

Les Parlers Bozo: Révision d'une Enquête Dialectologique

par Thomas & Bärbel Blecke

1. Introduction

Dans le Delta Intérieur du Niger au Mali, de Ségou au Lac Débo, sont situés les parlers bozo. Linguistiquement classifié comme Mandé du Nord-Ouest (< Mandé de l'Ouest < Mandé < Niger-Congo)¹ avec le soninké comme parent le plus proche, cet ensemble de parlers est géographiquement situé en marge nord-est du phylum Mandé. Etant restés très peu documentés jusqu'aux années 90, c'est une enquête dialectologique menée par un groupe de chercheurs de la Société Internationale de Linguistique sous la direction de Brad et Sue Smeltzer qui a le grand mérite d'avoir fourni la première vue d'ensemble des parlers bozo basée sur des données systématiques recueillies sur le terrain. Au sein de cette enquête une étude lexicostatistique constitue la base du calcul de proximité des différents parlers. Son résultat se résume brièvement en la reconnaissance de quatre langues bozo sans intelligibilité interne², dont une se subdivise en six dialectes nettement distincts.

Nous sommes très reconnaissants à Brad et Sue Smeltzer, qui, dans un esprit de libre échange scientifique, nous ont laissé le dossier entier de leur enquête des années avant sa publication. Si, dans ce qui suit, sa partie lexicostatistique sera soumise à une révision critique, nous espérons l'avoir menée dans ce même esprit. Tout en respectant leur travail systématique, notre but sera d'en augmenter la précision pour en optimiser l'approximation à la réalité linguistique. Sur la base de connaissances phonologiques et morphologiques approfondies d'un parler bozo précis (cf. Blecke 1996) et en mettant en valeur des données supplémentaires (DKS 1953), une correction importante des résultats sera proposée, qui mettra en question la répartition en quatre langues distinctes de l'ensemble bozo. En conséquence, la non-intelligibilité entre les parlers restera également à reconsidérer.

Un intérêt plus général de cette révision est qu'elle exemplifie des types de pièges qu'on rencontre d'une façon presque systématique lors d'un travail basé sur des listes de mots. Ainsi l'enquêteur pourra profiter des exemples donnés pour éviter, dans sa propre recherche, bon nombre de défauts potentiels.

2. L'Enquête de Smeltzer & Smeltzer 1995³

Le sujet de l'enquête de Smeltzer & Smeltzer 1995 (ci-après S&S'95) est la population bozophone, c.à.d. tous ceux dont la première langue est un parler bozo quelconque, des considérations d'ethnicité étant mises à part (S&S'95:58). Les objectifs de l'étude se répartissent en trois domaines, comme l'indique le titre de son compte-rendu:

¹ Selon Kastenholz (1996) et Williamson (1989);

² Selon Blair (1990:2) ce terme technique désigne "...the degree of understanding a speaker of one speech variety has of a related speech variety because of the degree of the genetic relationship between the two varieties."

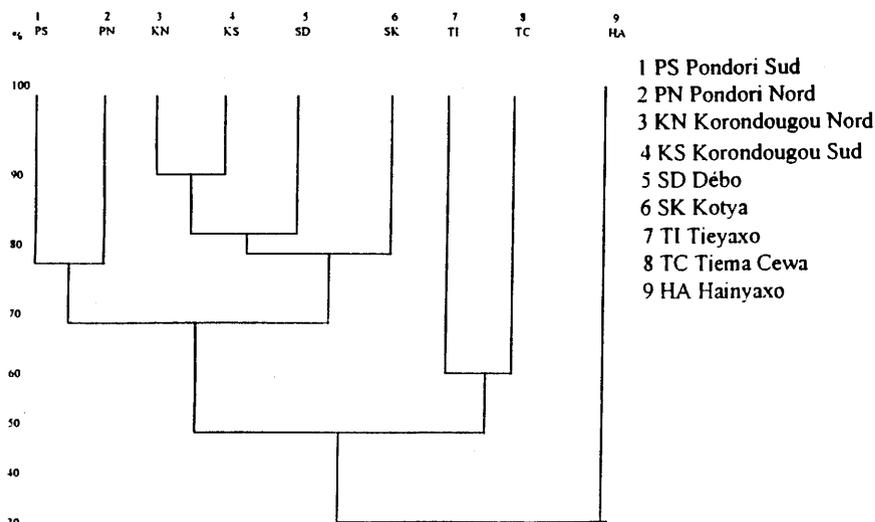
³ La recherche sur le terrain a été menée en deux étapes en 1990/1991 déjà

