Konstantin Pozdniakov,
INALCO – LLACAN

Etudes atlantiques comparatives : questions de méthodologie

Dans cet article, nous nous proposons de retracer quelques questions méthodologiques soulevées par les études comparatives dans le domaine africain, et en particulier concernant les langues atlantiques. Voila donc en trois points les réflexions que j'ai souhaité partager avec des collègues de domaines « marginaux » par rapport à la méthode historique comparée, c'est-à-dire travaillant hors du champ des études indo-européennes : 1) les langues africaines et la méthode comparative historique ; 2) les spécificités des études comparées dans le groupe atlantique des langues Niger-Congo ; 3) quelques éléments de « comparatisme quantitatif »[1].

1. Les langues africaines et la méthode comparative historique

La linguistique comparative tient-elle compte de la modeste expérience des études comparatives africaines ? A première vue, il semblerait que ce ne soit pas le cas. Les principaux progrès de la théorie des études comparées au XXème siècle ne viennent pas de données nouvellement accessibles aux linguistes, y compris des africanistes, qui auraient conduit à réexaminer le paradigme néo-grammairien. L'évolution de la méthode comparative au XXème siècle a été marquée par des postulats nouveaux, essentiellement saussuriens, sur la langue comme système (ce qui suppose la nécessité de tenir compte de la synchronie dans la reconstruction), sur le fait que les langues sont intrinsèquement susceptibles de diversité dialectale (ce qui suppose l'exigence d'une reconstruction pluridimensionnelle des proto-langues) et par une série d'autres idées théoriques relevant du domaine de la synchronie mais aujourd'hui totalement intégrées dans les études comparatives, grâce à Saussure, Troubetskoy, Jakobson, Benveniste, Meillet, Greenberg, Swadesh, Hjelmslev, Diakonov, Starostin, Makaev, Bloomfield, et de nombreux autres grands chercheurs dont les idées sont et seront à l'ordre du jour chaque fois qu'il est question de méthode comparative historique.

Malgré tout, certains aspects de la théorie ont été indiscutablement marqués par les données africaines et, plus généralement, non indo-européennes. Par exemple, les idées de Greenberg, qui ont beaucoup apporté à la linguistique comparative, étaient initialement formulées à partir de son expérience des données africaines. Elles ont attiré l'attention des comparatistes sur certains problèmes de classification génétique des langues qui, dans les études indo-européennes, n'auraient jamais été mises au premier plan (Porkhomovsky, 1982).

Le bagage théorique accumulé par les études comparatives africaines reste encore à évaluer. De fait, l'aspect épistémologique des études comparatives africaines est rarement examiné. C'est à cet examen que l'on voudrait se livrer ici.

De fait, il faut souligner que ce qui a permis d'aboutir aux résultats les plus marquants dans le comparatisme africain est l'application stricte de la méthode comparative historique classique. Ainsi, par exemple, les travaux d'inspiration néo-grammairienne de C. Meinhof sur le bantou n'ont rien perdu de leur actualité dans le domaine du comparatisme bantou. Par ailleurs, la spécificité des données africaines a poussé et pousse encore les comparatistes du domaine à rechercher et à appliquer de nouvelles méthodes d'analyse, et parfois à infléchir la théorie. En quoi réside donc la spécificité du matériau africain par rapport au domaine indo-européen ? Arrêtons-nous sur quelques points :

[1] Je tiens à remercier A. Donabédian et G. Segerer pour leurs remarques judicieuses.

1. La plus grande macro-famille d'Afrique (et du monde) est la famille Niger-Congo, qui compte environ 1500 langues. Le groupe atlantique est l'une des 20 filiations Niger-Congo ; il compte approximativement autant de langues qu'il existe aujourd'hui de langues indo-européennes vivantes. Ainsi, à l'échelle africaine, une famille comme la famille indo-européenne serait considérée comme un groupe de langues plutôt modeste et relativement homogène.
2. En Afrique, et particulièrement en Afrique occidentale, les langues ne sont pas seulement nombreuses, elles sont aussi particulièrement mal décrites. Dans le groupe atlantique, sur lequel je travaille, on ne dispose pratiquement d'aucune étude sur les langues safen, lehar, nyun, biafada, karon, bayot, pepel, nalu, mbulungish, sua, bullom, krim, plusieurs langues baga. Si les deux dernières décennies ont vu paraître les premiers éléments de description des langues bijogo, ndut, noon, badyaranke, buy, kisi, manjaku ou konyagi, on ne dispose de sources fiables que pour une dizaine de langues d'un groupe qui en compte une cinquantaine.
3. Dans les études comparées africaines il est pratiquement impossible de s'appuyer sur une tradition écrite. Et bien que de nombreux chercheurs appellent à ne pas exagérer le poids de l'écriture dans la reconstruction, il est difficile d'imaginer ce que serait aujourd'hui la grammaire comparée indo-européenne sans les sources écrites latines, grecques et hittites.
4. La particularité essentielle de la majorité des langues africaines provisoirement réunies dans une unité génétique comparable à l'indo-européen réside dans le fait que la distance entre ces langues, considérées comme « proches », est significativement supérieure à celle qui sépare deux langues indo-européennes quelconques. On sait que toute langue indo-européenne conserve environ 50% du lexique de base de la protolangue. Entre deux langues quelconques de groupes indo-européens différents, le pourcentage de lexique apparenté selon la liste de Swadesh est en règle générale de 30%. De l'avis de la majorité des spécialistes, ces chiffres indiquent que les langues indo-européennes ont divergé il y a environ 6000 ans. Malgré une parenté si éloignée, un débutant comparant les listes de Swadesh, par exemple en rromani et en islandais, est en mesure non seulement de découvrir environ 30% de concordances, mais même d'établir un certain nombre de correspondances phonétiques régulières entre ces deux langues. Que dire alors de la situation suivante, caractéristique du domaine africain : les ressemblances au sein de la liste de Swadesh entre deux « dialectes » (!) de la langue diola, le kwaatay et le bayot, ne sont que de 16% ! (Carlton, E. M. & Sharon R.R, 1993). Soulignons bien qu'il s'agit d'idiomes considérés comme des dialectes d'une seule et même langue, elle-même membre du groupe des langues bak, lui-même rattaché à la branche nord des langues atlantiques, elles-mêmes membres de la macro-famille Niger-Congo ! Comment interpréter ces chiffres ? D'après la table de statistique lexicale la plus connue des langues atlantiques (Sapir, 1971), entre le wolof et le temne, autre langue atlantique, le taux de ressemblance est de 3% ! Ces données sont très fréquemment citées, re-publiées, et on en déduit le plus souvent (cela est d'ailleurs évident) que ces langues atlantiques ont divergé plus tôt que, par exemple, le wolof et le ndut (18% de ressemblances, la valeur maximale pour la langue wolof !). Cependant, parmi les quantités d'ouvrages et d'articles publiés en études atlantiques comparées, aucun ne tente une interprétation théorique de statistiques aussi désespérantes. Soulignons une fois de plus qu'il s'agit bien de 3% de ressemblances pour deux langues d'un même groupe, sachant que la famille Niger-Congo compte une vingtaine de groupes de ce type.
5. L'Afrique, du moins occidentale, est caractérisée par un contexte sociolinguistique très particulier. De nombreuses études décrivent l'importance de la notion de centre et de périphérie dans la nature des changements diachroniques. Dans les zones linguistiques frontalières, où les locuteurs sont au moins bilingues, les changements ne se produisent

pas de la même manière que dans les zones centrales. En revanche, l'existence de frontières linguistiques claires facilite le travail du comparatiste. Il est manifeste que les correspondances régulières entre le malgache et les langues indonésiennes ont été favorisées par un facteur essentiel : ces langues sont séparées par l'Océan Indien, et se sont donc formées indépendamment l'une de l'autre, les contacts entre locuteurs de proto-malgache et de proto-indonésien ayant été brusquement et définitivement interrompus. On retrouve cette même situation pour les langues polynésiennes, et à un moindre degré, pour de nombreuses langues indo-européennes, bien que ce ne soit bien sûr pas le cas de toutes. La situation de l'islandais, qui s'est développé dans un isolement complet de plusieurs siècles, et, par exemple, de l'albanais, où les réflexes (les aboutissements) de racines indo-européennes sont particulièrement peu nombreux, relève de conditions sociolinguistiques totalement différentes. Concernant l'Afrique, il ne faut pas oublier que les comparatistes doivent souvent travailler avec des langues dont les locuteurs sont, par exemple, de petits groupes de nomades qui, d'une part sont multilingues (sachant de plus que l'inventaire des langues parlées par un locuteur peut changer à chaque génération), et d'autre part sont amenés à quitter leur territoire, et parfois à y revenir seulement après des siècles ou même des millénaires, dans un paysage linguistique totalement différent.

