For Ehret (3) is 'more a guideline than a dictum' which is based on the 'logical unidirectionality' of the link between two meanings."* At the same time, this weakening renders them inapplicable since they are no longer the expression of a rule but only a possibility. Calling upon them merely suggests that the fact of invoking them does not constitute a source of improbability in the elaboration of the hypothesis...but the fact of using these non-unlikely features is much too weak to justify a hypothesis.

The third is empirical and concerns the need for a detailed preliminary semantic analysis of the lexical fields in the languages being covered and, more precisely, in the distinctive features of their usage.

Discussing Ehret, Rottland acknowledges a certain validity in the criterion of semantic innovation for the establishment of a group. He also observes that: *"The difficulty lies in establishing for each individual case whether or not we have a semantic innovation and if so, what it consists in. Given the fact that we have to rely entirely on inferences from present-day languages this can never be established beyond doubt. I therefore agree with Ehret in relying on probability as a guideline. The next step should be to establish the semantics of the individual attestation in greatest detail"*.

This is precisely the question. The general principles concerning frequently observed regularities and logical inferences (bringing one back to a fall-back position based on typology, logic and cognition) are one thing. They provide justification for new directions of research. However, the work of justification remains ahead and still needs to be done in detail.

The practice of comparison binds together the need to refer to principles of this kind, the need to remain coherent with the "global hypotheses" regarding the relationship between the compared languages, and the need to justify, as precisely as possible, the details regarding item shifts.

Problems of Grouping and Subgrouping: the Question of Songhay

As a first conclusion, one can observe that in certain types of languages which do not lend themselves readily to the classic form of comparative analysis and for which lexical material becomes important, it is necessary to set up "safety guidelines" in the quest for relationships.

Three methodological principles are of importance here:

1) A refusal to use lexical items as opaque materials and the need to process them as language objects.

2) The need to *justify* the "transition steps" which are proposed for the establishment of a "chain of relationships" (A resembles B, B resembles C, C resembles D, does D resemble A? v. Nicolai 1994a). Building a "chain" of this kind in fact involves building a "construction hypothesis" on the relationship that the connected forms maintain with one another. In the resemblances method, this amounts to establishing a "conceptual category". However, in the process of development, this is not necessarily the case.

3) The need to study the entire dissemination field of the forms concerned by the relationship chains.

It is the combined analysis of the "construction hypotheses" concerning the relationship chains and of the dissemination fields in reference to the "global hypothesis" being demonstrated which enables one to justify an "attribution hypothesis" regarding the genealogical relationship of the language itself or the organisation of the sub-group.

Concerning Certain Relationships

The references supplied by Ehret regarding his sub-grouping of Nilo-Saharan are for the most part in press, making it difficult to refer to them. Nonetheless, from various recent articles and manuscript documents, one can ascertain that he has established a sufficient number of lexical comparisons to confirm the contested membership of Songhay within Nilo-Saharan and to modify the grouping in that direction. Thanks to Franz Rottland, I have been able to gain access to some of these resemblances with a list of 185 items which, a priori, are crucial to the justification of this hypothesis. It would therefore be useful to undertake an analysis of them. This can be done by comparing these results with other available data. On the one hand, this material would include documentation concerning all the known Songhay dialects and, on the other, some of the most reliable sources concerning all the languages in the Sahelo-Saharan area, namely Berber (Kabyle, Tamazight, and Touareg), Mande (several northern and southern dialects), Hausa, Saharan (Kanuri, Teda, Tubu), Sudanese Arabic and Peul. These compiled items currently represent more than 150,000 cards in SAHELIA. They have all

Robert Nicolai

been interconnected with one another in a "*Corpus des références croisées*"² (Nicolai 1994b) which makes it possible to undertake particularly rapid and complex searches in both the semantic and phonological areas. The data base which I will use is therefore not particularly well-informed in the Nilo-Saharan field, except for Songhay and the Saharan languages, which raises a problem. However, the abundant documentation concerning Berber and Mande enables one to extend the field of comparison and to conduct a more effective appraisal of questions regarding lexical dissemination.

Providing an indication rather than a proof, 70 of these 185 items (Appendix) would also appear to show — and sometimes much more clearly — resemblances with Berber. The size of the store of data that can be called on to justify the place of Songhay within the Nilo-Saharan languages is therefore diminished by a corresponding amount. Correlatively, one can find a very large number of other items attested in all the dialectal varieties (roughly 350) which are also candidates for comparison with Berber and which apparently are not present in Nilo-Saharan (cf. Nicolai 1990, v. also Appendix). To this number, one must also add a large quantity of items which are limited in their extent to eastern and western Songhay and which are as a result more likely to be interpreted as borrowings, as well as the stock of shared borrowings from Arabic, the status of which will be different depending upon the hypothesis adopted.