- délimiter nombre et territoire des parlers (cartographie)
- mesurer les relations linguistiques entre les parlers respectifs: taux de parenté et intelligibilité interne (lexicostatistique)
- explorer la vitalité des parlers, les attitudes des locuteurs par rapport à leur langue, le niveau de bilinguisme (sociolinguistique)

Comme notre révision ne portera que sur la partie lexicostatistique de l'enquête, nous limiterons par la suite à ce qui y est directement liée. A part la familiarité avec l'état actuel des connaissances du sujet de recherche et des préparations techniques, la démarche poursuivie dans l'enquête bozo comprend deux étapes principales. En premier lieu, une recherche sur le terrain a été menée pour recueillir des listes de mots à chaque endroit jugé représentatif d'un parler. Ceci a été réalisé par deux voyages de recherche. D'abord une liste de 30 mots fournissait des critères supplémentaires⁴ pour déterminer des villages échantillons. Ceux-ci établis, une liste de 300 mots, relevée pendant une deuxième visite à chacun de ces endroits, constituait la base de données principale pour calculer le taux de parenté entre les parlers.

Après cette recherche sur le terrain, la deuxième étape de la démarche générale comprend l'évaluation des données obtenues. Celle-ci a été effectuée par un calcul schématique de similarité linguistique (S&S'95:63-68). Selon son résultat, l'ensemble bozo consiste en quatre langues distinctes, dont une se subdivise en six dialectes. En voici un diagramme⁵ représentant les taux de similarité lexicale et les relations linguistiques qui en sont déduites:

tableau 1: Diagramme de l'Ensemble Bozo selon Smeltzer & Smeltzer 1995



⁴ Le choix des villages représentatifs a été basé sur des enquêtes déjà disponibles, ainsi que sur un questionnaire et la liste de 30 mots (cf. S&S'95:59-61).

⁵ Le diagramme est produit à l'aide de Lexistat de Schadeberg (1990).

3. La Phase d'Evaluation: Pièges et Erreurs

3.1. Considérations Préliminaires

En exploitant des listes de mots recueillies, on a intérêt à distinguer deux choses, bien qu'elles ne soient pas opposées l'une à l'autre. On obtiendra d'abord les items tels qu'ils sont transcrits, c.à.d. au niveau phonétique brut. On pourra s'en arrêter là pour en tirer les similarités entre lexèmes et estimer l'intelligibilité interne entre les parlers. Ceci implique pourtant le risque presque inévitable de sauter les liens entre lexèmes comparés, qui sont souvent sous-jacents de par leur caractère phonologique ou morphologique. Si l'on se base essentiellement sur ce niveau et qu'on fait un calcul schématique des données, on surestimera la distance linguistique entre les variétés comparées.

Alternativement, on peut soumettre les données recueillies à une étude linguistique et établir d'abord une esquisse phonologique et morphophonologique avant de faire le calcul des similarités lexicales. On éliminera ainsi des facteurs fortuites de la surface phonétique et obtiendra des relations génétiques relativement fiables. Le risque qu'on court en se plongeant dans l'étymologie, c'est de perdre de vue l'intelligibilité interne. Car le niveau diachronique néglige la réalité linguistique telle qu'elle est perçue par les locuteurs, puisque les liens de parenté entre lexèmes peuvent être si indirects, qu'ils ne sont à reconnaître que par un spécialiste.

Ceci étant dit, on rappellera que dans l'enquête bozo, le but de l'approche lexicostatistique était „d'établir les relations génétiques entre les parlers bozo" (S&S'95:63), c.à.d. d'y inclure explicitement des considérations diachroniques. En plus de cela, les résultats lexicostatistiques allaient servir à estimer l'intelligibilité interne entre les parlers (cf. S&S'95:56), ce dernier but partant d'une covariation entre relations génétiques et intelligibilité interne⁶. On notera maintenant que „le degré de similarité phonétique entre deux mots a été déterminé selon la méthode de Blair" (S&S'95: 64, cf. appendice D pour la méthode), donc à un niveau essentiellement phonétique, augmentée par certains traits phonologiques bozo (cf. op.cit.:appendice E). En procédant ainsi au niveau phonétique, on arrivera difficilement à éviter bon nombre de faux jugements de parenté. Car le caractère d'une transcription phonétique brute dans un continuum dialectal ne manque pas de traits plutôt futiles que pertinents et est par conséquent extrêmement difficile à évaluer.