Il est clair que ces 5 facteurs, l'abondance des langues, le fait qu'elles sont mal étudiées, l'absence d'écriture, leur divergence renforcée par un contexte sociolinguistique complexe, tout cela fait que les méthodes et approches de l'étude génétique des langues d'Afrique occidentale se distinguent considérablement des approches indo-européennes. Il n'est pas exagéré de dire que le domaine africain développe ses traditions propres de linguistique comparative. En quoi sont-elles spécifiques ? Par de nombreux points, dont les suivants :

1. Nous avons déjà cité plusieurs fois le nom de Swadesh. Un indo-européaniste confirmé peut aisément ignorer cette référence, sans que cela l'empêche pour autant de travailler de manière fructueuse. Dans le domaine de la reconstruction des langues africaines, certains considèrent au contraire que comparer les cent mots de la liste de Swadesh et en déduire que deux langues ont divergé le matin du 1er janvier de l'année 5432 avant J.-C. relève justement de la linguistique comparative la plus classique. Et cela est moins exagéré qu'on pourrait le croire.
2. Dans le domaine indo-européen, comme cela a été dit plus haut, les problèmes de classification génétique des langues n'ont jamais figuré au tout premier plan. La configuration de l'arbre indo-européen est relativement simple, et elle est établie depuis longtemps. En revanche, pour le domaine africain, compte tenu de l'immense quantité de langues mal étudiées et à l'histoire inconnue, ce sont justement ces problèmes de classification génétique qui ont été d'emblée prioritaires, et sur ce plan, les apports de Swadesh et Greenberg sont déterminants. C'est peut-être pour cela que les africanistes, semble-t-il, ont accueilli avec beaucoup d'enthousiasme, et surtout avec moins de réserve, les théories de la « parenté éloignée », qui n'ont émergé dans les études comparatives classiques qu'à la fin du vingtième siècle, notamment avec les idées de la linguistique nostratique.
3. Dans un contexte aussi défavorable, où on ne parvient pas à trouver deux « bonnes vieilles langues » dotées d'une longue tradition littéraire, comme le latin et le grec, pour faire apparaître des correspondances régulières entre elles, les africanistes plus que les indo-européanistes sont tentés par des voies nouvelles, comme le recours aux méthodes statistiques pour la reconstruction, la sélection de « langues pilotes » et « langues-étalons » pour vérifier la parenté, et l'utilisation plus large des données typologiques et structurelles en linguistique comparative.

4. Contrairement aux études indo-européennes, en linguistique comparative africaine il est encore permis de publier des étymologies qui ne s'appuient pas sur les résultats d'une analyse comparative. Cela est particulièrement sensible dans les travaux sur l'ethnonymie et la toponymie africaines. Les spéculations sur les noms des pays et des peuples de l'Afrique pourraient donner lieu à un véritable roman. L'explication la plus souvent avancée est que la plupart de ces noms proviennent de malentendus entre voyageurs européens naïfs et autochtones innocents. Quelques exemples. Un navire européen jette l'ancre dans une baie pittoresque au Cap Vert. Des barques d'indigènes viennent à lui. Le capitaine demande aux autochtones comment s'appelle leur pays – et le malentendu arrive : l'indigène, suivant du regard l'index tendu du capitaine, voit une autre barque et répond avec assurance : **sunu gaal**, 'notre pirogue' dans la langue wolof. Et voila que sur les cartes européennes apparaît le nom du Sénégal. « Comment s'appelle cette ville ? » - poursuit le capitaine. Cette fois-ci, son index se dirige vers un bel arbre sur le rivage et quelques instants plus tard il note sur son carnet de bord le mot wolof **daqaar** 'tamarinier'. Ainsi naît le nom de la capitale du Sénégal, Dakar. Un bataillon d'Européens débarque sur le rivage et organise une expédition dans le continent. « Comment s'appelle votre peuple ? » - demandent les Européens aux Africains. **Bɔ am bara!** « Partez de chez nous! » répondent des maliens, jaloux de leur liberté. Et les cartes enregistrent le nom du peuple *bambara*. Une autre expédition demande dans le même temps à une femme du peuple *susu* en Guinée : « Qui es-tu ? ». Elle répond honnêtement : **gine** 'une femme' (elle aurait d'ailleurs aussi bien pu être grecque…). Le capitaine reporte sur les cartes le nom de Guinée. Et nous pourrions poursuivre sans fin. Soulignons qu'il s'agit là d' « étymologies » soutenues par de nombreux spécialistes. Le mécanisme est simple et transparent : on cherche à « élucider » un mot, on trouve dans le lexique d'une des langues concernées un terme phonétiquement proche, on impute la différence de sens à un malentendu culturel, et au fait que sur ce rivage poussent des tamariniers. Cette méthode, qui serait difficilement imaginable aujourd'hui dans la grammaire comparée indo-europénne, est la règle dans le domaine africain.

Ainsi, la linguistique comparative africaine a été, beaucoup plus que la grammaire comparée classique, confrontée au problème suivant : comment tirer des connaissances, même indirectes, à partir de rien ou presque rien ? Et sur ce plan, l'expérience des africanistes ne peut pas être dénuée d'intérêt épistémologique, si nous cherchons à avoir une idée du développement futur de la linguistique comparative au XXIème siècle. Mais soulignons malgré tout que, quel que soit l'intérêt de cette expérience, les résultats auxquels sont parvenus les africanistes sont à ce jour très modestes. Le seul acquis des études comparées africaines est la grammaire historique des langues bantu, aujourd'hui très élaborée. Pour ce qui est des autres domaines, en particulier de la grammaire comparée des langues atlantiques, des langues mandé, des langues voltaïques (gur) et de nombreuses autres, les succès sont extrêmement modestes, malgré les efforts de plusieurs générations de linguistes, parmi lequels des personnalités aussi éminentes que G. Manessy, J. Doneux, W. Welmers, etc. Certes, ces derniers domaines sont considérés comme « jeunes » dans les études comparées, mais il faut bien reconnaître que c'est là une jeunesse qui s'éternise…

Que se passe-t-il donc ? Pourquoi la théorie « patine-t-elle » ? Tous les spécialistes reconnaissent l'existence d'un arbre Niger-Congo, mais la connaisance de la configuration de cet arbre n'est pas mieux connue aujourd'hui qu'il y a un siècle. Le progrès accumulé durant le siècle dernier s'est limité à la description de langues, aux prémisses de grammaires de sous-groupes compacts, c'est-à-dire des branches les plus minces de l'arbre, mais n'a pas atteint de branches assez solides. Mais alors, dans quel sens parlons-nous de groupe et dans quel sens parlons-nous de parenté entre langues de différents groupes ? Dans quel sens peut-on, non seulement rattacher

une langue à un groupe, mais même plus généralement, parler de parenté ? Cette question n'est pratiquement jamais examinée.
Pourquoi ne dispose-t-on pour aucun groupe de langues, excepté le bantu, d'un système fiable de correspondances phonétiques, et par conséquent, des reconstructions convaincantes ? A qui la faute ? Aux linguistes, à la méthode, ou aux langues elles-mêmes ? Peut-être, après tout, les langues que nous présentons comme constituant un groupe, ont-elles divergé il y a plusieurs millénaires ? Ou bien peut-être sommes-nous incapables d'identifier les morphèmes apparentés, tant ils ont été transformés pour telle ou telle raison ? Ou encore, peut-être faut-il réévaluer une théorie qui refuse de fonctionner pour l'Afrique ?

Examinons à partir de quelques exemples concrets les difficultés tant théoriques que pratiques qui surgissent dans la grammaire comparée atlantique en particulier, et plus généralement dans le domaine Niger-Congo.

2. Les spécificités de la grammaire comparée des langues atlantiques

2.1. Le rôle des alternances consonantiques dans la nature et la vitesse des changements diachroniques

Dans la plupart des langues de la branche atlantique Nord, on trouve des systèmes très developpés d'alternances de consonnes initiales de radical, qui se rapprochent typologiquement des alternances celtiques, et servent à marquer formellement les catégories du nombre, de la classe nominale, et d'autres.

ex. en sereer : **bes** 'jour' / **pes** 'jours'
raay 'être exilé' / **taayu** 'exilé'

Les alternances concernent la plupart des phonèmes consonantiques, et, pour certaines langues, tous les phonèmes. Elles supposent en principe un changement de mode d'articulation de la consonne, mais pas de point d'articulation. Les consonnes sont groupées en séries ternaires. Chaque élément d'une série caractérise un *degré* d'alternance. La convention est de noter les degrés en chiffres romains dans l'ordre décroissant : III-II-I. Dans certaines séries, plusieurs degrés peuvent être confondus.
Voici un fragment de la reconstruction des séries alternantes des consonnes sonores et sourdes et leurs réflexes dans deux langues atlantiques, le sereer et le konyagi (Pozdniakov 1987) :

degrés	III-II-I	III-II-I	III-II-I	III-II-I	III-II-I
*proto	*mb-b-w	*nd-d-l	*nc-j-y	*ng-g-γ	*ngw-gw-γw
sereer	mb-p-b	nd-t-d	nj-c-j	ng-k-g	mb-b-w/f
konyagi	mp-b-w/β	nt-d-l	nc-j-y	nk-g-w/y	nkw-gw-w

degrés	III-II-I	III-II-I	III-II-I	III-II-I	III-II-I	III-II-I
*proto	*p-p-f	*t-t-r	*c-c-s	*k-k-(h)	*kw-kw-hw	*q-q-x
sereer	mb-p-f	nd-t-r	nj-c-s	ŋg-k-x	mb-b/w–f	nG-q-x
konyagi	p-p-ϕ	t-t-r	c-c-s	k-k-x	kw-kw-xw	k-k-x?

