



HAL
open science

Mesurer les musiques pour parler du passé : la comparaison des musiques du Gabon comme source d'informations historiques

Jérémy Gardent

► **To cite this version:**

Jérémy Gardent. Mesurer les musiques pour parler du passé : la comparaison des musiques du Gabon comme source d'informations historiques. Anthropologie sociale et ethnologie. Museum national d'histoire naturelle - MNHN PARIS, 2017. Français. NNT : 2017MNHN0003 . tel-01715159

HAL Id: tel-01715159

<https://theses.hal.science/tel-01715159>

Submitted on 22 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Muséum National d'Histoire Naturelle

École Doctorale Sciences de la Nature et de l'Homme – ED 227

Année 2017

N° attribué par la bibliothèque

THÈSE

Pour obtenir le grade de

DOCTEUR DU MUSÉUM NATIONAL D'HISTOIRE NATURELLE

Spécialité : Évolution culturelle

Présentée et soutenue publiquement par

Jérémy Gardent

le 30 juin 2017

Mesurer les musiques pour parler du passé La comparaison des musiques du Gabon comme source d'informations historiques

Volume I : Introduction

Sous la direction de : Évelyne HEYER et Sylvie LE BOMIN

JURY :

Dana Rappoport,	Directrice de Recherche, CNRS, Paris	Rapporteuse
Stephan Shennan,	<i>Professor</i> , University College, Londres	Rapporteur
Olivier Tourny,	Chargé de Recherche, CNRS, Paris	Examineur
Olivier Morin,	<i>Research group leader</i> , Max Planck Institute, Jena	Examineur
Évelyne Heyer,	Professeure, MNHN, Paris	Directrice de Thèse
Sylvie Le Bomin,	Maître de Conférence, MNHN, Paris	Co-directrice de Thèse



Travail sous licence Creative Commons
Attribution – Partage dans les Mêmes Conditions
4.0 International

Table des matières

Volume I : Introduction	7
Remerciements	7
Préambule	9
Introduction	13
1 L'inférence historique par comparaison des structures formelles	15
1.1 De quoi s'agit-il?	17
1.2 Historique disciplinaire	18
1.3 Description de méthodes d'inférence historique	24
1.4 Conclusion du chapitre	47
2 Les méthodes issues de l'inférence phylogénétique appliquées aux données culturelles	49
2.1 État de l'art disciplinaire	53
2.2 Considérations générales sur ces transpositions interdisciplinaires	75
2.3 Tableau comparatif de quelques études	81
2.4 Conclusion du chapitre	103
3 Systématique musicale	105
3.1 Outils théoriques et méthodologiques	107
3.2 Des outils au service de la comparaison interculturelle	114
3.3 Réflexions autour de la notion de <i>modèle</i>	118
3.4 Conclusion du chapitre	128
4 Musique, comparatisme et histoire	129
4.1 Quelle place pour la comparaison?	130
4.2 La comparaison dans la pratique de l'ethnomusicologie	132
4.3 La comparaison musicale pour l'inférence historique	139
4.4 Conclusion du chapitre	143
Conclusion de la première partie	145
Volume II : Résultats — Discussion	153
Résultats	153
5 Essai d'inférence phylogéographique	155
5.1 Contexte et objectifs	157
5.2 Méthodes	158
5.3 Résultats	165
5.4 Discussion	180
5.5 Perspectives	184