Si l'on se base principalement sur des listes de mots pour établir à la fois le taux de parenté et l'intelligibilité interne, une solution qui nous semble être viable, c'est de prendre comme point de repère la réalité linguistique telle qu'elle est perçue par les locuteurs. Ceci signifie qu'on prendra comme base du calcul de parenté lexical ni le niveau phonétique ni le niveau étymologique. On a plutôt intérêt à étudier d'abord les données obtenues pour en tirer toute indice phonologique et morphologique. On mettra également en valeur les traits caractéristiques du (sous-)groupe linguistique, surtout en phonologie. L'ensemble de ces phénomènes permet d'estimer, si imprécis que ce soit, la perception du locuteur: on peut supposer qu'il reconnaîtra un item comme correspondant ou non à un autre item moins courant selon les principes phonologiques qui lui sont familiers. C.à.d., il associera à deux lexèmes une même signification lorsqu'il est habitué au type de correspondances qu'ils exposent. Ainsi, des

⁶ Cf. Blair (1990:2) et Simons (1983 chap. 5.2) à ce propos.

items parfois très divers à la surface peuvent être intelligible sans problèmes pour des individus d'un même ensemble de parlers.

3.2. Sources d'Erreurs

Dans cette section, toute la partie lexicostatistique sera examinée en détail, y compris l'adaptation des questionnaires au contexte linguistique Mandé et l'exploitation des données obtenues pendant la recherche sur le champs. A titre de documentation, des détails parfois très spécifiques devront être mentionnés ci-dessous. Ils seront regroupés et introduits par un aide-mémoire précédé d'un gros point (●).

3.2.1. Méthode

- base phonétique du calcul

Partir de transcriptions phonétiques brutes de plusieurs chercheurs et en calculer des similarités lexicales d'une manière schématique, c.à.d. sans considérer chaque cas individuellement implique bon nombre de facteurs fortuits et ne mènera guère à des résultats réalistes.

- non-exploitation des données disponibles

Parmi le peu de données sur l'ensemble bozo, DKS (1953) et GTB (1980) auraient mérité attention, malgré leurs inconvénients, dès la planification de l'enquête et surtout pour une évaluation efficace des listes de mots.

L'enquête elle-même fournit un riche matériau, qui n'a pourtant pas été épuisé. Un exemple étonnant en est que les deux listes recueillies, celle de 30 mots et celle de 300 mots, n'ont pas été comparées entre elles, bien qu'il s'agisse ici d'un excellent outil pour se familiariser avec la variation en jeu et pour évaluer les données obtenues. Ainsi, le fait curieux qu'à Kouakourou (sorogama: pondori nord) un calcul de similarité lexicale entre les deux listes (et donc limité à 30 mots) expose un taux de 33% de mots apparentés seulement, est resté inaperçu⁷.

- concepts vagues

La liste de 300 mots utilisée est l'extrait d'un standard pour l'Afrique (S&S'95:63). Or, elle contient des lexèmes dont le relèvement n'est pas sans problèmes⁸. D'une manière générale, il ne nous paraît pas recommandable de relever des mots en isolation, surtout lorsqu'ils expriment des notions de mesure ou des concepts relativement vagues, car le risque de confusion pendant un recueil rapide des données sur le champ y est très grand. Pour cette raison nous avons éliminé de la liste de 300 mots les (groupes d') items suivants, parce que les lexèmes obtenus comme réponse étaient divers, exposant toute un éventail de traductions possibles de l'item questionné. Par ce fait, les taux de parentés lexicales de l'enquête ont été diminués artificiellement:

⁷ L'explication probable en est que le premier informateur parlait tieyaxo, le deuxième sorogama.

⁸ Dix items problématiques avaient déjà été éliminés par les chercheurs eux-mêmes (cf. la liste dans S&S'95:65)

haute taille (91), long (93) et grand (238);
grand (238) et gros (239);
exacte (246) > droit, moral, mesure, etc. ?
tourner (288) > retourner, traduire, faire le tour de, etc. ?