Dans certaines langues atlantiques, une même consonne apparaît seulement dans l'une des séries (konyagi), et dans d'autres, elle peut apparaître dans plusieurs séries (le sereer, où ***mb-b-w > b-p-b, *p-p-f > mb-p-f, *ngw-gw-γw > mb-b-w/f** c'est-à-dire que **mb** occupe le degré III dans les trois séries, alors que **b** intervient aux divers degrés de l'alternance).
Au total, les correspondances directes (degré par degré) des consonnes labiales du sereer et des consonnes du konyagi apparaissent comme suit :

degrés	III	II	I	III	II	I	III	II	I	III	II	I
*proto	mb	b	w	p	p	f	ngw	gw	γw	kw	kw	xw
sereer	**mb**	**p**	**b**	**mb**	**p**	**f**	**mb**	**b**	**w/f**	**mb**	**b/w**	**f**
konyagi	mp	b	w	p	p	ϕ	nkw	gw	w	kw	kw	xw

Ainsi, par exemple, **mb** peut régulièrement correspondre en sereer aux consonnes **mp, p, nkw, kw** du konyagi. Soulignons que la plupart de ces correspondances régulières ne sont pas dues ici à des changements phonétiques diachroniques (hormis la perte des vélaires labialisées en sereer, qui sont devenues des labiales), mais à un changement de série des phonèmes de la protolangue restés stables jusque là. Ce changement résulte de la réorganisation des caractéristiques paradigmatiques des phonèmes de la protolangue, qui dans chacune des langues-filles ont acquis de nouvelles fonctions grammaticales. Le phonème ***b** n'a pas disparu en sereer. Simplement, dans la protolangue il se trouvait essentiellement dans les mots au singulier désignant des personnes humaines (à la classe nominale des noms animés du singulier, était associé dans la protolangue le degré d'alternance II, qui suppose une opposition occlusive ~ fricative entre les formes du singulier et du pluriel, et en sereer, du fait d'une réorganisation des rapports à l'intérieur du paradigme des classes nominales, le phonème **b-** a disparu des formes du singulier, mais est apparu dans les formes du pluriel (degré I), et actuellement l'opposition entre formes du singulier et du pluriel des noms animés est marquée par l'opposition occlusive sourde ~ sonore.
De plus, dans chaque fonction marquée par une alternance (opposition de nombre pour les noms, opposition de nombre pour les verbes, formation de diminutifs, de noms abstraits, etc.), les caractéristiques morphologiques des oppositions ont leurs particularités, ce qui a des conséquences catastrophiques pour un comparatiste : la consonne **b-** du sereer peut régulièrement correspondre à n'importe laquelle des consonnes du konyagi figurant dans le tableau ci-dessus **(mp, b, w, p, ϕ, nkw, gw, w, kw, xw)** selon la fonction grammaticale en synchronie.
Pour compléter le tableau, il reste à ajouter que les systèmes du type du sereer sont si complexes, et l'équilibre des caractéristiques morphonologiques des nombreuses fonctions si fragile, que la moindre atteinte à cet équilibre provoque une réaction en chaîne qui se traduit par une cascade de changements des consonnes initiales de radical.

Une telle situation augmente considérablement le nombre d'étymologies leégitimes pour chaque mot du sereer, et par conséquent, réduit la fiabilité des étymologies proposées. Cela soulève une série de problèmes méthodologiques concrets qui ne sont pratiquement pas examinés par les théoriciens : comment représenter un tel système de correspondances en grammaire historique ? En utilisant un moyen traditionnel de représentation des correspondances régulières qui manifestement ne sera pas opérationnel ? En fait, il semble absurde de produire des tableaux de plusieurs pages contenant des indications comme :

sereer	b	b	b	b	b	b	b	b	b	b	b	b
konyagi	mp	b	w	p	p	ϕ	nkw	gw	w	kw	kw	xw

Faut-il construire un tableau distinct pour les correspondances régulières de chaque fonction, avec leur variation de langue à langue ? Cela donnerait un tableau encore plus terrible que le précédent. Faut-il se limiter à la formulation (juste !) selon laquelle toute consonne labiale de la langue 1 peut régulièrement correspondre à toute labiale de la langue 2 ? Mais comment, avec une telle liberté, proposer des étymologies lexicales ?

Tout cela signifie qu'au sens strict, tout rapprochement étymologique dans les langues à alternance consonantique suppose non seulement une justification phonétique, mais aussi un

fondement morphonologique circonstancié, ce qui complique sérieusement l'analyse comparative.

2.2. Les classes nominales

Arrêtons-nous sur deux aspects présentant à mon sens un intérêt pour la théorie du comparatisme: 1) les problèmes de la reconstruction des systèmes de classes nominales et 2) les problèmes de reconstruction du lexique dans les langues à classes nominales.

2.2.1. Problèmes méthodologiques liés à la reconstruction des classes nominales

Personne ne doute, semble-t-il, que les marqueurs proto-atlantiques de classes nominales (au moins sous leur forme indéfinie) étaient représentés par des préfixes de structure CV. Commençons par une question qui, il est vrai, est loin d'être la plus complexe dans la reconstruction des classes : comment reconstruire des classes de structure CV- sur la base du bedik et du wolof, si le bedik a historiquement perdu toutes les consonnes dans les préfixes (celle qui s'y trouvent sont une innnovation), c'est-à-dire *CV > V, et que le wolof a perdu les voyelles originelles (auxquelles se sont substituées des voyelles distancielles : **i** 'proche', **a** 'éloigné', **u** – neutre), soit *CV > C ?

Ces difficultés sont heureusement surmontables. Le bedik fait partie d'un groupe de langues étroitement apparentées (le groupe tenda), d'après lesquelles on peut restituer les consonnes perdues du bedik (cf. Ferry & Pozdniakov, 2001). Néanmoins, comment, sur la base de tels critères, peut-on déduire qu'un marqueur de classe nominale dans une langue (bedik, *CV-) est apparenté au marqueur de telle ou telle classe d'une autre langue (wolof, C) ?

Quand on compare les racines lexicales dans des langues apparentées, on dispose dès le départ de points d'appui tangibles (la signification des mots comparés est connue ; la structure canonique des racines atlantiques étant CVC, on peut compter au moins sur deux consonnes pour comparer les mots supposés apparentés), mais lorsqu'il s'agit de comparer les classificateurs nominaux, la situation se complique : il devient très difficile de reconstruire la sémantique de la classe nominale (en tout cas, on peut être sûr que les classes comparées de deux langues offriront un large inventaire de champs sémantiques, qui sera en outre très variable selon les langues choisies) ; par ailleurs, les classificateurs étant de structure CV, voire C ou bien V, les possibilités de rapprochement sont pratiquement illimitées.

Bien souvent, même dans les publications des spécialistes les plus reconnus, le seul argument pour un rapprochement étymologique de marqueurs de classes nominales est une correspondance formelle (peut-être fortuite) de consonne (Cf. par exemple Doneux 1975, une référence classique des études atlantiques comparées).

Par ailleurs, il ne faut pas oublier que les marqueurs de classe sont très souvent soumis à des changements par analogie, et que par conséquent même les marqueurs consonantiques sont loin de constituer une base toujours fiable pour des rapprochements étymologiques. Ainsi par exemple dans les langues tenda, presque tous les marqueurs de classe nominale du pluriel contiennent une consonne labiale, ce qui est sans aucun doute l'effet d'une assimilation selon le point d'articulation, ayant fait du trait de labialité la marque formelle du pluriel.

Heureusement, malgré les problèmes posés par la sémantique des classes, la situation n'est pas toujours aussi désespérée. Il est vrai qu'il reste très difficile de rendre compte de la sémantique des classes en synchronie, au point que de nombreux linguistes préfèrent justement parler d'un « asémantisme » des classes nominales. Mais le problème ne se pose pas vraiment en diachronie. En pratique, on constate que la répartition en classes de certains types de valeurs sémantiques est relativement stable. Ce qui est directement comparable, c'est la présence ou non, dans une classe donnée, d'un certain ensemble de notions.

Examinons un exemple concret. Si dans une langue donnée, une classe réunit d'une part des noms signifiant « œil », « dent », « sein » « étoile » et d'autre part des noms désignant des fruits, la première hypothèse qui se présente au spécialiste est qu'il s'agit d'un réflexe de la classe Niger-Congo 5 ***di~li**. Et si, en l'absence à ce jour de base fiable pour une comparaison génétique, nous nous risquons malgré tout à comparer directement les classes nominales d'une langue bantu (le swahili) et les classes d'une langue atlantique moderne (le peul), alors, découvrant en peul les mots **yite-re** 'oeil', **nyiinde-re** 'dent', **hoor-re** 'étoile', **dibinoo-re** 'datte', nous pourrions supposer avec un taux de vraisemblance très élevé qu'un jour, on sera en mesure de présenter des preuves de la parenté de la classe swahili **ji-** (classe 5) et de la classe peul **nde/de/re/e**.

Mais là encore il reste des problèmes. La pratique montre qu'on observe des regroupements de mots sémantiquement peu compatibles dans une même classe même quand le réflexe du morphème originel de classe ne s'est pas conservé dans la langue. Ainsi, en proto-bantu, on reconstruit un petit groupe de noms entrant dans la classe **ku-** (15) dans le cadre de la corrélation de nombre **ku-** (15) ~ **ma** (6N). Ce groupe, très réduit, contient des noms de parties du corps fonctionnant par paires (« main », « pied », « oreille », « genou »), mais aussi un nom signifiant « lune, mois, clair de lune ». Dans un certain nombre de langues d'Afrique occidentale, les noms ayant ces sens sont réunis dans une même classe, même lorsque le réflexe de la classe ***ku** est absent. Ainsi, même l'existence d'inventaires notionnels analogues dans deux langues est loin d'être toujours un fondement suffisant pour des rapprochements étymologiques de morphèmes de classes.