6	Un aperçu de la diversité des musiques du Gabon	187
6.1	Introduction	188
6.2	<i>Diversity and distribution of musics in Gabon</i>	191
6.3	Conclusion et perspectives	245
7	Quelles données, quelle(s) histoire(s) ?	247
7.1	Positionnement épistémologique	249
7.2	Croisements d'histoires	251
7.3	Les indicateurs de choix des marqueurs historiques	268
7.4	Conclusion du chapitre	274
	Conclusion de la deuxième partie	275
	Discussion	277
8	Pistes de recherches	279
8.1	Comprendre la transmission et la transformation des musiques	280
8.2	Approfondir l'étude des formules rythmiques	287
9	Réflexions sur la pratique de la recherche	295
9.1	La relation aux données	297
9.2	Le rôle d'analyste	301
9.3	Limites de l'interdisciplinarité	304
9.4	La transposition de méthodes	306
	Conclusion générale	315
	Volume III : Glossaire — Annexes	327
	Glossaire argumenté	327
	Annexes	370
A	Conventions	373
A.1	Citations et traductions	373
A.2	Sources et références bibliographiques	374
B	Suppléments	377
B.1	Chapitre 3 : Citations autour de la notion de <i>modèle</i>	378
B.2	Chapitre 5 : Inférence phylogéographique	387
B.3	Test de Mantel	421
C	Bibliographie	425
C.1	Publications consultées	425
C.2	Autres publications citées	445
D	Listes des figures et des tableaux	457
D.1	Liste des figures	458
D.2	Liste des tableaux	460
E	Table exhaustive des matières	461

Volume I
Introduction

Remerciements

Ces quatre années de doctorat s'achèvent par la présente thèse, exercice individuel qui laisse mal transparaître tout ce qu'elle doit au travail collectif. Les pages qui suivent n'ont pu être écrites que grâce au concours de nombreuses personnes et institutions que je remercie sincèrement.

Merci tout d'abord à mes directrices de thèse, qui ont porté ce projet du début à la fin. Évelyne Heyer a cru en ce sujet et a su le défendre chaque fois qu'il l'a fallu. Elle a aussi su me donner les conseils, avisés et pragmatiques, pour me sortir des impasses scientifiques dans lesquelles je m'embourbe facilement. Sylvie a supporté pendant toutes ces années mes doutes et questionnements sur le sujet qui m'occupait, sachant toujours les écouter et les orienter vers des réalisations concrètes qui me redonnaient confiance. Elle s'est aussi attachée, chaque fois qu'il le fallait et avec bienveillance, à me faire prendre du recul sur mes certitudes tout en m'aidant à acquérir les bases théoriques et pratiques nécessaires pour traiter mon sujet, et a pris le temps de me guider et de m'accompagner dans la découverte du terrain.

Cette thèse doit aussi beaucoup au laboratoire *Éco-anthropologie et ethnobiologie* (UMR 7206, CNRS-MNHN), qui en a fourni le cadre matériel et scientifique. Florence Loiseau et Taouès Lahrem ont fait preuve d'une grande disponibilité et je les en remercie. L'équipe *Systématique et catégorisation culturelle*, en particulier, m'a apporté à la fois une formation rigoureuse, un espace de réflexion et de partage sur nos recherches et les outils et moyens pour les réaliser. Merci donc à tous les membres de l'équipe, notamment à Susanne Fűrniß dont la rigueur d'une part et la bienveillance d'autre part venaient, tour à tour et au moment opportun, renforcer mon travail ou mon moral. Je remercie aussi chaleureusement Pierre Darlu, qui a toujours accepté de consacrer sa rigueur et son exigence scientifique à me conseiller et relire mes travaux. Guillaume Lecointre et Michel Raymond, eux aussi membres de mon comité de thèse, apportèrent leur œil extérieur (et perspicace!) m'aidant à prendre des décisions à certains moments clés et je leur en suis reconnaissant. Dominique Grimaud-Hervé a su m'aider à concrétiser et terminer un travail qui risquait d'être sans fin, merci! Je tiens aussi à remercier les jeunes chercheur-e-s animant le séminaire *Atelier terrain—théorie en anthropologie*, ainsi que Frank Alvarez-Pereyre qui anime le séminaire *Interdisciplinarité*, car les échanges suscités en ces occasions m'ont été précieux et m'ont permis de prendre du recul sur mes recherches. Un très grand merci, enfin, à toutes les personnes qui ont passé du temps à relire mon travail ou m'ont aidé tout au long de ces années.

Préambule

« Et pourquoi les musiques du Gabon ? » Une fois passé le premier étonnement, parfois émerveillé parfois dubitatif qu'il existe des personnes dont le métier puisse être d'étudier la musique gabonaise, voilà la question qui m'est bien souvent posée juste après. Oui pourquoi les musiques du Gabon ? Le « choix du terrain » et plus généralement celui du sujet de recherche est bien souvent celui d'un parcours dans lequel le hasard a sans aucun doute sa place. Ignorant tout du Gabon et de ses musiques comment en suis-je venu à décider d'y consacrer trois années de ma vie¹ ?