- inconsistances dans l'exploitation

Dans plusieurs cas, le calcul de similarité selon Blair, qui aurait mené à une association de lexèmes, n'a pas été strictement appliqué:

lune (26): xay (transcription alternative⁹: xai) = keu
arbre (47): juru = jugu
verser (80): fie = pii

Dans d'autres cas, l'application correcte de ce même calcul mène à de faux jugements: les données obtenues pour ya:lo 'femme' (20) sont dissociées en deux groupes¹⁰, tandis que dans lolō-ya:lo 'petite soeur' (191), qui contient la même racine, intégrée dans un mot composé, toutes les données sont réunies à cause de la longueur du mot seulement. Ici, le calcul schématique fait tort aux relations linguistiques.

En hainyaxo, il y a un élément /-ŋō/ qui apparaît très souvent en fin de mot, sans pouvoir être analysé à partir des données disponibles. Dans le calcul, cet élément n'a pas été traité d'une façon systématique. Ainsi, dans les items 7, 75 et 94 sa présence a été ignorée, tandis qu'elle a menée à des dissociations dans les autres occurrences (8, 14, 44, 48, 106, 113, 146, 150, 179, 181, 232, 302). Nous avons décidé de le considérer provisoirement comme suffixe et de l'ignorer dans notre calcul (cf. section 4.).

- erreurs de recensement

Dans les items suivants il y a eu une mauvaise réponse dans au moins un des parles. Chacun de ces malentendus aboutit à une dissociation des lexèmes en question, malgré l'existence d'une forme apparentée:

(pré-)nom (24): dyāmū 'nom de famille' en hainyaxo
voir (119): tonyo 'observer' en tiema cewe
accoucher (121): kire 'recevoir' en hainyaxo
âme (197): bodio 'cœur' en tiema cewe
arriver (284): be 'venir' dans trois variétés sorogama

3.3.2. Le Niveau Lexical

- synonymes non relevés (taux d'erreur: 8,6%)

Une source de déviation considérable au niveau de la sémantique lexicale est causée par le fait d'obtenir, dans différents parlers, des synonymes d'un item questionné, tandis que le lexème apparenté qu'on cherche est aussi en usage. Sans tenir compte de cela, on risque d'accroître artificiellement la divergence lexicostatistique entre les parlers¹¹. Nous avons pu découvrir dans l'ensemble bozo 25 cas de synonymes, à la

⁹ La transcription alternative n'est pas identique, mais raisonnablement proche de celle dans l'enquête, la différence entre les deux correspondant en plus à une variation phonétique courant du moins en tieyaxo. C'est ce genre d'examen individuel qui nous semble indispensable.

¹⁰ A tort d'ailleurs, car on a ignoré une chute consonantique.

¹¹ Dans son manuel d'enquêtes linguistiques, F. Blair (1990:29) propose une revue de la liste de mots avec plusieurs locuteurs: "Variants, near-synonyms, and synonyms should be discussed at this time, as they occur; and a decision should be made about which form to include in the word list."

place desquelles un mot apparenté usuel existe. Ceci diminue donc sensiblement la divergence entre le tiyaxo et d'autres parlers bozo, c.à.d. une variété sorogama dans la grande majorité des cas¹².

- terme général versus spécification

Certains lexèmes sont problématiques dans le contexte des langues Mandé de l'Ouest, parce qu'on doit s'attendre à obtenir plusieurs désignations spécifiques plutôt qu'un seul terme général, comme le montre un coup d'oeil dans un dictionnaire bambara ou, pour les parlers bozo, la consultation du lexique de DKS (1953). En voici un exemple: Pour l'item 'tortue' (111) l'enquête donne quatre formes différentes, leur calcul menant donc à autant de subdivisions dialectales. Seulement, DKS (1953:209) nous apprend que trois de ces quatre formes sont des sous-espèces de tortues: kũnã kerɛɛ 'petite tortue terrestre', dyembe kũũ 'grosse tortue terrestre', le troisième, kũũ 'tortue terrestre', étant le terme général. Le quatrième lexème (korokara 'grosse tortue') est un emprunt au bambara. Il en va d'une manière analogue pour lance (112), fourmis (206), araignée (207), scorpion (208), hache (228), houe (233). En diminuant les taux de parentés, chacun de ces cas a causé des distorsions dans le calcul lexicostatistique.