2.2.2. Problèmes de reconstruction du lexique dans les langues à classes

Les substantifs atlantiques présentent très souvent des cas de réévaluation morphologique de la racine, où un préfixe de classe ayant perdu sa productivité se met à être perçu comme un segment de la racine lexicale et se voit préfixer une nouvelle marque de classe. Naturellement, des cas de ce type rendent très difficile la comparaison lexicale, puisque des morphèmes lexicaux apparentés ne sont plus identifiables. Même dans l'approche la plus audacieuse des ressemblances entre mots, personne ne se risquera, dans des langues étroitement apparentées comme le bedik et le konyagi, à rapprocher les lexèmes **ɓenyakuɗ** (bedik) ~ **xu-ɗɛx** (konyagi) 'feu'. Pourtant, les racines sont bien apparentées. En bedik, la racine atlantique ***dox** est reflétée par la consonne **ɗ**, **ɓe-** est un préfixe de pluriel (le mot entre dans la classe des *pluralia tantum*), **nya-** est un préfixe de singulier (sachant que la forme **nyakuɗ** n'est pas attestée telle quelle en bedik), **ku-** est un réflexe de la classe Niger-Congo 15 ayant perdu sa productivité en bedik et intégré à la racine (alors qu'en konyagi la classe **xu-** est restée productive). Par ailleurs, la consonne initiale radicale **ɗ** fait partie en bedik de la série d'alternances initiale **n-/ɗ-**, et en konyagi, de la série **nd-/ɗ-/r-**. Et puisque, comme nous l'avons dit plus haut, chaque classe est associée dans ces langues à un des trois degrés d'alternance de la consonne initiale, la coïncidence entre les réflexes **ɗ** en bedik et en konyagi est totalement fortuite : l'intégration d'un nom à une nouvelle classe conduit automatiquement au changement de la consonne radicale, et dans le mot **benyakuɗ** il pourrait tout à fait ne s'être conservé aucun élément de la protoforme. Soulignons que cet exemple n'est pas un cas exotique et marginal, c'est un type assez standard de reflet d'un lexème proto-atlantique dans une langue atlantique.

La racine peut intégrer également un suffixe de classe nominale, ce qui crée parfois curieusement une fausse impression de métathèse. En proto-atlantique, une forme nominale définie avait, semble-t-il, une marque de classe de structure confixale CV-...-VC. Dans une série de langues du groupe mel (par exemple en temne), on observe une série d'exemples montrant l'évolution suivante : ***kə-kot-ək** 'pied' > **ko-tək** (avec chute du préfixe dans le cas où sa consonne coïncide avec la consonne initiale radicale, et par la suite réinterprétation de la racine).

Dans d'autres langues atlantiques, où ce phénomène ne s'est pas produit, c'est la racine originelle **kot** qui est représentée. Il est curieux que la même évolution soit observée dans un mot grâce auquel les langues sud-atlantiques sont souvent appelées langues mel (proposition de D. Dalby). Dans la plupart des langues de ce groupe, c'est la racine **mel** qui signifie 'langue (anat.)', alors que les langues nord-atlantiques conservent une des racines Niger-Congo les plus stables (si ce n'est la plus stable) **dem ~ lem**. Le plus souvent, on considère justement que la racine **mel** résulte d'une métathèse. Cependant, la comparaison avec d'autres mots semblables permet de proposer la reconstruction suivante : ***lə-lem-əl** 'langue' > **le-mel.** Ajoutons que les traces de la forme originelle définie des substantifs sont le plus souvent conservées dans les noms de parties du corps.

.Des problèmes de ce type apparaissent dans les études comparatives des langues atlantiques, mais aussi de beaucoup d'autres. Il est cependant des problèmes qui semblent spécifiques aux langues atlantiques, puisque pour autant que je sache, seules ces langues ont un système de classes nominales aussi étroitement lié aux alternances radicales morphonologiques observées plus haut.

Examinons le cas du wolof, qui semble unique (la question est approfondie dans Pozdniakov, 1991). Les résultats de l'analyse statistique et la reconstruction interne des consonnes initiales du proto-wolof font apparaître un phénomène intéressant pour la théorie des études comparatives. En wolof il n'y a plus de marques de classes dans les substantifs, sauf quelques marques figées. Les marques de classes ne fonctionnent que dans les formes dépendantes, y compris les déterminatifs définis postposés aux substantifs. Cependant, la qualité de la consonne initiale d'une base lexicale dépend de sa classe nominale. Par exemple, malgré l'absence d'alternances consonantiques à l'initiale en wolof, les substantifs dans les classes nominales **M** et **L** ont une consonne forte à l'initiale : une prénasalisée pour les sonores et une occlusive pour les sourdes. En revanche, la plupart des noms de la classe **W** ont une consonne faible à l'initiale – une sonante ou bien une fricative sourde : **ndox mi** 'eau', **po mi** 'jeu', **ndab li** 'calebasse', **cin li** 'marmite', **waas wi** 'carpe', **fas wi** 'cheval'.

Ce genre de corrélation entre classes et consonnes initiales des bases lexicales est typique des langues atlantiques. Les alternances consonantiques ont pour base le mode d'articulation (occlusives, fricatives, prénasalisées, etc.). Ce qui fait la particularité du wolof, c'est que les changements s'y font *aussi* en fonction du lieu d'articulation : la consonne initiale des substantifs a une tendance très marquée à adopter le lieu d'articulation de la consonne de la classe nominale ! Ainsi, en wolof il n'existe pas un seul substantif de la classe **L** qui ait, par exemple, une consonne labiale à l'initiale indépendamment du degré d'alternance attaché à cette classe : aucun nom ne commence par **mb-**, **p-**, **m-**, **f-** ou **w-**. L'affectation d'un nom proto-wolof dans la classe **L** a été suivie d'une double modification de sa consonne initiale : celle-ci devient dentale et adopte le degré d'alternance III (fort) requis par la classe **L**. Ainsi, ***b-** > **mb** > **nd** ; ***w** > **mb** > **nd** ; ***f** > **p** > **t** ; ***g** > **ng** > **nd**, etc.

De même, dans les classes dont les marques comportent une labiale, les vélaires initiales des racines lexicales se sont labialisées. Par exemple, dans la classe **M** : ***g** > **ng** > **mb** ; ***d** > **nd** > **mb** ; ***k** > **p** ; ***n** > **m**, etc., dans la classe **W** : ***j** > **y** > **w**, etc., dans la classe **J** (degré II d'alternance) : ***ng** > **g** > **j** ; ***k** > **c**, etc.

Ces deux processus parallèles (changement de série suivi d'un changement d'ordre) rendent possible la correspondance entre une consonne initiale wolof et n'importe quelle consonne initiale d'une autre langue atlantique, surtout si l'on tient compte des multiples reclassements des substantifs en fonction de critères sémantiques.

La reconstruction de ce phénomène propre au wolof permet de rapprocher, par exemple :

- dans la classe **B** (degré II d'alternances) : wolof **bopp b-** 'tête' – atl. ***gop** ; wolof **bët b-** 'œil' – atl. ***ngit** ; wolof **bir b-** 'ventre, intérieur' – peul **nder** 'à l'intérieur' (phonétiquement il est possible également de rapprocher wolof **bir** avec peul **ɓer-nde** 'cœur') ;
- dans la classe **W** : wolof **weer w-** 'lune' – wolof **leer g-** 'lumière' (atl. ***l-**) ; wolof **ween w-**

'sein' – wolof **dënn w** 'poitrine' (atl. ***d-**) ;

- dans la classe **G** : wolof **guy g-** 'baobab' – atl. ***bok** (wolof **buy b-** 'fruit de baobab') ; wolof **guddi g-** 'nuit' – atl. ***bud-** ;
- dans la classe **S** – wolof **suuf s-** 'terre' – atl. ***lof**
- dans la classe **L** : wolof **ndab l-** 'calebasse', récipient' – sereer **o gamb ol** 'calebasse (sp.) – peul (Est) **gummb-al** 'gourde'.

On comprend que ces changements diachroniques ne facilitent guère la comparaison du lexique du wolof et d'autres langues atlantiques.

2.2.3. Changements par analogie

Une des difficultés principales de la reconstruction atlantique est liée aux nombreux cas de changements par analogie à grande échelle. Avant tout, ils sont observés dans le système des classes nominales, dans les séries des pronoms et dans la dérivation verbale.

Les changements par analogie sont un type universel de changement linguistique. On en trouve de nombreux exemples dans toute grammaire historique. Malgré tout, il semble que l'étude des changements par analogie soit liée à une des perspectives théoriques essentielles de la linguistique comparative du XXIème siècle. Quelles unités linguistiques sont soumises aux changements par analogie, dans quels cas se produisent-ils, quelle est leur fonction, quels sont les facteurs qui appellent un changement donné ? Malgré les dizaines de publications consacrées au changement par analogie dans diverses langues, il n'existe pas de théorie de ces changements.

Examinons l'exemple connu des flexions verbales sémitiques (d'après H. Koch, 1966, p. 220) :

	1Sg.		2Sg.M
Akkadian	-ku		-ta
Arabic	-tu	←	-ta
Ethiopic	-ku	→	-ka
Proto-Semitic	*-ku		*-ta

Les flèches indiquent l'orientation du changement par analogie[2].

La théorie linguistique ne fournit pas de réponse à la question suivante : qu'est-ce qui change exactement par analogie en arabe ou dans les langues éthiopiennes ? Quelle unité ? A quel niveau linguistique ? Quel est le statut de **t-** en arabe et **k-** dans les langues éthiopiennes pour les paradigmes concernés ? Ce ne sont pas des phonèmes, puisque l'on peut attribuer à ces éléments un signifié (« locuteur »). Ce ne sont pas des morphèmes : en arabe, on ne peut pas segmenter **-tu** en deux morphèmes **-t-** et **-u**.