C'est d'abord l'histoire d'un parcours universitaire. Biologiste de formation m'étant spécialisé de par mes choix dans la biologie de l'évolution, j'avais un attrait tout particulier pour les concepts et théories sous-jacents à cette discipline, mais aussi pour ses méthodes et les questions épistémologiques qu'elle soulève². Découvrant le plaisir d'apprendre des langues tout d'abord puis spéculant sur certains rapprochements théoriques entre l'évolution des langues et des espèces j'ai pris connaissance de l'utilisation de méthodes venant de la biologie en linguistique historique. C'est par cette fenêtre très étroite que j'ai découvert l'étude de l'évolution culturelle³. J'aurais voulu alors pouvoir étudier chaque pan de la culture, mais je me tournais vers l'étude de la musique, poussé sûrement par mes plaisirs de musicien amateur et sans doute aussi par l'intuition qu'un certain formalisme de la musique, semblable à celui du langage, devrait faciliter un tel travail.

Mais cette thèse doit surtout énormément à une rencontre, celle de Sylvie Le Bomin qui deviendra ma directrice. Ethnomusicologue étudiant depuis plus d'une dizaine d'années les musiques du Gabon, Sylvie s'attachait avec Évelyne Heyer et Guillaume Lecointre depuis les années 2010 à utiliser une méthode venant de la biologie (la méthode cladistique) pour analyser ses données musicales. Compte tenu de la proximité des recherches de Sylvie avec mes préoccupations,

1. ↑ J'ignorais alors qu'il m'en faudrait quatre pour mener à bien ce projet. Mais peut-être l'avais-je senti fort de l'expérience de mes nombreuses et nombreux collègues étant passé par là avant moi ?

2. ↑ Sans aucun doute déformé par une tendance exagérée à la théorisation et à l'abstraction que je m'imagine en toute méconnaissance de cause bien française. Cette addiction sur-théorique dont j'essaie difficilement de me défaire ne pourra pas passer inaperçue à la lecture de cette thèse et je m'excuse d'avance pour les lourdeurs quelle génère, les faits simples inutilement compliqués par une théorisation malvenue et, surtout, par les extrapolations hasardeuses et peut-être dommageables qui en sont issues.

3. ↑ Traditionnellement étudiée par l'anthropologie. Ce n'est que bien plus tard que je me rendrais compte de la longue tradition d'étude historique de la culture en général et de la musique en particulier, dont aujourd'hui encore j'ignore presque tout.

la rencontre pouvait sembler inévitable. Elle se fit pourtant difficilement puisqu'entre un stage dans un laboratoire d'ethnomusicologie et deux années de formation en musicologie classique mises à profit pour me renseigner sur ce sujet qui m'attirait, il m'a fallu presque trois ans pour prendre connaissance des travaux de Sylvie. La matrice de données qu'elle avait constituée me procurait le moyen de commencer rapidement mes recherches.

Pourtant les difficultés ne tardèrent pas à arriver. Si mes connaissances des méthodes d'analyse phylogénétique en biologie nécessitaient certainement d'être consolidées, mon bagage ethnomusicologique était quant à lui quasiment nul et devait être renforcé de toute urgence. De plus l'utilisation de méthodes venant de la biologie dans une discipline aussi différente que l'ethnomusicologie ne pouvait manquer de poser un certain nombre de questions théoriques et pratiques. Des disciplines telles que la linguistique historique ou l'anthropologie de la culture matérielle ayant une pratique plus ancienne de telles transpositions, ces expériences pouvaient être mises à profit dans mes tâtonnements et méritaient à ce titre d'être explorées⁴. Enfin, je ne pouvais ignorer la longue tradition ethnomusicologique qui s'était souvent posée la question de l'histoire invoquant de nombreuses théories et méthodes différentes⁵.