Having said this, although roughly 70 of Ehret's items can be related to Berber, it is nevertheless a fact that once one has eliminated certain relationships involving Mande, there are about a hundred which appear to be "irreducible" and which could well be of Nilo-Saharan origin (v. Appendix).

The idea of relating Songhay with certain Nilo-Saharan languages situated much further east is also interesting. A project is currently being envisaged to conduct a comparison with certain Nilotic languages³ which, to be meaningful, should take into account all these irreducible, presumed Nilo-Saharan lexical items. We are hoping to obtain some clues, not to define "at all costs" the position of Songhay in the Nilo-Saharan tree model, but to assist in the construction of the framework of preliminary knowledge within which a hypothesis could be established...not necessarily the one I proposed recently, even if it appears obvious that it will be necessary to take into account a complex genetic, but not exclusively genealogical, structure⁴. One can remark here the importance of an extensive lexical comparison, even if a translinguistic study into the spread of the lexical forms has not been totally completed. Even if the results can be

Problems of Grouping and Subgrouping: the Question of Songhay

deceptive, this comparison nevertheless appears to be necessary. However, it should be pointed out that these preliminary counts are not sufficient by themselves to tell us anything about genealogical relationships and their nature.

It will not be possible in this setting to discuss in detail all the relevant items. Therefore I will limit myself to a few illustrations grouped according to whether they concern a narrow semantic field or a broad one, and according to whether they do or do not structure a potential morpholexical field in their languages. These examples will allow us to push our semantic observations a little further.

Items with a Narrow Semantic Field

I will use two items from Ehret which are characterized by a relatively "stable" and "narrow" semantic field and which enable comparisons with Berber: *fofe* (Ehret 661) :

<u>Uduk</u>	aÂphoÂ	back, on the top
<u>Kunama</u>	afuna	bosom; beside
<u>S.</u>	fôÂfé	breast
<u>Daju</u>	pete	nipple, udder
<u>Surma</u>	a:pòti	breast
<u>Nilotic : Dinka</u>	pwon	heart

Ehret remarks: Kunama -n nominal suffix ; Daju, Surma -t ; Songhay : Duplication **afofo* - e > *fofe*).

Ehret relates *fofe* to several Nilo-Saharan forms. I am not in a position to assess the justification for these relationships in this domain, neither from the semantic nor from the phonetic viewpoint. However, *fofe* is attested in the entire Songhay field with the stable meaning of 'sein' and more precisely 'mamelle, pis, trayon' and even possesses a further specific semantic extension, as in Kaado: 'régime de bananes' which is also largely attested in the Berber domain with the same meaning. So, in the Touareg dialects, one finds the form *efaf*, as in Tamazight and Kabyle in which this same item (*iff*) is attested alongside a doublet form *bubbu* which appears to have a less "functional" meaning ('sein, petit sein' as opposed to the form *iff* which designates something closer to 'trayon').

The chain of semantic relationships given by Ehret is as follows: 'back, on the top/ bosom; beside/breast/nipple, udder/heart'. This shows the problems involved in semantically based comparisons.

The involvement of a much more general root than Nilo-Saharan or Afro-Asiatic could be invoked. It does not solve the problem of this type of root, which, on account of

Robert Nicolai

its general nature and extension, does not allow one to draw inferences acceptable from the "resemblance" point-of-view, but which can nevertheless be used with no difficulty in a detailed treatment and in precise comparisons.

tasa (Ehret 735):

<u>Kunama</u>	desa	liver
S.	<i>taĀsā</i>	<i>liver</i>
<u>Nilotic</u> : <u>Teso</u>	a-tid	spleen

Like the previous form, this one is widely attested in Songhay and Berber. It refers to an internal part of the body and also to certain feelings ('courage, fear'...).

So, in Songhay, one finds *tasa* in all the dialects to designate 'liver'. One also finds this form in Berber where it designates both the liver and also feelings involving oppositions ('peur, courage'), and again the notion of maternal love ('entrailles maternelles'). Cf. Kabyle: 'foie, courage au travail, audace, attachement pour ses enfants' ; Touareg (Tahaggart et Timbuktu) : 'ventre ; côté maternel' ; Tawwellemmet : 'foie'; Tayrt 'foie, peur incontrôlée'⁵.

The possibility of other relationships with the Berber forms connected with the Western Songhay root and meaning 'foie, bile, aigreur d'estomac, vésicule, rate' can also be envisaged. It is even this particular form which could be related to the Saharan items (Tubu, Teda, Daza) which mean 'peur, crainte': *aus*, *awozer*, etc.

In Berber, the three forms (*tasa*, *adis/tadist*, *awza*) coexist in several dialects (cf. Tahaggart). Finally, in Bozo (Daget) one finds *tassa* which could be a borrowing, and *disi* 'poitrine' in Maninka (Cl. Grégoire), without being able to check whether it is a borrowing or a form emanating from the Mande lexical stock.