Les problèmes liés aux submorphes sont examinés en détail dans (Pozdniakov, 2003). Arrêtons-nous seulement sur quelques points essentiels qui concernent non seulement la question des submorphes en synchronie, mais également l'interprétation théorique des changements diachroniques par analogie.

1. Plusieurs paradigmes morphologiques (marques de classes nominales, pronoms, affixes de dérivation verbale, marques aspectuelles, etc.) s'organisent en système grâce à des signes qui sont distincts des morphèmes. Ces unités paradigmatiques, les submorphes, sont représentées le plus souvent par des segments de morphèmes ; ainsi, en anglais, **th-** à l'initiale est une consonne caractéristique des pronoms et des articles, **b-** dans plusieurs langues du groupe atlantique est une consonne caractéristique des classes nominales du pluriel, **-u** en haoussa est une marque des pronoms du pluriel (**mu, ku, su**), etc. Ces submorphes, qui sont porteurs

[2] On trouvera de nombreux exemples de ce type dans les systèmes pronominaux d'autres langues africaines dans (Pozdniakov & Segerer, 2004)

d'une valeur, peuvent être représentés par d'autres unités linguistiques que des segments des morphèmes : la langue peut charger de valeur sémantique une structure syllabique : en français, par exemple, ainsi qu'en russe, les pronoms de 3ème personne ont une structure V(C) – **il, elle, on** – tandis que les autres pronoms (1ère et 2ème personnes, les interlocuteurs) ont une structure CV (**je, tu, nous, vous**). En peul, en mende ou en irlandais, des traits distinctifs de phonème (« surdité », « fricativité », « nasalité ») sont chargés de signification grammaticale. Autre exemple d'emploi d'un trait distinctif de phonème comme signe : en wolof, devant le suffixe d'inversif **–i** la consonne finale de la racine verbale est « renforcée » : la consonne finale devient une géminée occlusive : **ub** 'fermer' ~ **ubb-i** 'ouvrir, **fas** 'nouer' ~ **fecc-i** 'dénouer'. Il existe de nombreux autres types de submorphes.

2. Ces marques submorphémiques[3] se constituent le plus souvent à partir de changements diachroniques par analogie, donc hors changements phonétiques réguliers, mais apparaissent pourtant avec une régularité étonnante dans les langues. Ces changements par analogie, qui sont bien connus des comparatistes, n'ont pour ainsi dire jamais été étudiés du point de vue de leur fonction synchronique. Dans la théorie, on ne pose pas la question de savoir pourquoi des morphèmes qui ont une signification proche se transforment phonétiquement en harmonisant leurs segments.
3. Comme j'ai essayé de le démontrer en utilisant des données variées venant de langues de familles différentes, la fonction essentielle de ces marqueurs submorphémiques est la même que celle des neutralisations morphologiques. Une neutralisation morphologique (l'absence de distinction des membres d'une opposition dans un contexte – **je *leur* donne**..., Datif = ***leur* maison**, Possessif, et le maintien de cette distinction dans un autre contexte – **il *nous* donne**..., Datif, mais ***notre* maison**, Possessif) opère un « agrafage » entre deux paradigmes différents (dans notre exemple, les pronoms possessifs s'intègrent dans le système des pronoms, grâce à la neutralisation de l'opposition Datif ~ Possessif à la 3ème pl.). Les submorphes fonctionnent de la même façon. Cependant, le changement par analogie portant sur des segments submorphémiques est une opération moins radicale que la neutralisation morphologique, qui suppose l'unification totale de deux morphèmes dans certains contextes.
4. La distribution des neutralisations morphologiques, tout comme celle des unifications submorphémiques, n'est pas arbitraire, puisque soumise à une règle qui les fait dépendre de l'opposition grammaticale entre membres marqués et non-marqués. Par exemple, ce n'est pas un hasard si dans la plupart des langues à classes nominales ou à genres, l'inventaire des marques plurielles est plus réduit que celui des marques du singulier, et que l'inventaire des pronoms pluriels est plus restreint que celui des pronoms singuliers (en français **le / la** ~ **les / les**, en wolof 1ère personne pl. = 3ème personne pl. dans la série des pronoms Sujet – **nu**, 2ème personne pl. = 3ème personne pl. dans les pronoms Objet – **leen**, ou les pronoms possessifs – **seen**). La raison en est la suivante : dans un signe «cumulatif» ou « polyfonctionnel » (comme c'est le cas pour les pronoms, qui cumulent plusieurs grammèmes ou catégories : personne, nombre, genre, fonction, inclusivité, statut énonciatif, et d'autres encore), il existe une répartition très nette des rôles. Un signe qui « travaille » plus pour une catégorie (la catégorie de nombre, par exemple), en étant le membre marqué, c'est-à-dire porteur d'une valeur (la plupart des marqueurs du pluriel dans des oppositions de nombre), « travaille » moins pour une autre catégorie (la catégorie de personne, par exemple). Lorsque ce signe travaille moins pour une autre catégorie, il perd sa force distinctive et ouvre son espace à des neutralisations qui exercent leurs fonctions intégratives (par exemple les neutralisations des marqueurs du pluriel dans l'opposition de personnes ou de genres).
5. Les neutralisations morphologiques et submorphologiques ont tendance à être en distribution complémentaire. Si, par exemple, les neutralisations morphologiques sont chargées d'intégrer dans le système les marques de fonctions (en français, **nous** Sujet = **nous** Objet, **vous** Sujet = **vous** Objet), et les membres de l'opposition de nombre (en français, **se** sg. ~ **se** pl. Réfléchi), alors les unifications submorphologiques touchent un autre domaine – celui de catégorie de personne (par exemple, 1ère ≈ 2ème **–ous** dans **n-ous**, **v-ous**, **-otre** dans

[3] R. Jakobson qui les a décrites dans les paradigmes de déclinaison en russe les appelait des «indices» (*primety* en russe) (Jakobson, 1958 ; malheureusement cet article fondamental n'a pas été inclu dans la version française des œuvres choisies de R. Jakobson), et A. Meillet – dans le même contexte – les appelait des «consonnes caractéristiques de désinence».

n-otre, votre, -os dans **n-os, v-os** ; $1^{ère} \approx 2^{ème} \approx 3^{ème}$: **-ien** dans **m-ien, t-ien, s-ien**).

6. L'étude des fonctions synchroniques des unifications totales (morphèmes) ou partielles (submorphes) des membres d'un paradigme amène à un constat important pour les études comparatives : les changements diachroniques par analogie ne sont pas arbitraires, ils sont destinés à maintenir un équilibre fragile entre ces unifications pour un état donné de la langue. Selon la configuration de cet état, on peut s'attendre à ce que les changements par analogie portent sur telle ou telle partie du paradigme. Tout se passe comme si la synchronie donnait une instruction ou, du moins, une permission pour tel ou tel changement par analogie. On voit donc que l'étude des fonctions synchroniques des changements par analogie n'est pas étrangère à la théorie des études comparatives.

2.2.3.1. Les changements par analogie dans les classes dimensionnelles du peul

On trouve un exemple particulièrement clair d'organisation submorphémique stricte dans le paradigme des classes nominales dimensionnelles du peul. Le système développé de diminutifs et d'augmentatifs est une innovation certaine du peul. Il est apparu et s'est formé en peul de manière isolée, après que le peul se soit séparé des autres langues atlantiques. Les classes dimensionnelles du peul ne sont pas exclusivement fondées sur la sémantique. Elles se distinguent des autres classes par de nombreux traits formels, qui font de ces classes dimensionnelles un sous-système de fait.

Le peul est une langue représentée par un grand nombre de variétés dialectales. Globalement, deux systèmes de classes dimensionnelles différents existent dans les parlers peuls : le système des parlers occidentaux et centraux d'une part, et le système des parlers orientaux de l'autre. Examinons ces deux systèmes[4].

Parlers non orientaux :

Diminutifs :	
Sg.	*Pl.*
II … ngel	III … koy (FJ) ~ III … kon (FT)
II … kel (FJ) (dim. intensif)	
II … kun (FJ)	
II … kal (partitif)	
Augmentatifs :	
II … ngal (FT, M)	II … (pas de marque spécifique)
II … ngil / ngii (FJ) (animaux, péjoratif)	

Parlers orientaux :

Diminutifs :	
Sg.	*Pl.*
II … ngel	III … kon (koñ,kol) ~ III … koy (Centre)
II … ngum (dim. intensif)	
II … kal (partitif)	
Augmentatifs :	
III … nga	III … ko
III … ngi (Ansongo)	

Examinons les marques formelles qui permettent l'intégration des classes dimensionnelles en un système structuré.

Pour les parlers non orientaux :

[4] Les tableaux sont basés sur (Koval, 1997). Parlers : FJ – Fouta Jallon, FT – Fouta Toro, M – Macina.

1. Toutes les marques ont la structure CVC ; de plus, cette structure n'est pas la même que celle des classes non-dimensionnelles (-CV).
2. Toutes les marques contiennent une consonne vélaire à l'initiale : **k-**, **ng-**.
3. Toutes les marques ont une sonante en finale : **-n**, **-l**, **-y**.
4. Toutes les classes du singulier sont marquées par le degré II d'alternance (occlusives).
5. Les classes sg. s'opposent aux classes pl. par la qualité de la voyelle.

Dans les parlers orientaux, le système est structuré différemment :

1. La qualité vélaire de la consonne initiale reste une marque générale, ainsi que le contraste des voyelles – dans les diminutifs comme dans les augmentatifs.
2. La voyelle [o] devient une marque spécifiques des classes dimensionnelles du pl..
3. La structure des augmentatifs (CV) est opposée à celle des diminutifs (CVC).
4. Les alternances s'organisent différemment : on a d'une part une opposition de degrés sg. II ~ pl. III dans les classes diminutives, et d'autre part une identité de degrés pour les classes augmentatives (III – sg. et pl.). Ce dernier trait caractérise également les parlers non orientaux : opposition II/III pour les diminutifs, et absence d'opposition (cette fois degré II) dans les augmentatifs.