Tout ceci m'a conduit à explorer des domaines très variés que j'ai jugé intéressant de rapporter dans la première partie de cette thèse⁶. Si les spécialistes de l'un ou l'autre des domaines n'y apprendront sans doute que peu de choses, elle a pour objectif de rendre compréhensible la problématique interdisciplinaire qui m'occupe et donner aux scientifiques de différentes traditions les moyens de comprendre et évaluer mon travail.

Ce bagage bibliographique et méthodologique, s'il m'a sans doute guidé et donné un peu plus de confiance, n'a certainement pas suffi à affronter sans encombre l'épreuve de la pratique. Après de nombreux tâtonnements, ma première analyse aboutie me confronta aux limites théoriques et empiriques qui rendaient impossibles d'évaluer la pertinence de mes résultats. Cette constatation a donc été l'occasion d'une réorientation importante de ma thèse : il apparaissait indispensable, avant toute tentative d'inférence historique, de mieux comprendre et décrire les données à ma disposition d'une part⁷ et d'affiner les fondements théoriques et méthodologiques de l'inférence historique à partir de données musicales d'autre part⁸. Cette première analyse exploratoire, la description de la diversité des musiques du Gabon et les réflexions théoriques qui en sont issues constituent les principaux résultats de cette thèse et seront présentés dans la deuxième partie⁹.

Ce cheminement sinueux s'est accompagné de nombreux questionnements et le cadre offert par mon laboratoire m'a permis de prendre du recul sur ma propre pratique de recherche. Sur ce qu'était mon terrain, objet si cher et personnel de l'ethnomusicologie dont je me sentais dépourvu et, par là même, sur les problèmes que posait le fait de me décharger de la collecte

4. ↑ Pages 49–103.

5. ↑ Voir chapitre 4 pages 129 à 143.

6. ↑ Pages 15–146.

7. ↑ Voir chapitre 6 pages 187 à 245.

8. ↑ Voir chapitre 7 pages 247 à 274.

9. ↑ Pages 155–276.

des données pour me concentrer sur leur analyse. Ces réflexions liées à la difficulté de maîtriser tous les aspects de mon sujet de recherche m'ont conduit à réfléchir sur les pratiques et limites de l'interdisciplinarité, en particulier en ce qui concerne la réutilisation d'une méthode dans une autre discipline. Ces discussions et introspections constituent la dernière partie de ma thèse¹⁰.

Le contenu principal de ma thèse présenté ci-dessus et constituant les deux premiers volumes est complété par un dernier volume contenant le glossaire et les annexes. Le glossaire détaillé est d'abord conçu comme un outil pour clarifier le sens que j'ai choisi d'utiliser pour chaque terme et astreindre ainsi mon vocabulaire à une plus grande rigueur. Mais il s'est rapidement avéré être aussi un outil essentiel à la bonne lecture de cette thèse. En effet, la perspective interdisciplinaire de ce texte fait qu'il est amené à être lu par des personnes aux divers bagages disciplinaires, une même précision pouvant être à la fois essentielle pour l'une et superflue voire pesante pour l'autre. Afin de faciliter autant que possible la lecture tout en permettant que le texte soit le plus accessible, j'ai choisi de proposer un glossaire détaillé et argumenté expliquant le sens des termes les plus importants de mon travail. Je ne peux que vivement encourager la consultation du glossaire dès lors qu'une expression est inconnue ou que son usage paraît surprenant. Les mots expliqués dans le glossaire sont signalés en violet dans le corps du texte, ce qui ne devrait pas trop nuire à la lecture.

Le annexes regroupent les informations qui peuvent être utiles si l'ont veut approfondir certains aspects. On y trouvera la bibliographie détaillée, que j'ai choisi de séparer entre les travaux que j'ai consultés directement d'une part et ceux dont j'ai pris connaissance par l'intermédiaire d'une autre source d'autre part¹¹.

10. ↑ Pages 279–313.

11. ↑ L'organisation de la bibliographie est expliquée plus en détail dans l'annexe A.2 page 374.