These forms are known throughout the Berber-speaking area with more general and more diversified meanings than those found in Songhay. One is led to conclude that, on this point, the relationship between the two languages is very strong but that its definition as borrowing, cognate or other cannot be made.

The chain of semantic comparisons made by Ehret is as follows : 'foie/mauvaise humeur', probably for reasons which are 'bilieuses'. There seems to exist no particular semantic relationship in either Songhay or Berber with *tasa* whose extension we have already studied. However, perhaps this needs to be verified.

Items with Broad Semantic Fields

These involve above all items which are widespread outside Nilo-Saharan and whose form is possibly stratified. Several examples can be provided⁶: (815) *tanka/tuku* ;

Problems of Grouping and Subgrouping: the Question of Songhay

(801) *tuusu* ; (440) *gong* ; (630) *ferek* ; (540) *bita*; (69) *buka* ; (583) *bur* ; (589) *fari* ; (1038) *kosu*, etc. Stratification, in this study, refers to the fact that it is possible, with no justification internal (structural) or external (dialectological) to the language, to relate several items which are distinguished by the same type of regular variation or addition of the formative.

I will present the example of the item *bita* (540). According to Ehret, this item involves the semantic unit 'to rise' ('lever, hausse, ascension, augmentation: houle, crescendo; se gonfler: épaissir, grossir'), attested in particular by the following examples:

<u>Centr. Sud.</u>	*pi	to increase
<u>Kunama</u>	bido-	to rear upon (of animal)
<u>S.</u>	biÁtâ	to be thick, pasty
<u>Dongolawi</u>	bitaan	child, offspring, son, fruit, seed
<u>Proto-Nilotic</u>	*pít̃	to grow (based on Shilluk)
<u>Kuliak</u>	*ibit	to grow
<u>Berta</u>	biÁd	to fly

The relationship attempted by Ehret is risky, for it is not obvious that the general semantic unit he identifies is the one which characterizes the Songhay item he studies. Otherwise one needs to keep to the notion of 'épaississement' which would result from an "innovation"⁷.

In Songhay, this lexeme seems to be organized in a stratified series of interesting relationships which I will attempt to demonstrate. For instance, in Songhay, there exists a lexematic form *bita* which is found in all the dialects. It has a general meaning of 'bouillie de farine de mil ou de blé', generally a very liquid one. One can also observe that there are different phonetic variations of the form which, in some dialects, seem to be semantically attributed. Cf. Kaado: *bita* 'être pâteux, en bouillie' ; *bibita* 'préparation de mil proche de la bouillie' ; *betbeta* 'bouillie peu épaisse (sp)'

Alongside this form with its voiced initials, one finds another form with an unvoiced fricative initial: *foto*. It is attested in western Songhay with the meaning 'too liquid'. One also finds *foto* in eastern Songhay meaning 'bouillie de farine et brisure, plus grossière que *bita*'. A reduplicated form is attested in western Songhay: *fotofoto* 'mets préparé avec du riz écrasé en farine'. These forms all relate to a semantic unit designating 'un liquide trouble, plus ou moins épais, plus ou moins grumeleux'. They do not relate specially to the notion of 'pâte qui lève', the only notion which could connect 'croissance' with 'liquidité'.

Robert Nicolai

At the same time, one finds voiced forms with final vocalisation designating another type of 'liquide trouble', and suggesting 'boue'. Thus one finds *botonte* 'vaseux, boueux', *botoboto* 'marécage, vaseux' in Western Songhay and in Eastern Songhay: *botobotu* 'former *potopoto*, où on enfonce'; also *buranta* 'préparé pour être filtré'. Once again there is the same reference to the 'liquide et trouble' quality....

Outside the Songhay (S) domain, in Mande (M), there are fewer attestations. The following have been found:

Bozo: *bòtòbòtò* 'boue', *bèrè* 'bouillie',

Soninke: *fonde* 'bouillie de farine de mil', *buda* 'troubler l'eau',

Maninka: *bòrò* 'boue'; *budun* 'se troubler (liquides)',

Soso: *booraa* 'boue'

Bozo: *buruburu* 'dépôt solide dans un liquide'; *bore* 'cuisson, lorsqu'il s'agit d'un met ramolli / cuire à point'.

Lastly, to be exhaustive as far as food is concerned, it is interesting to note that in Dendi one finds the attestation *bodobodo* to designate European bread. The same is attested in Kaado. These are the only two cases in which the notion of 'gonflement' could appear, but the indigenous quality of these lexemes is not at all obvious, given that the generic noun for bread is *buura*.