- emprunts non reconnus

Si on cherche à mesurer le degré d'apparenté linguistique dans une perspective généalogique, on aura intérêt à éviter des emprunts à des langues d'une autre branche linguistique. Dans deux cas seulement, des emprunts à l'arabe n'ont pas été reconnus: barika 'fort' (243), daawa 'animal' (114). Dans un cas, un emprunt au français n'a pas été reconnu (120). Pour estimer l'intelligibilité interne, un deuxième but du recensement, des emprunts sont pourtant sans importance. Des emprunts au bambara, langue apparentée au bozo, sont plus difficile à évaluer. Nous avons éliminé un mot bambara là où un terme bozo est attesté.

- euphémismes

Ce qui peut également accroître artificiellement la divergence entre parlers sont des euphémismes, pourvu que le mot non-euphémique correspondant existe également et que ce dernier est relevé dans quelques uns des parlers. Les cas suivants sont attestés en tiema cewe:

kɔdɔ-teu 'dos-chair' > 'fesses' (158),
byɛ-ni 'revenir-CAUS' > 'vomir' (63).

¹² Les chiffres suivants renvoient aux items de la liste de 300 mots, pour lesquels des synonymes ont été trouvés: 11, 33, 52, 64, 95, 123, 131, 138, 139, 148, 157, 220, 230, 240, 244, 250, 276, 294.

3.2.3. Le Niveau Phonologique

C'est ici qu'entrent en jeu des particularités phonologiques communes non seulement aux parlers bozo, mais probablement connues des langues Mandé du Nord-Ouest en général, à en juger par notre connaissance du bambara. Ceci facilite une contextualisation phonologique relativement élaborée de la liste de 300 mots au préalable. Parmi les caractéristiques générales du niveau phonologique nous constatons d'abord une extrême variation, qui complique la formulation de règles rigides, et qui nous semble exclure un calcul mécanique des taux de parentés entre lexèmes. On obtiendra de meilleurs résultats en considérant chaque cas individuellement, tout en respectant des principes de calcul explicites¹³.

L'étendue de la variation en jeu sera d'abord illustrée par l'item (6) 'oreille', selon les cas 'oreille-enfant', qui apparaît une deuxième fois comme déterminé du mot composé 'oreille-grand', désignant l'éléphant (42):

tableau 2: Les Items 'oreille' et 'éléphant'

parler bozo	'oreille' (oreille-enfant)	'éléphant' (oreille-grand)
S. Kotya	to-riō	too-ro
S. Débo	too-riē	taa-burō
S. Korondougou N	two-riō	twa-burō
S. Korondougou S	tuo-liō	twa-bulō
S. Pondori N	twɔ	tuo-rō
S. Pondori S	twɔ	tua-boro

Pour compléter cette liste, le lexème apparenté en tigemaxo connaît les trois variantes tala, tɛɛ et tɔɔ, qui ont en plus conservé la latérale à l'intervocalique.

Quant à la variation vocalique, on notera d'abord que nasalité et longueur, bien que pertinentes du moins dans des cas de paires minimales, sont presque systématiquement instables, surtout en fin de mot¹⁴. Ceci étant un phénomène bien connu dans le groupe linguistique, nous nous passerons d'en donner des exemples.

Il y a parmi les lexèmes de structure CVCV un nombre important à voyelles identiques, environ 50% en tigemaxo. C'est ce type de lexèmes qui subit couramment l'enchaînement des processus suivants: une chute de la consonne intervocalique est suivie par la dissimilation des voyelles apparentées, V1 devenant une (semi-)voyelle d'ouverture minima, conservant le trait antérieur ou postérieur de son prédécesseur¹⁵.

¹³ Nos propres principes de calcul adaptés aux parlers bozo seront proposés sous section 4

¹⁴ Il est aussi évident que la perception de ces traits et sa transcription exacte ne sont pas sans difficultés et peuvent avoir produit des déviations, surtout comme l'enquête a été réalisée par plusieurs chercheurs.