Première partie

Introduction

Chapitre 1

L'inférence historique par comparaison des structures formelles

1.1	De quoi s'agit-il ?	17
1.2	Historique disciplinaire	18
1.2.1	En critique textuelle	18
	Naissance de la stématisation	18
	Formalisation des méthodes	19
	Raffinements méthodologiques	19
1.2.2	En biologie	20
	La conception évolutionniste et la filiation entre les espèces	20
	Naissance de la systématique biologique	20
	Premières méthodes d'inférence historique	20
	Le développement des méthodes automatisées d'inférence	21
1.2.3	En linguistique	22
	L'évolution et la filiation des langues	22
	Premières propositions comparatistes	22
	Le développement de méthodes rigoureuses	23
	Le développement marginal d'autres méthodes	24
1.3	Description de méthodes d'inférence historique	24
1.3.1	Objectifs	24
	Un travail préparatoire à l'analyse	24
	Un outil pédagogique au service de l'interdisciplinarité	24
	Une base de comparaison	25
1.3.2	Méthodes et limites	25
	Diversité des méthodes	25
	Choix des méthodes présentées	26
1.3.3	La méthode comparative (linguistique historique)	26
	Principe général	26
	Identification des cognats	27
	Établissement de correspondances phonologiques	27
	Reconstruction de la proto-phonologie	27

	Reconstruction des proto-morphèmes	29
1.3.4	Les méthodes stématiques	29
	Principe général	29
	Classification des lieux variants	29
	Construction de la chaîne reliant les manuscrits	31
	Nœuds intermédiaires et orientation de la chaîne	31
1.3.5	La compatibilité des caractères	31
	1.3.5.1 Principe	32
	1.3.5.2 Méthodes de recherche	32
	Cas de caractères binaires	32
	Cas des caractères à états multiples	34
1.3.6	Les méthodes cladistiques	34
	Principe général	34
	Choix des caractères	35
	Identification des états de caractères	35
	Polarisation des états de caractères	35
	Recherche de l'arbre le plus parcimonieux	36
1.3.7	Les méthodes probabilistes	36
	1.3.7.1 Les méthodes de maximum de vraisemblance	39
	1.3.7.2 L'inférence bayésienne	39
	Principe général	39
	Analyse bayésienne empirique	39
	Analyse bayésienne hiérarchique	40
1.3.8	La biogéographie historique des taxons	40
	1.3.8.1 Optimisation de Camin–Sokal	40
	1.3.8.2 Optimisation de Fitch	42
	1.3.8.3 Optimisation pondérée	42
	1.3.8.4 Analyse de dispersion–vicariance (DIVA)	42
1.4	Conclusion du chapitre	47

1.1 De quoi s'agit-il ?

Les objets matériels, quels qu'ils soient, sont le produit d'une histoire unique. Cette histoire laisse une marque dans les caractéristiques observables de ces objets. À partir de la description et de l'analyse de ces traces historiques, les **sciences historiographiques** (qui ont pour objet l'étude du passé) cherchent à formuler des hypothèses rationnelles, plausibles et vérifiables concernant les événements et les objets passés. Ce processus, permettant de passer d'objets matériels observables actuellement à des hypothèses sur le passé, est l'**inférence historique**.

Dans certains cas, un objet observé seul peut donner des indices utiles à l'inférence historique¹. Par exemple, dans le domaine de la géologie, un ensemble de roches présentant toutes les caractéristiques de roches trouvées sous les océans² sont observables au mont Chenaillet, dans les Alpes françaises, à plus de 2000 mètres d'altitude. Cette observation permet de proposer une hypothèse : ces roches étaient autrefois présentes sous un océan.

Dans d'autres cas, l'objet observé seul ne permet de faire aucune inférence historique³ et seule sa comparaison avec d'autres objets de même nature le permet. Par exemple, le simple fait de savoir qu'on utilise en français le mot *chèvre* pour désigner l'espèce *Capra hircus*, ne permet pas de faire d'hypothèse sur le passé de la langue française. En revanche comparer les mots utilisés pour désigner cette même espèce dans différentes langues permet de faire une inférence historique⁴. Ainsi le mot français *chèvre* (prononcé /ʃevr/) est exprimé en italien par le mot *capra* (prononcé /kapra/) et en espagnol et en portugais par le même mot *cabra* (prononcé dans les deux cas /kabra/). On observe que les mots utilisés pour exprimer la même signification sont plus proches entre le portugais et l'espagnol qu'entre n'importe quelle autre paire de langues. On observe aussi que le mot italien ressemble plus au mot espagnol et portugais qu'au mot français. En prenant en compte d'autres types de connaissances⁵, on peut supposer que l'espagnol, le portugais et l'italien sont issus d'une même langue et que la séparation entre l'espagnol et le portugais est plus récente que celle entre l'italien et une des deux autres langues⁶. De la même manière, la seule description de l'espèce biologique *Pinus halepensis* (le pin blanc de Provence) ne permet pas d'émettre d'hypothèses sur le passé. La comparaison de cette espèce végétale avec les trois suivantes permet par contre de le faire : le brome dressé (*Bromus erectus*), le crocus d'automne (*Crocus nudiflorus*) et la bruyère callune (*Calluna vulgaris*)⁷ : on observe chez ces