Finalement, chaque segment de chaque marque de classe est chargé d'un rôle intégratif ou distinctif dans les divers paradigmes de la catégorie de la dimension. Une fois de plus, on constate que l'évolution des phonèmes dans les marques des classes nominales n'est pas déterminée par des changements phonétiques réguliers, mais plutôt par la pression structurelle vers une distribution équilibrée des ressemblances et des distinctions synchroniques. On a là de nouveaux arguments pour supposer une transformation radicale du proto-système, dont même les contours sont totalement masqués par les multiples changements par analogie.
Pour conclure sur ce point, soulignons une fois encore que l'idée de submorphe reste très marginale dans la théorie des études comparatives, qui préfèrent opérer avec les variantes des unités de niveau (allophones, allomorphes). Je souhaite ici attirer une fois encore l'attention des linguistes sur les éléments « hors niveau » qui, en tant que signes paradigmatiques spécifiques, jouent un rôle essentiel dans les changements par analogie[5].

Les alternances atlantiques, les classes nominales et les changements par analogie à grande échelle illustrent bien les difficultés qui apparaissent dans les études comparatives atlantiques. Pour dépasser ces problèmes dans la construction de grammaires comparées « jeunes », la statistique peut se révéler un instrument puissant. La troisième partie est consacrée à quelques aspects du « comparatisme quantitatif ».

3. Eléments de comparatisme quantitatif

Le champ dans lequel le recours aux méthodes statistiques est le plus répandu est un domaine relativement périphérique des études comparées, et qui sort d'ailleurs du cadre des procédures traditionnelles de la méthode comparative historique : les études glottochronologiques, qui visent à dater la séparation de langues apparentées. Mais les méthodes statistiques peuvent également être utiles dans d'autres domaines clés de la linguistique comparative, et parfois même y jouer un rôle de premier plan.

Le potentiel énorme des méthodes statistiques dans les études comparatives apparaît clairement si on prend en considération le fait que chaque différence significative entre des langues apparentées qui résulte d'un processus d'innovation entraîne automatiquement une

[5] V.V. Ivanov, qui, depuis de longues années, poursuit les idées pionnières de Jakobson sur les submorphes, considère leur étude comme une des perspectives les plus prometteuses pour la linguistique de demain. L'un de ses derniers ouvrages, intitulé "La linguistique du troisième millénaire : interrogeons l'avenir ", comporte un chapitre particulièrement intéressant : "Les submorphes et leur importance pour la diachronie" (Ivanov, 2004). Je profite de cette occasion pour remercier V. V. Ivanov d'avoir attiré mon attention sur les travaux de Bogoroditskij (1917) qui, semble-t-il, fut le premier (bien avant Jakobson) à se pencher sur le rôle des submorphes dans le système casuel de l'indo-européen.

différenciation des caractéristiques quantitatives des unités qui les constituent. Ainsi, la palatalisation des consonnes vélaires devant une voyelle antérieure entraîne une augmentation de la fréquence des palatales et une diminution de la fréquence des vélaires dans le lexique. Cela signifie que la répartition statistique actuelle des morphèmes, des lexèmes, des phonèmes, et d'autres unités de langues apparentées reflète, directement ou indirectement, les principaux processus de changement linguistique (à quelques rares exceptions), dont la reconstruction est justement l'objet principal des études comparées.

On peut en déduire, à l'inverse, que **toute divergence des caractéristiques quantitatives de deux systèmes apparentés doit trouver une explication dans le domaine qualitatif**. Cela signifie qu'au lieu d'un domaine de faits, le comparatiste dispose de deux domaines contigus : celui des divergences qualitatives et celui des divergences quantitatives.

D'une part, il arrive fréquemment que les méthodes qualitatives de comparaison et de reconstruction aboutissent à des contradictions pour lesquelles aucune solution ne s'impose. Dans cette situation, bien connue de tout comparatiste, l'outil statistique peut apporter des arguments convaincants pour réduire le choix des possibilités.

D'autre part, les méthodes statistiques permettent d'évaluer, parfois même avec précision, la qualité des résultats obtenus par les méthodes qualitatives.

Enfin, les méthodes statistiques se révèlent efficaces dans les nombreux cas où les méthodes traditionnelles sont limitées par le niveau insuffisant de description de la langue.

On peut distinguer au moins cinq domaines principaux de l'analyse comparative pour lesquels le recours aux méthodes statistiques donne des résultats positifs :

1. La mise en évidence de caractères morphologiques ou morphonologiques archaïques.

Par exemple, dans de nombreuses langues ayant un système de classes nominales développé (comme les langues bantoues ou atlantiques), la comparaison des fréquences phonémiques lexicales dans les substantifs et dans les verbes permet de reconstituer des séries d'alternances consonantiques perdues (cf. l'exemple du wolof examiné ci-dessus), mais également de mettre en évidence des traces de classes figées à l'intérieur des racines nominales. Pour mettre en évidence de tels préfixes figés il suffit parfois de comparer les fréquences des phonèmes initiaux des bases nominales et des bases verbales. Ainsi par exemple en noon (sous-groupe cangin), la fréquence du **p-** initial est de 12% pour les noms, et de 5% seulement pour les verbes. La pratique montre que des divergences de ce type apparaissent là où la racine nominale comporte en réalité un préfixe de classe nominale ayant perdu sa productivité. C'est ainsi que nous avons analysé plus haut la racine Niger-Congo ***dem** 'langue'. En noon, la racine désignant cette notion est **perem**, mais le recours à la statistique pour la reconstruction interne permet de soupçonner qu'en proto-noon ce mot était segmenté en **pe-rem** avec le marqueur de classe ***pV-** et la racine **rem** (cf. dans les autres langues atlantiques : 'langue' – manjaku (dial.) **pə-remt**, mancagne **pə-diim**, pepel **pə-rɛmtɛ**, diola (dial.) **fu-rim** ~ **hu-rim**, jaad **pɛ-dɛɛma**, biafada **bu-dɛɛma**, nyun **bu-lemuc**, limba **fi-liŋ**, etc. Bien entendu, toutes les disparités fortes de fréquence entre racines nominales et verbales ne signifient pas que la racine nominale présente des traces d'un ancien marqueur de classe. Ainsi, la comparaison des fréquences en noon aboutit au résultat suivant pour les séries de sonores prénasalisées et de sonantes (les fréquences sont indiquées en pourcentage) :

	mb	nd	nj	ng	total	w	l	y	total
nom	5	4	2	4	15	2	8	2	12
verbe	1	2	2	1	6	5	12	3	20

Les substantifs présentent systématiquement plus de prénasalisées (en moyenne 9% de plus que pour le verbe) et les verbes plus de sonantes (les noms en présentent en moyenne 8% de moins, soit environ la même différence). Une telle répartition complémentaire témoigne, en règle générale, non pas de l'existence de classes archaïques dans les langues atlantiques, mais de la nécessité de reconstruire les alternances de consonnes initiales radicales dans la protolangue.

2. La mise en évidence des contours généraux d'un système de correspondances régulières sur la base d'une comparaison des fréquences phonémiques lexicales dans des langues apparentées.

L'approche la plus efficace est de croiser plusieurs méthodes quantitatives complémentaires, ce qui permet de mettre d'emblée en évidence les contours des systèmes de correspondances, avant même de comparer des lexèmes en particulier. Prenons un cas concret : la comparaison des fréquences lexicales des consonnes initiales en guro et en yaure (langues mande) donne la répartition suivante (dans le tableau, le signe « + » désigne une fréquence consonantique significativement supérieure dans une langue par rapport à une autre langue ; le signe « – » signifie une fréquence significativement moindre) :

	b	p	v	f	g	k	g'	d	t	z	s
guro	+	–	+	–	+	–	+	+	–	+	–
yaure	–	+	–	+	–	+	–	–	+	–	+

Pour établir des correspondances phonétiques régulières entre le guro et le yaure, il n'est pas nécessaire de comparer le lexique de ces langues. Pour chaque langue prise séparément, les fréquences montrent sans ambiguité une correspondance en système entre les consonnes initiales sonores du guro et les sourdes du yaure. Donc, on peut prévoir, avant même de consulter la documentation lexicale, qu'aux consonnes sourdes du yaure correspondront des consonnes sonores en guro.

3. La reconstruction interne du système phonologique de la proto-langue et des processus de changement phonétique à partir des données d'une langue, sur la base d'une comparaison des fréquences phonémiques lexicales dans les différentes positions. Le recours à la méthode statistique de reconstruction interne, qui est particulièrement important dans le cas de langues isolées dont l'apparentement génétique n'est pas établi.

4. La statistique, qui permet non seulement d'obtenir des résultats dans le domaine des études comparées, mais aussi de vérifier leur fiabilité.