In view of this 'liquidité' semantic unit, one also finds phonetically similar forms in Songhay which relate to the notion of 'brumeux, laiteux, nébuleux', then possibly 'gris, couvert, obscur'. So in E. S. *bitébité* 'être brumeux, laiteux'; *bitibata* 'gris, couvert'; W. S. *biti* 'être nébuleux'. Finally, there are also variants with a final M: *bitim bitima* 'être brumeux, ciel couvert', a meaning which one also finds in Soninke with *bita* 'être obscur'.

The Saharan field also offers interesting forms for comparison. For example, Teda and Daza possess, respectively, the forms *boda*, *bede* with the meaning 'nourriture' whereas Kanuri, with the form *bèri* (Bilma: *huri*; Manga: *bri*) attests 'bouillie, nourriture, 'Aish; couscous; pâte de mil' and, in the Kanuri dialects, alongside this form, one finds a semantically and phonetically related form characterized by the addition of the consonant m: *balam* 'bouillie, mélange généralement de riz et de lait'; *balambo* 'plat préparé à base d'un mélange de farine et du jus de l'arbre marga', etc.

The Berber domain, for its part, attests two roots: *BR* with the Kabyle 'mouiller' and 'marécage'; *BZ* with the Kabyle 'être trempé, être mouillé, enfonce, couler'. But also the Tamazight 'plonger, immerger (dans un liquide), enfonce, s'enfoncer dans un liquide'.

Problems of Grouping and Subgrouping: the Question of Songhay

This presentation, organized around *bita/foto/boto* could usefully be compared with another similar set structured around the *botogo/fotogo* reference units and with a similar semantic unit which is characterized more particularly by the presence of a third formative (a velar occlusive).

botogo/fotogo/ferege : this lexical structure is similar to the previous one both by virtue of its meanings and of its phonetic structure. The two first syllables are identical. To them are added a formant with an initial velar. The relationship between the two structures - and consequently the processing of this relationship - will necessarily be different depending on whether one adopts the hypothesis involving a potential genealogical relationship or borrowing, or whether one assumes the existence of a particular morphism or any other hypothesis.

For myself, and for the moment, I believe it is too early to draw up hypotheses on this level. But this does not preclude groupings on the basis of relationships organized around transformation criteria the credibility of which, as a classification operation exclusively, is formally acceptable, given our experience of the attested language shifts.

botogo: this form is largely attested in Zarma, Kaado and Dendi (with a metathesis: *bokoto*) to signify 'boue, argile, marne'. It is duplicated in Kaado and Zarma in the form *fotogo/fotogu* which means 'être liquide, être surabondant'. One also finds the form *butugu* 'trouble, troubler, en parlant d'un liquide' in Zarma as well as in Eastern Songhay in which the duplicated form *butugubutugu* signifies 'barboter dans l'eau'. Throughout the field, one finds the designation of a brick with the form *fareeje, ferey/feraji/feraw*. Having said this, western and eastern Songhay attest the form *batakara* to mean 'la boue, la terre mouillée ou l'argile des bas-fonds'.

Extensive Research

All of these semantically related forms seem to demonstrate a certain parallel with bi-syllabic forms presented above. However, the tri-syllabic form extends beyond the Songhay field. It is attested in Berber as well, though to a lesser extent, in Mande, and in Saharan. Consequently, in Kabyle and Tamazight one finds the root *bl*⊙, which corresponds exactly with the semantic unit 'mortier, imbiber', then 'souiller; dépôt, fond de liquide, barbouiller'. It is also found in Touareg (Timbukto, Tahaggart, Tayrt, Tawallemmet) *ebdeg*, etc. 'mouiller, gargouiller', etc.

An extension of these forms is found in Northern Songhay (forms belonging to the Berber lexical stock) with *abla*⊙ meaning 'brique'; also *birgi* which is attested in

Robert Nicolai

Tasawaq by Lacroix 'briques moulées rectangulaires', although the question of confusion with the French word 'brique' must necessarily be raised. Nonetheless the bricks in question are quite obviously traditional sun-dried clay bricks.

The Saharan field also has the same forms. In Tubu (Lukas), for instance, one finds: *berège* 'argile, glaise', *barègè* 'marécage' ; *firki* 'marécage, mare, flaque' and in Daza (Le Coeur) : *barege* 'mare'. There is a probable semantic and phonetic relationship here. As for the Mande field, the following forms have been found: *bòrògò* / *bòròkò* 'boue', *borogoborogo* 'boue', *bagakuru* 'boulette de farine pour faire la bouillie' ; in Soninke : *baxare* 'bouillie de mil et de lait caillé', *baga* 'bouillie de farine de mil'.

I haven't yet exhausted the potentially interesting tri-syllabic items. In Berber, there is still another useful root, which is *BZG*. It is found in both Tamazight and Kabyle with the following meanings: 'être mouillé//enfler, être enflé, faire gonfler à l'eau//mollir, devenir mou'. Finally, to round off the consulted items, one should note that in Sudanese Arabic the form *baagiir* means 'marais, marécage'; *buriye* 'marécage, marais, sol à coton lourd et bourbeux pendant les pluies, sec durant la saison sèche'.