1. ↑ Il nécessite bien sûr d'être mis en relation avec des connaissances et hypothèses extérieures. J'entends ici *seul* en opposition au cas où plusieurs objets du même ordre sont comparés les uns aux autres (toujours à la lumière d'autres connaissances extérieures), comme je l'explique un peu plus loin.

2. ↑ Ces roches présentant les caractéristiques de la lithosphère océanique et observables émergées sur un continent sont appelées des *ophiolites*.

3. ↑ Même lorsqu'il est mis en relation avec son contexte.

4. ↑ Cet aspect est détaillé dans la partie 1.3.3 page 26 et sur la figure 1.1.

5. ↑ Par exemple l'idée que les langues se transforment au cours du temps par modification progressive et peuvent se séparer en plusieurs langues différentes.

6. ↑ Cet exemple est bien sûr approximatif et incomplet et ne vise qu'à montrer comment la comparaison structurelle de plusieurs objets de même type permet d'apporter des informations utiles à l'inférence historique, qui ne seraient pas disponible sans le processus de comparaison. La partie 1.3 explique plus en détail différentes méthodes permettant d'effectuer de telles inférences historiques.

7. ↑ Cet aspect est détaillé dans la partie 1.3.6 page 34.

trois dernières espèces la présence de fleurs, absentes chez le pin. Cette observation permet de suggérer⁸ que l'ancêtre commun à ces trois espèces est plus récent que celui commun aux quatre espèces de l'exemple.

La comparaison systématique des caractéristiques d'objets semblables afin d'apporter des informations sur l'histoire de ces objets a été utilisée par plusieurs disciplines distinctes et de manière relativement indépendante. Dans ce chapitre, je présenterai rapidement l'historique de ces approches au sein de trois disciplines⁹ : la **critique textuelle** qui s'intéresse à l'histoire des textes écrits, la **biologie de l'évolution**, qui s'intéresse à l'histoire des êtres vivants et la **linguistique historique** qui s'intéresse à l'histoire des langues, avant de décrire un certain nombre de méthodes développées par chacune de ces trois disciplines¹⁰.

1.2 Historique disciplinaire

1.2.1 En critique textuelle

Naissance de la stématique. L'idée d'une filiation entre les différents manuscrits d'un même texte remonte probablement à l'antiquité et sans aucun doute à la renaissance puisqu'elle est évoquée par Érasme de Rotterdam en 1508¹¹. Dès lors, le développement de la critique textuelle est stimulé par le souci de proposer des éditions des textes anciens les plus proches possibles d'un supposé texte originel. Ainsi Scaliger propose au XVI^{ème} siècle une méthode rationnelle pour établir cette filiation¹². La première représentation sous forme d'un arbre reliant différents manuscrits¹³ est proposée en 1827¹⁴ par Collin et Schlyter [A95]¹⁸, à partir de différentes copies de textes de Loi du *Västergötland*¹⁹.

Lachmann [A231]²⁰, durant la première moitié du XIX^{ème} siècle, met en avant la nécessaire rigueur méthodologique et cherche à être le plus objectif possible. Il présente son approche au

8. ↑ Toujours en relation avec des connaissances extérieures, par exemple concernant les mécanismes biologiques d'évolution et de spéciation.

9. ↑ Voir partie 1.2 page 18.

10. ↑ Voir partie 1.3 page 24.