Cet aspect de l'analyse statistique est particulièrement important dans les études comparatives africaines. Pour de nombreux groupes de langues, les recherches sont conduites dans le meilleur des cas par un ou deux spécialistes, et les résultats obtenus ne font généralement pas l'objet de vérification ou de contradiction de la part de la communauté scientifique. La statistique peut ici jouer le rôle d'arbitre. Comment vérifier quelle série de correspondances est moins valide qu'une autre ? Peut-on dire quelle langue fournit les données les moins fiables dans un dictionnaire étymologique comparatif ? Je pense qu'il est possible de tenter de répondre à ces questions en s'appuyant sur la statistique. La méthode repose sur un axiome simple. Dans chacune des langues d'un ensemble génétique à reconstruire, les réflexes d'un protophonème donné ont une fréquence stable qui est égale à la fréquence des réflexes de ce phonème dans la protolangue. Quels raisons avons-nous pour adopter cet axiome comme point de départ ? Supposons que la protolangue X avait entre autres les phonèmes ***k-** et ***c-**, et que leurs fréquences dans le lexique étaient respectivement de 10% et 5%. Supposons que dans une langue 1, issue de la protolangue X, ces phonèmes n'ont pas changé, c'est-à-dire que ***k** > **k**, ***c** > **c**. Par ailleurs, la langue 1 a perdu 50% du lexique de la protolangue. La question est : à quelles fréquences pourrions-nous nous attendre dans la langue 1 pour les réflexes des phonèmes ***k-** et ***c-**? Il semblerait logique de supposer qu'elles sont inchangées : si la langue 1 n'a conservé que 1500 des 3000 racines de la protolangue, qu'est-ce que cela peut changer aux fréquences relatives des réflexes ***k-** et ***c-**? Si les proportions étaient sensiblement différentes de celles de la protolangue, nous devrions considérer que les réflexes d'un phonème donné sont, dans cette langue 1, plus « volontiers » remplacés par les emprunts et innovations que les réflexes d'un autre phonème, ce qui semble ne pas avoir d'explication logique. On attendrait donc 150 réflexes de **k** et 75 de **c**.

Un autre exemple : supposons que dans une langue 2, ***c** > **c**, ***k** > **k** devant /u, o, a/, mais ***k** > **c** devant /i, e/. La question est : à quelles fréquences pourrions-nous nous attendre dans

cette langue 2 pour les *réflexes* des protophonèmes ***k-** et ***c-** (et non, soulignons le bien, pour les phonèmes **k-** et **c-**) ? Comme pour la langue 1, elles doivent rester stables, indépendamment de la quantité de lexique de la protolangue qui a été perdu dans la langue 2.

En tout cas, si le dictionnaire étymologique comparé de ce groupe de langues est établi, et qu'il est construit sur la base du système de correspondances phonétiques régulières entre les langues, on peut aisément compter les fréquences des réflexes de chaque protophonème dans chaque langue. Il est également aisé de comparer ces fréquences avec celles des protophonèmes correspondants dans le lexique du dictionnaire étymologique. Les différences de fréquences (qu'elles soient inférieures ou supérieures) correspondront très vraisemblablement à des erreurs de l'analyse comparative.

En pratique, cette méthode montre généralement que dans les dictionnaires étymologiques comparés, environ 10% du lexique est mal interprété. Souvent, la méthode statistisque incite, sinon à mettre en cause les résultats concrets de l'analyse comparative, du moins à y réfléchir. Prenons un exemple concret. Dans les recherches comparatives, on suppose implicitement que plus nous trouvons d'exemples pour une correspondance donnée, plus cette correspondance est fiable. En réalité, ce n'est pas toujours le cas, comme j'ai pu personnellement le constater en examinant les données du dictionnaire nostratique de V. Illich-Svitych (précisons que d'après le critère ci-dessus, le matériau nostratique présente un degré élevé de fiabilité, puisque son taux d'erreur est inférieur à 10%, ce qui est un excellent résultat même pour des groupes génétiques de langues beaucoup plus jeunes). Le dictionnaire étymologique de Illich-Svitych est destiné, on le sait, à démontrer la parenté existant entre 6 grands groupes de langues : indo-européen, afro-asiatique, kartvélien, ouralien, dravidien et altaïque. En protonostratique, Illich-Svitych reconstruit deux occlusives vélaires sourdes : ***k_1-** et ***k_2-**. La distinction entre les réflexes de ces deux consonnes est fondée exclusivement sur les langues afro-asiatiques (on considère que dans les cinq autres branches ces réflexes ont convergé). Cela signifie qu'on ne reconstruit ***k_2-** que quand il existe un réflexe afro-asiatique. Il s'ensuit que d'après le dictionnaire, l'afro-asiatique a conservé 100% des mots nostratiques en ***k_2-**, alors que globalement (selon le dictionnaire) l'afro-asiatique a perdu 50% du lexique nostratique ! D'autre part, dans les nombreux cas où il n'existe pas de réflexe afro-asiatique, ce qui ne permet pas de choisir entre ***k_1-** et ***k_2-**, Illich-Svitych reconstruit conventionnellement ***K-** (occlusive vélaire sourde non spécifiée). Il est tout à fait naturel que parmi les réflexes de ***K-** le dictionnaire ne comporte pas un seul réflexe afro-asiatique, bien que l'afro-asiatique ait perdu 50% du lexique nostratique et non 100% ! Une répartition complémentaire de ce type en afro-asiatique (trop de réflexes de ***k_2-** et absence complète de réflexes de ***K-**) montre sans équivoque la chose suivante : parmi les exemples que Illich-Svitych fait remonter avec précaution à ***K-**, de nombreuses séries de correspondances remontent à ***k_2-** et non à ***k_1-** (Pozdniakov, 1993).

Cette méthode permet également de déduire que : 1) le matériau qui est traité avec le moins d'exactitude dans le dictionnaire nostratique est le kartvelien (environ un quart des réflexes kartvéliens du dictionnaire est douteux) ; les données les plus exactes sont celles de l'altaïque et de l'indo-européen (taux d'erreur respectif de 6% et 8%) ; 2) les réflexes les moins exactement représentés dans le dictionnaire sont ceux des affriquées et des occlusives sonores ; il existe également des « zones non fiables » pour chacune des branches : ainsi, on peut considérer que dans le lexique de la protolangue altaïque il « manque » environ 30% des réflexes des consonnes nasales nostratiques.

5. Enfin, et c'est une des applications les plus connues de la statistique dans le domaine du comparatisme, il est possible d'établir la chronologie relative de la séparation de langues apparentées sur la base d'un traitement statistique du matériau lexical. Plus efficace que la méthode lexicostatistique de Swadesh, fondée sur les ressemblances entre unités lexicales, l'approche tout autre de S. Starostin s'appuie justement sur le lexique apparenté en utilisant une technique de calcul très différente (Burlak, S. & S. Starostin, 2005). Il existe d'autres méthodes qui donnent des résultats prometteurs. A mon sens, l'une d'elles est celle que proposent P. Boyeldieu et P. Nougayrol pour la classification génétique des langues sara-bongo-baguirmiennes. Cette méthode suppose des calculs distincts pour les différentes

couches du lexique (Boyeldieu, P. & P. Nougayrol, 2005).

Cet inventaire des sphères d'application des méthodes statistiques englobe de fait tous les aspects les plus importants de la linguistique comparative, ce qui permet de considérer les « études comparées quantitatives » comme un champ des études comparatives.

En conclusion, arrêtons-nous encore sur un domaine d'application de la statistique élémentaire dans les études comparatives, et notamment dans le traitement des célèbres et nombreuses exceptions aux changements phonétiques. En travaillant sur la reconstruction atlantique avec G. Segerer (Pozdniakov, K. & G. Segerer, à paraître), nous avons remarqué un fait étrange : dans ces langues, où la structure des racines lexicales est majoritairement de forme CVC, il existe très peu de racines combinant des consonnes ayant le même point d'articulation. Nous avons procédé à une évaluation quantitative de cette observation empirique. On peut aisément calculer pour chaque langue une fréquence théorique (une « norme ») pour la combinatoire des consonnes : la consonne initiale a une fréquence donnée dans le dictionnaire, la finale également, il est donc facile de compter dans chaque cas le nombre de combinaisons de consonnes dans le lexique s'il n'y a pas de corrélation entre ces consonnes. En calculant la norme et les cas de non-conformité à la norme, nous avons obtenu un résultat inattendu : il apparaît extrêmement rare que deux consonnes de même point d'articulation se trouvent combinées de part et d'autre d'une voyelle. Dans une moindre mesure, on observe également des restrictions combinatoires entre les labiales et les vélaires d'une part, entre les dentales et les palatales d'autre part. Ainsi, sur la base de ces restrictions combinatoires, peut-on identifier deux superclasses de consonnes : les «centrales», regroupant les palatales et les dentales, et les «périphériques», regroupant les labiales et les vélaires. Notre plus grande surprise a été le fait que cette restriction existe non seulement dans les langues atlantiques, mais dans la plus grande partie des langues du monde. Il est important de souligner que ces restrictions statistiquement significatives à la combinatoire des consonnes de même ordre se rencontrent malgré les tendances contraires bien connues (assimilation, réduplication). Prenons un exemple en russe moderne. En tant que locuteur natif de russe, je n'ai pas conscience d'une quelconque restriction à combiner les consonnes **b** et **p** lorsqu'ils sont séparées par une voyelle. La statistique, elle, démontre pourtant l'existence de cette restriction, et c'est la statistique, et non l'intuition du locuteur, qui se révèle juste. La fréquence des phonèmes /b/ et /p/ est très élevée en russe. Cela signifie que le dictionnaire devrait présenter des centaines de mots ayant à l'initiale une structure **bVp-**. Voila pourtant leur liste exhaustive : **baptist** 'baptiste', **biplan** 'biplan', **bipol'arnyj** 'bipolaire'. Et c'est tout ! Il est clair que ces trois mots sont empruntés, et par ailleurs, dans deux d'entre eux /b/ et /p/ sont séparés par une frontière morphologique. On devrait avoir en russe plusieurs centaines de mots (compte tenu des fréquences des phonèmes concernés) commençant par **pVb-**. En effet, il existe un nombre important de mots commençant par le préfixe **po-**, comme **po-bezhat'** 'commencer à courir', **po-bit'** 'battre', etc.. Hormis ces mots, je n'ai pu trouver qu'une seule racine présentant cette structure, dans les emprunts **publika** 'publique', **pablisiti** 'publicité' et **pab** 'pub'. Est-il possible de ne pas tenir compte, dans la grammaire historique du russe, d'une restriction aussi indubitable ? De quels processus diachroniques peut-elle résulter ? Visiblement, il faut la rapporter au type de changement qui a conduit à la formation de l'espagnol **arbor** 'arbre', c'est-à-dire à un changement marginal (du point de vue de la théorie comparative) connu sous le nom d'assimilation/dissimilation distante des liquides et des sibilantes. Mais des changements de ce type ne correspondent pas à un type de changement phonétique régulier. On ne peut pas les considérer non plus comme des changements par analogie.