¹⁵ Un phénomène très semblable est attesté en bambara, où on trouvera p ex : ɲɔŋɔ > ɲwɔ 'semblable'; dɔŋɔ > dwá 'semaine'.

tableau 3: Chute Consonantique et Dissimilation Vocalique

item	Tigemaxo	S. Korondougou N
'lait' (18)	sombo ~ sōbo ~ sōmbo	suo
'courir' (287)	xeele	kie
'verser' (80)	fidi ~ fie	pii
'peur' (127)	kōna	kuā
'sel' (115)	xōlō	kuo

L'enquête bozo tient compte de la plupart des chutes consonantiques entre deux voyelles identiques. La liste des correspondances régulières¹⁶ doit pourtant être complétée. Il y a d'abord les chutes de consonnes suivantes, toujours à l'intervocalique: /l/ > ∅ entre le tigemaxo et le sorogama, ainsi que, moins fréquents, /s/ > ∅ et /n/ > ∅ entre le sorogama et le tieyaxo¹⁷ et /ŋ/ > ∅ entre le tieyaxo et le sorogama.

Ce qui n'a pas été reconnu dans l'enquête bozo, tout en ayant un poids statistique important (5,8% de taux de parentés supplémentaires), c'est le processus de dissimilation entre les voyelles décrit ci-dessus, qui se produit souvent après la chute consonantique (cf. le tableau 3).

Un autre phénomène, qui n'a pas été identifié dans l'enquête sont des correspondances de consonnes au niveau interdialectal. Il s'agit ici de tendances fortes, même si elles ne sont pas régulières à cent pour-cent:

tableau 4: Correspondances de Consonnes entre Parlers Différents

parler de départ	parler comparé	règle	exemple attesté
Tiema Cewe, S. Kotya	autres parlers	/j/ > /s/ /V_V	beaucoup (148), montrer (129), presser (139)
Tiema Cewe Sorogama	autres parlers autres parlers	/r/ > /l/ /V_V /g/ > /l/ /r/ /V_V	oreille (6) femme (20), arbre (47)
Sorogama	autres parlers	/s/ /c/ > /k/ /x/ /#_	endre (38), pou (46), tomber (79)

En tenant compte de tous ces phénomènes de variation et traits phonologiques énumérés ci-dessus, une modification sensible des critères du calcul est nécessaire pour établir des taux de parentés plus réalistes entre les parlers bozo.

¹⁶ Cf. S&S (1995: appendice E).

¹⁷ Ces correspondances sont également attestées à l'intérieur du tieyaxo.

3.2.4. Le Niveau Morphologique¹⁸

Des connaissances morphologiques jouent un rôle important dans l'estime des relations lexicales, le taux d'erreur remontant à 9% ici.

- suffixe du participe résultatif non reconnu

Dans deux cas, un suffixe du participe résultatif n'a pas été reconnu, ce qui a mené à une dissociation des lexèmes des autres formes: gulā-nā 'chaud' (86), mūnā-nā 'froid' (87).

- postposition non reconnue

Dans d'autres cas, une postposition non reconnue a également donné lieu à une dissociation par erreur: sībo-pa 'côté droit' (247), kwōni-pa 'côté gauche' (248).

- radical simple vs. mot composé

D'une manière systématique, „les mots composés (comprenant des items déjà inclus dans la liste)” (S&S'95:65) ont été supprimés du calcul. Les formes suivantes n'ont pourtant pas été analysées comme mots composés:

larme (186) = oeil (1) + eau (34): na-ji

barbe (173) = menton (172) + poil (155): logō-tī

Un deuxième cas est constitué par des mots composés recueillis, dont le premier terme est une traduction suffisante de l'item en question, tels que:

fesses (158): bu(:)-kudu 'fesses -morceau' (cf. DKS 1953:168,209)

corps (183): tū-teu 'corps-chair' (cf. op.cit.:265,259)

matin (231): su(b)a-so 'matin-moment' (cf. op.cit.: 25 0,246)

Ici nous avons supprimé le terme déterminant le lexème.

En plus de cela, le corpus contient six mots composés, dont le deuxième terme (i.e. le déterminé) est ju3 'enfant'. Ce dernier étant obligatoire aux cas des items 'langue' (4) et 'abeille' (205), il est pourtant facultatif dans 'oeil' (1), 'oreille' (6), 'ongle' (9) et 'soleil' (28)¹⁹. Le fait de pouvoir alors supprimer ce terme dans quatre items signifie pour le calcul des taux de parenté que d'autres dissociations artificielles ont pu être corrigées.