11. ↑ Don Cameron [A125, p. 231], citée par Bordalejo [49, pp. 39–40].

12. ↑ Timpanaro [A416], citée par Platnick et Don Cameron [339, p. 381].

13. ↑ Cette représentation sera plus tard appelée un *stemma*.

14. ↑ Soit dix ans avant l'ébauche d'arbre évolutif griffonné par Darwin [A109]¹⁵ en 1837 et plus de 30 avant le schéma publié dans l'origine des espèces. Concernant la linguistique, Schleicher [A375]¹⁶ publie en 1963 un arbre de l'évolution des langues indo-européennes, mais la première représentation *arborescente* de relations historiques entre des langues aurait peut-être été publiée autour de 1800 [A29]¹⁷. L'apparition simultanée de cette représentation en biologie et en critique textuelle correspond probablement à un développement parallèle (sans doute dû en partie à un contexte scientifique et culturel commun) plutôt qu'à une influence directe [353, pp. 93–94]. Un article de O'Hara [325] analyse plus en détail l'histoire de la figure d'arbre comme représentation de l'histoire en systématique et en critique textuelle.

15. Citée par Atkinson et Gray [27, p. 516] et Robins [353, pp. 93–94]. L'année du carnet dans lequel a été dessiné cet arbre diffère selon ces deux articles : 1836 pour Robins [353] et 1837 pour Atkinson et Gray [27].

16. Citée par O'Hara [325, pp. 84–85].

17. Citée par O'Hara [325, p. 84].

18. ↑ Citée par O'Hara [325, p. 85].

19. ↑ Le *Västergötland* (ou Gothie occidentale) est une province historique du sud-ouest de la Suède.

20. ↑ Citée par Timpanaro [A415, pp. 85–88], elle-même citée par Robins [353, p. 91].

cours de ses différents travaux, mais ne formalise pas vraiment une méthode générale de critique textuelle [A416, pp. 85–88]²¹ [366, pp. 18–19]. Ces premières propositions seront formalisées et modifiées au cours de la fin du XIX^{ème} siècle et jusqu’au début du XX^{ème} siècle, en particulier par Quentin [A343]²². La méthode **stemmatique** naissante propose ainsi de séparer la reconstruction d’un archétype hypothétique (*recensio*, étape se voulant objective) de la proposition d’un texte corrigé (*emendatio*, correspondant à la tâche de l’éditeur). L’étape de *recensio* est fondée sur le regroupement des manuscrits en familles à partir de l’identification et de la comparaison des « erreurs » contenues dans chacun d’eux [49, p. 40].

Formalisation des méthodes. Le début du XX^{ème} siècle est cependant marqué par un recul de la **stemmatique**²³. Tout d’abord en raison des critiques faites à la méthode, notamment par Bédier [A39]²⁴ qui remet en cause la prétendue impartialité du jugement éditorial, ainsi que la notion d’auteur unique qui n’a pas de sens dans certaines traditions littéraires [A345]²⁴. L’intérêt des universitaires se portera par ailleurs vers d’autres approches, notamment synchroniques [353, p. 100]. Ceci vient, entre autres, de ce que les méthodes statistiques, en plein développement et suscitant alors un fort engouement, sont difficiles à appliquer à ces études diachroniques [353, p. 101]. La critique textuelle se tournera alors vers des approches subjectives comme celle du *meilleur texte* proposée par Bédier [A39]²⁴.

Maas [A272]²⁵ formalise en 1958 une procédure rigoureuse d’inférence historique inspirée des analyses de Lachmann qui sera largement utilisée en critique textuelle par la suite²⁶. Cette méthode est fondée sur l’identification d’« erreurs » dans les différents manuscrits puis sur la comparaison de ces erreurs entre les différents manuscrits. Elle permet ainsi de retrouver les manuscrits apparentés, d’établir les filiations directes et d’identifier d’hypothétiques manuscrits non retrouvés. L’objectif final consiste à reconstruire le contenu de ces manuscrits hypothétiques, jusqu’à parvenir au texte du supposé manuscrit originel, que Maas nomme l’*archétype*.