4. Discussion

Pour la méthode comparative, tous les changements diachroniques sont égaux, mais en pratique, certains changements sont plus égaux que d'autres. La linguistique comparative introduit de fait une hiérarchie stricte entre les différents types de changements diachroniques :

- On considère en premier lieu les changements phonétiques, dont la réalité est déduite des correspondances phonétiques régulières observées.
- En l'absence de correspondance régulière, c'est l'hypothèse de l'emprunt qui est le plus souvent évoquée pour expliquer une ressemblance. Comme l'a justement dit Helmslev, les emprunts « ...sont comme des minorités pour lesquelles il existerait une législation spéciale. (...) l'explication par l'emprunt est une "réserve" dont le linguiste dispose toujours quand toutes les autres explications font défaut » (Hjelmslev, 1991, p. 87, 89) .
- Si l'hypothèse de l'emprunt ne peut pas être retenue, on pourra expliquer une des formes observées comme résultant d'un changement par analogie.
- Si cette solution échoue, il reste une issue de secours : considérer la forme comme iconique (par exemple, un idéophone).
- Enfin, au bas de la hiérarchie, on trouve les explications de type ethnolinguistique : lexique tabou, formes d'évitement, langues secrètes...

La pratique de la comparaison des langues d'Afrique occidentale confirme la remarque très importante de Saussure, qui disait que « ... l'histoire de chaque langue permet de découvrir un fourmillement de faits analogiques accumulés les uns sur les autres, et, pris en bloc, ces continuels remaniements jouent dans l'évalution de la langue un rôle considérable, plus considérable même que celui des changements de sons. (...) La langue est une robe couverte de rapiéçages faits avec sa propre étoffe ». (Saussure, 1972, p. 235). Dans la littérature comparatiste, les changements par analogie occupent sans aucun doute une position inférieure aux changements phonétiques systématiques. A ce jour, les données recueillies sont suffisamment nombreuses pour « réhabiliter » les changements par analogie.

Sur ce point, il serait souhaitable de considérer que les changements phonétiques systématiques et les changements par analogie constituent deux aspects opposés du changement diachronique. Cette dichotomie oppose en fait des changements **syntagmatiques** (changement des allophones dans un contexte phonétique donné, y compris dissimilation distante des consonnes de même point d'articulation) et des changements **paradigmatiques** (changement par analogie conduisant à la formation de « signes paradigmatiques »). Si on y ajoute les changements « marginaux » (lexique expressif, etc.), nous obtenons une typologie des changements diachroniques phonétiques qui rappelle de manière étonnante (et probablement non fortuite) la structure du signifié dans laquelle on distingue au moins trois facettes : syntagmatique, paradigmatique (*valeur* chez Saussure), et **modal**.

Pour terminer, j'émettrai ici une hypothèse prudente : Ne peut-on pas prévoir que l'élaboration d'une typologie des changements diachroniques montrera que les divers groupes de langues aux différentes étapes de leur évolution sont plus « enclins » que d'autres à certains types de changement ? Par exemple, il y a un facteur objectif qui limite les changements phonétiques « classiques » dans la diachronie des langues à système d'alternances consonantiques développé, comme les langues atlantiques : si la langue obéit à des règles strictes d'utilisation des signes distinctifs des phonèmes dans la grammaire, le prix de tout changement « mécanique » est trop élevé puisque tout changement de ce type romprait l'équilibre fragile des relations entre système phonétique et fonctions grammaticales.

Références

Bogoroditskij, V. *Краткий очерк сравнительной грамматики ариоевропейских языков* (2ème éd.). Kazan', 1917.

Boyeldieu, P. & P.Nougayrol. L'expansion sara-bongo-baguirmienne, communication au XIIème colloque du Réseau Méga-Tchad « Migrations et mobilité dans le bassin du lac Tchad », Maroua, 2005.

Burlak, S. & S. Starostin. *Сравнительно-историческое языкознание.* Moscou, 2005 : Academia.

Carlton, E. M. & Sharon R.R. Enquête sociolinguistique sur les langues Diolas de Basse Casamance. Cahiers de Recherche Linguistique, 2. Dakar, 1993 : Société Internationale de Linguistique.

Doneux, J.L. Hypothèses pour la comparative des langues atlantiques // *Africana Linguistica,* 6, 1975.

Ferry, M.-P. & K. Pozdniakov. Dialectique du régulier / irrégulier dans la reconstruction des classes nominales. // *Leçons d'Afrique, Filiations, ruptures et reconstitution de langues, Un Hommage à Gabriel Manessy* (éd. R. Nicolaï). Paris, 2001: Peeters.

Hjelmslev, L. *Le langage.* Paris, 1991 : Gallimard.

Ivanov, V'ach. Vs. *Лингвистика третьего тысячелетия: Вопросы к будущему.* Moscou, 2004 : Jazyki slav'anskoj kultury.

Jakobson, R. Морфологические наблюдения над славянским склонением // *American contributions to the IV International congress of Slavicists.* 's-Gravenhage, 1958.

Koch, H. Reconstruction in Morphology // *The Comparative Method Reviewed* (éd. M. Durie, M. Ross). New York – Oxford, 1966 : Oxford University Press.

Koval, A. Именные категории в пулар-фульфульде // *Основы африканского языкознания* (éd. V. Vonogradov). Moscou, 1997 : Aspect Press.

Porkhomovsky, V. Проблемы генетической классификации в языках Африки. // *Теоретические основы классификации языков мира. Проблемы родства.* Moscou, 1982 : Nauka.

Pozdniakov, K. Развитие систем консонантных чередований в языках манде и в атлантических языках // *Африканское историческое языкознание (Проблемы реконструкции).* Moscou, 1987 : Nauka.

Pozdniakov, K. Perspectives of comparative studies on the Mande and West Atlantic : An approach to the quantitative comparative linguistics // *Mandenkan*, 22, Paris, 1991.

Pozdniakov, K. *Сравнительная грамматика атлантических языков* (éd. V. Dybo). Moscou, 1993 : Nauka.

Pozdniakov, K. Micromorphologie ou morphologie de paradigme? // *Bulletin de la Société de linguistique de Paris*, 2003, n° 1, vol. XCVIII.

Pozdniakov, K. & G. Segerer. Reconstruction des pronoms atlantiques et typologie des systèmes pronominaux. // *Systèmes de marques personnelles en Afrique* (éd. D. Ibriszimow et G. Segerer). Louvain, Paris, 2004: Peeters, Coll. Afrique et Langage n° 8.

Pozdniakov, K. & G. Segerer. Similar Place Avoidance: A Statistical Universal. *A paraître.*

Sapir, J. D. West Atlantic: An inventory of the languages, their noun class systems and consonant alternation // *Current Trends in Linguistics*, vol. 7 (éd. T.A. Sebeok). The Hague, 1971: Mouton.

de Saussure, F. *Cours de linguistique générale.* Paris, 1972 : Payot.

Résumé

Dans cet article, on aborde quelques problèmes résultant de l'application de la méthode comparative aux langues d'Afrique occidentale, et en particulier, aux langues atlantiques. Parmi les approches propres à la reconstruction des protolangues africaines, on s'attache tout particulièrement à montrer l'apport des méthodes quantitatives. On examine certains domaines marginaux de la linguistique comparatiste, qui n'ont pas encore reçu toute la place qu'ils méritaient dans la théorie, notamment le rôle des submorphes dans l'interprétation des changements par analogie, ainsi que le rôle des processus de dissimilation des consonnes de même point d'articulation dans une typologie des changements diachroniques, ce qui conduit à examiner certaines perspectives d'une telle typologie.

The present paper examines some specific methodological problems of comparative studies, among them the application of the standard comparative method to West African languages in general and Atlantic languages in particular, reconstruction of proto-languages in Africa and perspectives of "quantitative comparative studies". A special focus is made on certain marginal aspects of comparative historical linguistics, so far neglected by the standard theory, i.e. the role of submorphs in analogical changes, similar place avoidance within the general typological framework of diachronic changes and the potentials of this approach.

В статье анализируются некоторые специфические проблемы приложения классического сравнительно-исторического метода к языкам Западной Африки, и в частности, к атлантическим языкам. Обсуждаются перспективы «кантитативной компаративистики». Рассматриваются некоторые маргинальные области сравнительно-исторического языкознания, которые, как представляется, пока не заняли подобающего им места в теории: в частности, роль субморфов в интерпретации изменений по аналогии, место диссимилятивных процессов у согласных общего места образования в типологии диахронических изменений. В связи с этим обсуждаются и некоторые перспективы такой типологии.