- ignorance de faits phonologiques et morphologiques

L'item pour coeur (101) est dans tous les parlers un mot composé de 'coeur-enfant', la forme étant bogo-liō en sorogama (dialecte korondougou sud). Sur le plan formel, il y a eu une chute de consonne à l'intervocalique dans plusieurs parlers: bogo > bo(:), la forme tieyaxo étant bo-jo(:), par exemple. Sans reconnaître ces deux faits linguistiques, les items ont été regroupés d'après le nombre des syllabes, dissociant les dissyllabiques (bo-jo, etc.) des trisyllabiques (bogo-liō, etc.). Ce même type d'erreur s'est produit avec les items endre (38): xo -lu 'endre-poudre' et doigt (182): su-xō-ne'main-os-?' (tigemaxo: su-xō-ju3 'main-os-enfant').

¹⁸ Les exemples cités ci-dessous se réfèrent au tieyaxo.

¹⁹ Ces informations proviennent d'une comparaison des listes de 30 et 300 mots ainsi que de DKS (1953).

- bases en alternance d'un même verbe non reconnues

Dans bon nombre de verbes, une dichotomie de transitivité (monovalent, bivalent) est exprimée par deux bases distinctes, l'une étant transitive, l'autre intransitive. L'ignorance de ce trait formel du système verbal a causée des erreurs du calcul de parenté du type suivant:

'frapper' (58): kwɔri (tiyaxo, hainyaxo),
 kwa (sorogama, tiema cewe),
 cf. DKS (1953:55) kɔre v.i / kwa v.t

Le même genre d'erreur s'est produit avec 'boire' (62), 'dire' (128), 'connaître' (130), 'enterrer' (168), 'prendre' (280), 'préparer' (281) et 'attacher' (297).

4. Le Calcul Lexicostatistique Révisé

4.1. Les Critères Modifiés

Notre révision du calcul lexicostatistique sera basé sur 250 données, le résidu ayant été éliminé soit à cause de du caractère douteux des données (cf. section 3, supra), soit à cause de la limitation du programme lexistat²⁰. Pour établir les similarités lexicales nous prendrons comme point de départ les critères de Blair (cf. S&S'95: appendice D) et les adaptations de l'enquête bozo (S&S'95:appendice E). Celles-ci seront modifiées par tous les traits linguistiques, phonologiques et morphologiques en majorité, que nous avons pu attester dans l'ensemble bozo (cf. section 3, supra). En voici un rappel bref:

- système vocalique: longueur et nasalité sont instables et considérées individuellement; une ouverture de plus d'un degré peut être considérée identique dans des structures CVV et CV selon les tendances générales du système vocalique;
- système consonantique: une chute de consonne intervocalique est une tendance commune sans limitation systématique, elle est prise en considération dans des structures CVV; des correspondances de consonnes entre variétés bozo (cf. section 3.2.3., tableau 4) sont systématiquement inclus dans le calcul;
- morphologie: dérivation et composition sont respectées; le cas non résolu de l'élément /ŋʒ/ en hainyaxo est considéré individuellement;

4.2. Le Résultat de la Révision

Après la révision décrite dans cet exposé, l'estime des taux de parentés présentée ci-dessous n'inclut plus d'écart d'erreur systématique. Les relations entre les parlers bozo sont résumées par les tableaux suivants²¹:

²⁰ Les items 295 à 302 ont été supprimé pour cette raison, lexistat de Schadeberg (1990) ne pouvant manier que des listes de 250 données au maximum.

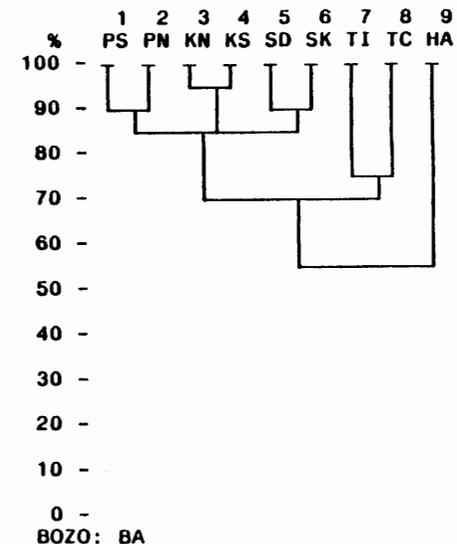
²¹ Calcul et représentation graphique sont produits par lexistat de Schadeberg (1990).