This presentation still needs to be expanded by comparison to a related formal structure designated here as *BUR* and which refers, apparently, to the general meaning 'cendre chaude'. Ehret (583) attests the following forms:

<u>Centr.Sud.</u>	*pu	'ashes'
<u>S.</u>	buÁroÁw	'ashes'
<u>Berta</u>	buÁbuÁdâ (bbb)	'ashes'

and refers to the form *burow* to designate ashes in Songhay. This is not the only one. There is also the form *boosu*. It would appear, if one limits oneself to lexical analysis, that the notion 'ashes' is not a simple one. Semantically and lexically, one needs to distinguish between 'cendre chaude' and 'cendre froide'.

I am going to organize the corpus of items likely to be of use in relationship with the preceding work and with a postulated semantic and - initially a priori - division. One draws up a first list of the useful forms relating to the notion 'nuage, tornade, brume'. These items are situated at a point of contact both with the notion 'cendre'(made of fine dust) and with the notion of 'espace troublé par des particules' which is also found in 'eau boueuse' as well as in 'air poussiéreux' or 'brumeux'.

Problems of Grouping and Subgrouping: the Question of Songhay

In Songhay, the form *buru/bur* meaning 'nuage' is attested throughout the Songhay field except for Northern and Eastern Songhay. In Zarma, one also finds the form *buuda* (with a long vowel and occlusive apical) 'brume'.

As far as the notion 'cendre' is concerned, the form *buraw/burow* considered by Ehret is attested in Zarma and in Kaado with the meaning 'tas de braises qui commence à se transformer en cendres'. The *w* which closes the final syllable deserves attention since it could well point to a former tri-syllabic structure (cf. the underlying hypotheses in the analysis of *feraw/ferey* 'brique' from **ferege*. One should also bear in mind the Eastern Songhay form : *buberey* 'noircir, carboniser en surface, griller à feu nu, flamber'. However, we have scarcely no clues and few parallel forms (as was the case with *botogo*) to justify this hypothesis and the corresponding Zarma, Kaado and W. S. forms (*bolbol*: 'cendre chaude'; *bubur*: 'rôtir qqch facilement et vite') do not possess the final *y*. Only one attestation in Maninka by Cl. Grégoire suggests *buruku* 'cuire à la cendre', which could constitute the starting point for an attempted justification or further exploration. Parallel to this set and with similar semantic designation, one finds forms like *berbere* designating 'flamber, passer à la flamme, griller à feu nu, etc.'

Analysing

A lexical unit can potentially be broken down into morphemic units within its language. But alongside this linguistic analytical breakdown, one can imagine other structural levels relating to a formal or semantic type of breakdown. One can therefore find potential organizations of the data which can be defined within the lexicon and which do not appear to have a synchronic backing within the scope of linguistic relevance. The determinants of these structures can be partially quantitative or can be based upon a form of "constructed correlationism". Perhaps they are the same determinants which underpin the construction of structures such as analogy or popular etymology...since at some stage or another one must analyse, on the one hand, the results of genuine popular etymology and, on the other, the parallel results of a "scholarly" etymology drawn up without the backing of philology and of texts, either from a "quasi-popular" etymology justified by "likely" inferences based upon empirical, cognitive and/or theoretical considerations on the basis of the store of knowledge that the researchers build, or which they already have available!

As a result one observes a kind of laminated, multilevel structure among the lexical units which has no established diachronic justification but which nonetheless exists and which results perhaps from a diachronic process which remains "non-demonstrable in the

Robert Nicolai

current state of our knowledge". If this laminated structure is based upon recurrent relationships, it then perhaps becomes possible to identify it as a characteristic phenomenon of the language in question...and therefore use it, for what it can provide, in one's comparisons.

Example: In Songhay, one can relate a series of lexemes which can be connected by a single phonetic differentiation and which also stand in potential semantic relationship. This is the case with certain categories of lexemes with initial labials (*b, f*) similar to the one we have just dealt with.

For example, around the notion of 'bouillie, boue, préparation liquide'

bita 'être pâteux, en bouillie' ; *bibita* 'préparation de mil proche de la bouillie' ; *betbeta* 'bouillie peu épaisse (sp)',
botogo, bokoto 'boue, argile, marne',
batakara 'la boue, la terre mouillée ou l'argile des bas-fonds'.
foto 'bouillie de farine et brisure, plus grossière que *bita*',
fotofoto 'mets préparé avec du riz écrasé en farine',
fotogo/fotogu 'être liquide, être surabondant',
fareeje/feray/feraji/feraw 'brique'.

However, other lexical units without semantic links with the preceding examples are: -

around the notion of 'filage, enroulement'

bibiri 'corder, rouler entre ses doigts'.
fifiri 'tourner la sauce'.