tableau 5: Pourcentages de Taux de Parentés

	1.PS	2.PN	3.KN	4.KS	5.SD	6.SK	7.TI	8.TC	9.HA
1.PS	-								
2.PN	86,0	-							
3.KN	82,4	85,5	-						
4.KS	84,1	86,7	93,5	-					
5.SD	82,2	83,3	88,1	87,8	-				
6.SK	82,2	83,7	86,6	87,9	86,6	-			
7.TI	72,4	72,8	71,9	72,7	73,6	76,1	-		
8.TC	69,0	67,5	70,8	70,7	73,9	72,7	78,5	-	
9.HA	55,0	55,1	52,7	52,1	53,4	53,9	65,8	57,6	-

1 Pondori Sud	7 Tiyaxo
2 Pondori Nord	
3 Korondougou Nord	8 Tiema Cewe
4 Korondougou Sud	
5 Sorogama Débo	9 Hainyaxo
6 Sorogama Kotya	

tableau 6: Diagramme des Relations de Parenté entre les Parlers Bozo



En comparant maintenant le calcul originel de l'enquête bozo et sa révision, la première chose à remarquer c'est que la répartition en quatre unités n'a pas changée. Ce qui a pourtant changé dans une large mesure, ce sont les taux de parentés, qui rapprochent donc l'ensemble bozo jusqu'au point où on doit reconsidérer s'il est justifié de les diviser en quatre langues. Une réponse à cette question étant remise à des recherches ultérieures, on considérera brièvement les quatre unités maintenant.

L'homogénéité interne du sorogama²² est considérable, les taux de parenté variant de 82,2% à 93,5% (versus un calcul originel de 69% à 89%) entre les deux variétés les plus éloignées géographiquement. Il est très probable qu'il y ait intelligibilité interne en sorogama, et ses subdivisions pourront être réduites aux trois variétés pondori, korondougou et débo-kotya.

Le parler le plus proche du sorogama est le tieyaxo, les deux unités partageant entre 71,9% et 76,1% (53%) de similarité lexicale, les taux correspondant toujours aux distances géographiques, un fait qui confirme la validité du calcul. Vu le fait que le tieyaxo partage 78,5% (60%) de lexèmes avec le tiema cewe, on pourra peut-être réunir ces deux variétés pour les considérer comme une seule langue. Comme le peu de données du tiema cewe sont moins fiables que la moyenne, une étude approfondie sera indispensable ici. Le parent le plus proche du tiema cewe est le débo (73,9% (51%) de mots partagés), variété de sorogama situé au sud de ce parler.

Les données du hainyaxo²³ sont probablement les moins fiables de l'enquête et n'ont pu vérifiées que par quelques items dans le lexique de DKS (1953). Il nous semble pourtant relativement sûr qu'on a ici affaire à un parler à part, dont le parent le plus proche est le tieyaxo avec 65,8% (48) de similarités lexicales.

Ouvrages Cités

- Blair, Frank 1990: Survey on a Shoestring. A manual for small-scale language surveys. Dallas: S.I.L.
- Blecke, Thomas 1996: Lexikalische Kategorien und grammatische Strukturen im Tigemaxo (Bozo, Mande). Köln: R. Köppe.
- DKS 1953 = Daget, Jacques & Konipo, M. & Sanankoua, M. 1953: *La Langue Bozo*. (Etudes Soudaniennes) Limoges: Institut Français d'Afrique Noire.
- GTB 1980: Guide de Transcription et Lexique Bozo. Bamako: DNAFLA.
- Kastenholz, Raimund 1996: *Sprachgeschichte im West-Mande. Methoden und Rekonstruktionen*. (Mande Languages and Linguistics 2) Köln: R. Köppe.
- Schadeberg, Thilo C. 1990: LEXISTAT (version 2.1). (A computer programme to calculate lexicostatistics and to construct tree models). Leiden.
- Simons, Gary 1983: Language Variation and Limits to Communication. Dallas: S.I.L.
- S&S 1995 = Smeltzer, Brad & Smeltzer, Sue 1995: Compte Rendu des Enquêtes Cartographique, Lexicostatistique et Sociolinguistique des Parlers Bozo. Avec la Collaboration de Mama Sabé. *Mandenkan* 30,55-101.
- Williamson, Kay 1989: Niger-Congo Overview. In: *The Niger-Congo Languages*. ed. by John Bendor-Samuel, 3-45. Lanham: University of America Press.

²² Les termes non3kr ou jenama nous semblent des alternatives plus courantes.

²³ Le terme alternatif kelinga est très courant.