- around the notion of 'perçement, ouverture'

biri 'tirer une flèche, os, arête',
beeri 'piocher, abattre, couper',
bidi 'variole'.
feri 'ouvrir, délier, décortiquer',
fiti 'ouvrir',
foto 'variole',
fiti 'sauter, se débattre',
fifiti 'couvert de petits boutons'.

- around the notion of 'éclater, gratter'

birgi 'bale de mil, faire voler la terre avec le pied',
birji 'agiter, troubler, remuer'.
firkiti 'frétiller, éclat de bois, bouger bruyamment', etc.

These blurred correlations are, in a way, characteristic of language, in this case of Songhay. Knowing of their existence enables one to better justify these comparisons and it is essential, at the outset, to study to what degree they can be used to elaborate correspondances between dialects or languages which are similar to one another. The results one obtains on this issue, whether positive or negative, will necessarily provide useful information⁸.

Problems of Grouping and Subgrouping: the Question of Songhay

The same comments can be made regarding the potential identification of "formative" structures. In Songhay, for instance, one finds phonetically and semantically similar units which are distinguished by the presence or absence of particular consonantal formatives. Obviously, it is not possible to give a meaning to these formatives - otherwise they would be morphemes. However, they can occasionally appear recurrently in numerous lexical fields that they help to structure. Nothing allows us to assert that they are defunctionalized or random morphemes, although they also form part of these blurred correlations.

For instance, around $BT(G)/BT(K)$, there appears to be a formal structure built as follows:

$$\begin{array}{ll} b/f-d/t & b/f-d/t-g/k \\ b/f-r/z & b/f-r/z-g/k \end{array}$$

In addition, this type of structure with formative extensions also exists in other series of lexemes. The important point is not to launch out into hypotheses concerning the origin of these laminations but to observe:

- 1) that they can be identified, or "shaped" in the manner of a gestalt,
- 2) that it is impossible to construct a lamination with just any lexical grouping.

As a result, the demonstration of the phenomenon can be employed with a useful "indicative" value in the context of inter-language comparisons. This means that a lexical study deserves to be performed along these lines. In the context of such a comparison, one might ask the question: When two languages are compared and when one of them contains laminated lexical sets, can one find a corresponding lamination in the compared language or not?

The answers would certainly contribute information to historical research into the genealogy or the dissemination of languages.

Conclusion

It would seem that it is necessary to undertake a preliminary analysis of the data and that the interest of this analysis goes well beyond the question of the quest for genealogical relationships. It could be argued that it will help to revitalize research into (modes of) lexical structuring and into the dynamics of the structures.

Several forms of relevance intervene here:

- linguistic structural relevance in its diverse relational dimensions (phonetico/phonological, semantico/semiotic and morphologico/formal).

Robert Nicolai

- cognitive and "rationalizing" relevance of the demonstration of potentially significant "logical" and "conceptual" structures.
- the relevance contingent upon demonstration of relationships from encyclopaedic knowledge resources.
- synesthetic relevance leading to partially motivated structures.
- interactional relevance in the framework of which the references justifying the construction of what is happening are effectively constructed.

The dynamics of language change result from the interaction of these relevances. Language development is an expression of these dynamics; realized languages are an expression of the outcome. Generally speaking, each of the dimensions possesses an explanatory potential. In fact, none explains anything on its own since none produces the changes independently of the others even if, in limited circumstances, at a given moment, one or other of these relevances appears to prevail. In these cases, the short-term impression is that the analytical model that it draws upon most heavily is capable of explaining everything... but, obviously, this is not a general phenomenon. No model relating to a single relevance is likely to account for the development of languages since languages, by virtue of their function as well as of their organization, are dependent upon several relevances.

Seen from this angle, the practical study of language development has a historical dimension. This does not mean that this form of study should adopt analysis methods used to account for historical facts or for archeological investigation.

It is necessary to distinguish between:

- a general approach to language development which is likely to develop into a broad, non-specialized approach based upon studies dealing with language change, historical linguistics and research into the dynamics of representations "in general". This approach aims at understanding and analyzing the dynamics of language change in a global manner.
- an empirical approach which involves imagining and reconstructing in detail the specific diachronic development of a given group of languages of which the initial unit is projected but without being divorced from a general stream of reflexion which it helps to nourish by feeding on its findings. Here, however, unlike "surface" analogies, this approach does not resemble archeological exploration since the latter is based upon objectifiable clues concerning what it is being reconstructed. A broken pot or an archeological site is not a hypothesis, whereas a resemblance is a hypothesis. A hypothesis about a broken pot is not a hypothesis on a hypothesis⁹.

Problems of Grouping and Subgrouping: the Question of Songhay

Finally, to return to the sub-grouping of Songhay. The immense amount of work undertaken and the diversity of the results lead me to believe that the problem is far from simple and is very unlikely to be resolved today.

It is always possible to draw up a list of "proofs" for or "proofs" against. One can reject the hypotheses by arguing point by point over each item used or one can discard the lot *en bloc*. This is merely a matter of individual "temperament". What is manifestly clear is that our assessment should bear upon the limitations of the method before we begin to criticize the results the method provides.

It is possible, even when starting from opposing critical standpoints, to acknowledge the obvious. As Bender wrote (1992:17) : "*les questions de la classification génétique et de la diffusion lexicale sont inextricablement entrelacées*" and that "*dans l'espoir de pouvoir faire ressortir des paradigmes et résoudre certains problèmes, on devrait travailler en passant constamment de l'un à l'autre*". All the evidence seems to indicate that the time is ripe for us not only to work in collaboration with one another regarding detailed items but also - and perhaps even more so - to construct new analytical paradigms which take the facts more fully into account and which, perhaps, might modify the nature of our questions.

The above remarks are intended as a step in this direction. The establishment of the position of Songhay lies ahead of us.

REFERENCES

- Bender, M. Lionel. 1989. Nilo-Saharan Pronouns/Demonstratives. In: M. L. Bender, ed.: *Topics in Nilo-Saharan Linguistics (Nilo-Saharan Linguistic Analyses and Documentation 3)*: 1-34. Hamburg: Buske.
- 1991a. Sub-classification of Nilo-Saharan. In: M.L. Bender, ed.: *Proceedings of the Fourth Nilo-Saharan Linguistics Colloquium (Nilo-Saharan Linguistic Analyses and Documentation 7)* (Bayreuth 1989): 1-35. Hamburg: Buske.
- 1991b. Compte rendu de R. Nicolai, Parentés linguistiques. In: *Journal of Pidgin and creole languages*: 6.2: 349-50.
1992. Classification génétique des langues nilo-sahariennes. In: *Linguistique Africaine*, N°9. 15-39. Paris
1996. (Forthcoming). *The Nilo-Saharan Languages: A Comparative Essay*. Munich: Lincom Europa.
- Dalby, David. 1966. Levels of relationship in the comparative study of African Languages. In : *African Language Studies*, vol. 7: 170-179.

Robert Nicolai

Erhet, Christopher. 1989. Subclassification of Nilo-saharan: A proposal. In: M.L. Bender, ed.: *Topics in Nilo-Saharan Linguistics: (Nilo-Saharan Linguistic Analyses and Documentation 3)*: 35-50. Buske, Hamburg.

1989 ms. The subclassification of Nilo-Saharan.

Manessy, G. 1990. Du bon usage de la méthode comparative historique dans les langues africaines et ailleurs. In: *Travaux 8, Cercle linguistique d'Aix-en-Provence*: 89-107.

Meillet, Antoine. 1958. *Linguistique historique et linguistique générale*. Paris: Champion. 335 p.

Nicolai, Robert. 1990. *Parentés linguistiques (à propos du songhay)*: Paris: Editions du CNRS. 209 p.

1993. Apparentements linguistiques: problèmes théoriques et méthodologiques. *Acta Universitatis Carolinae*: 57-74. Prague.

1994. Remarques concernant l'utilisation du lexique pour l'étude des relations entre les langues. In: *Deuxième Table ronde du Réseau Diffusion lexicale en zone sahel-saharienne*, fasc. 1: 24-42. Prague: Center for Theoretical Study.

1994. *Corpus des références croisées de SAHELIA* (document informatique). Université de Nice.

Rottland, Franz. 1994. On Lexical Relations: Nilotic-Songhay-Nilosaharan: In: *Deuxième Table ronde du Réseau Diffusion lexicale en zone sahel-saharienne*, fasc. 2: 43-50. Prague: Center for Theoretical Study.

SAHELIA. *Base de données lexicales et dialectologiques sur la zone sahel-saharienne* (document informatique). Nice.

Problems of Grouping and Subgrouping: The Question of Songhay

Appendix

Entrées songhay dans Ehret (avec développement possibles vers le berbère)

(Les entrées sans référence dans la colonne "Parentés..." renvoient à de rapprochements "possibles" établis depuis la parution de Nicolai "Parentés...", 1990).

ref. Ehret			Parenté
23	<i>boori</i>	to be good	46
27	<i>baka</i>	to put sth. to soak	45/52
42	<i>bey</i>	to know	36
43	<i>betu</i>	to wait for	26
45	<i>biri</i>	bone	39
51	<i>biiri</i>	to bring up	
59	<i>boro</i>	person	51
69	<i>buka</i>	to grind...	20
72	<i>buti</i>	to lift a garment (to urinate)	
92	<i>bute</i>	vulva	
97	<i>margan</i>	to join	267
105	<i>koms</i>	foot of cattle	
	<i>moosi</i>	nail, claw	
107	<i>me</i>	mouth	271
114	<i>meli</i>	lightning	272
138	<i>moy</i>	eye	277
152	<i>musey</i>	to massage	281
218	<i>deede</i>	to announce	
229	<i>din</i>	that	85
279	<i>zaanNo</i>	noon	388
289	<i>zarga</i>	bouillir de colère	393
322	<i>zefe</i>	to slash, chop	396
362	<i>zuru</i>	to flee	412
378	<i>nyukunyu</i>	to be compact	
394	<i>garu</i>	to intersect	144
440	<i>goNgon</i>	to be bent	
451	<i>gaasi</i>	to scratch	
469	<i>guusu</i>	be deep	161
471	<i>gaasu</i>	scratch ground	
507	<i>noNgon</i>	place	
540	<i>bita</i>	to be thick	
583	<i>burow</i>	ashes	
589	<i>fari</i>	field	109
	<i>faru</i>	cultivate	109
601	<i>feeme</i>	hearth	
603	<i>fesfese</i>	to gin cotton	114
	<i>fese</i>	to pick out, shift	119
630	<i>fifiris</i>		
	<i>firka</i>	to swerve	121
637	<i>futu</i>	to be angry	129

645	<i>kufu</i>	lungs	237
661	<i>fofe</i>	breast	126
693	<i>tamtam</i>	to grope	335
696	<i>te</i>	to come	
	<i>to</i>	to reach	
699	<i>taba</i>	to taste	337
702	<i>tibi</i>	to take a handful of	87
706	<i>toora</i>	to slaughter	(pl?)
716	<i>taru</i>	to hurry	
720	<i>tonton</i>	to add	365
735	<i>tasa</i>	liver	347
752	<i>tu</i>	wooden dish	
753	<i>tuti</i>	to push	
776	<i>duru</i>	piler	102
781	<i>taama</i>	to stamp with feet	335
787	<i>tetengi</i>	to stagger	
801	<i>tuusu</i>	to anoint	
815	<i>tanka</i>	shallow hole	368
817	<i>tu</i>	placenta	
856	<i>dusu</i>	être ankylosé	103
	<i>dusuNu</i>	to fall asleep	103
914	<i>tongo</i>	quiver	364
925	<i>telensi</i>	to slip, slide	353
954	<i>gebu</i>	to belch	
997	<i>kambu</i>	pincers	
1026	<i>kofi</i>	to hit on head	
1038	<i>koosu</i>	to scratch	
1042	<i>kuttun</i>	to bother s.o.	
1044	<i>kula</i>	gourd sp.	
1048	<i>kusow</i>	dust	
1064	<i>hay</i>	to strike with a spear	175
1073	<i>herow</i>	fem. kid	
1142	<i>sufu</i>	to wet slightly	
1170	<i>seele</i>	cut meat in strips	
1247	<i>didigi</i>	to roll up	
1402	<i>kooro</i>	hyena	226
1490	<i>hagey</i>	to winnow	
1562	<i>hey</i>	thing	
	<i>a</i>	she/he/it	

Robert Nicolai

Notes:

¹In order to avoid the charge of 'partiality' or of using truncated quotations, I should point out that Meillet opens the following paragraph by writing: "La difficulté qu'on éprouve à démontrer une parenté des langues en bien des cas ne suppose donc qu'il s'agisse de langues mixtes". Point taken. However, for an analysis of this statement, consult Nicolai (1992).

²The 'Corpus des références croisées de SAHELIA includes roughly 20 000 references and connects more than 150000 cards within the data base. It constitutes a useful tool for research and for the analogical comparison of items.

³ In collaboration with Fr. Rottland in the framework of GDRE 1172 "*Diffusion lexicale en zone sahelosaharienne*".

⁴ For the distinction between genetic and genealogical, which has not been 'theorised' in the literature in English, cf. Manessy (1990) or Nicolai (1993).

⁵In Raffanel, one finds (*dissi* 'estomac') which is an unreliable attestation. Very often, the (very old) items in Raffanel contain Touareg attestations. In this case, the relationship that needs to be made is with *tadist*, *adis* and *addas* 'ventre', but also 'grossesse, foetus' which is realized in Touareg, Tamazight or Kabyle.

⁶ The figures relate to Ehret's references. The forms suggested in the opening remarks correspond to R.N.'s references in SAHELIA.

⁷ One can see here the impact of subjectivity in the assessment. In the final analysis, what authorizes me to make this judgment?

⁸ Cf. Fr. Rottland (1994): "I base my own work on the assumption that the tree model and the concept of linguistic continuity that it entails are valid but not universally applicable. It follows from this assumption that the possibilities which the traditional approach offers should be exhausted before we look for alternative interpretations."

⁹ It should be noted however that 'a hypothesis about a broken pot' is only meaningful in close connection with the site on which it was found! Here again, the 'context' is fundamental.