Raffinements méthodologiques. En 1974, Dearing [A113]²⁷ propose une amélioration de la méthode formulée par Maas. En s’inspirant des travaux de Greg [A173]²⁷ concernant la généalogie des textes, il propose une typologie des positions variables du texte, en fonction du nombre de variantes trouvées à cette position dans tous les textes étudiés et du nombre de textes où l’on trouve chaque variante. Toutes les positions n’apportant pas la même information, il préconise de ne choisir que celles qui sont appropriées pour retrouver les relations de filiation entre manuscrits [366, pp. 30–35]. De plus Dearing propose de construire d’abord une *chaîne*

21. ↑ Citée par Robins [353, pp. 90–91].

22. ↑ Citée par Bordalejo [49, pp. 41–43].

23. ↑ Voir note 26 page 19.

24. ↑ Citées par Robins [353, pp. 100–101] et Bordalejo [49, pp. 43–45].

25. ↑ Citée par Platnick et Don Cameron [339, p. 381].

26. ↑ La méthode formalisée par Maas est fréquemment appelée *méthode de Lachmann*, bien qu’elle diffère substantiellement de la procédure utilisée par Lachmann [273, p. 229] [49, p. 41]. J’utilise le terme de *méthodes stemmatiques* pour désigner les méthodes rigoureuses d’inférence de la filiation de textes issus des travaux de Lachmann puis de Maas, la discipline correspondante étant la *stemmatologie* (voir le glossaire pour plus de détails).

27. ↑ Citées par Salemans [366, p. 19].

représentant les relations de proximité entre manuscrits, puis de chercher l'emplacement le plus vraisemblable de l'archétype au sein de la chaîne, permettant d'obtenir un *stemma*^{28,29} [366, pp. 30–35]. Cette méthode sera présentée dans la partie 1.3.4 telle qu'elle a été décrite par Salemans [366, pp. 30–35].

1.2.2 En biologie

La conception évolutionniste et la filiation entre les espèces. C'est aussi à la fin du XVIII^{ème} siècle et au début du XIX^{ème} siècle que se développe la pensée *évolutionniste* en biologie. On passe ainsi progressivement d'une conception *fixiste* des espèces, selon laquelle les espèces ont toujours existé telles qu'elles sont observées, à une conception *transformiste* envisageant une modification des espèces au cours du temps.

Simultanément l'idée d'une filiation entre les êtres vivants actuels se développe³⁰. La publication en 1859 de *L'origine des espèces* par Darwin [A110], proposant une théorie globale de la manière dont les espèces évoluent et descendent les unes des autres, modifie en profondeur l'ensemble des disciplines biologiques.

Naissance de la systématique biologique. La *taxinomie*, branche de la biologie se consacrant à la classification du vivant, n'est pas épargnée par ce changement³¹. Alors que les précédentes classifications étaient souvent basées sur des critères arbitraires, l'idée de filiation entre toutes les espèces aura pour effet d'unifier les critères de classification. La taxinomie cherche alors à classer les êtres vivants selon leur ordre « naturel », c'est-à-dire selon leur histoire évolutive : on regroupera les espèces ayant divergé le plus récemment. Ce tournant marque le début de la *systématique biologique*, qui cherche à créer une classification du vivant reflétant son histoire évolutive. Il ne s'agit donc plus seulement de classer les espèces, mais avant tout de retrouver les relations d'apparementement entre ces espèces.

Cependant la systématique biologique, à l'instar de la *stemmatique*, se développe peu au cours de la première moitié du XX^{ème} siècle et retourne vers des méthodes plus subjectives [353, pp. 100–102]. Les raisons sont d'ailleurs les mêmes : les approches synchroniques se développent d'avantage, les méthodes et les notions utilisées sont remises en question et des méthodes objectives sont difficiles à mettre en place [353, pp. 100–102].

Premières méthodes d'inférence historique. Les *phylogénies* proposées par Ernst Hæckel à la fin du XIX^{ème} siècle seront très critiquées, notamment parce qu'elles ne sont pas basées

28. ↑ Contrairement aux méthodes antérieures qui construisaient directement un *stemma*, incluant la position de l'archétype, sans passer par une chaîne intermédiaire.

29. ↑ Cette étape de positionnement *a posteriori* de l'archétype au sein de la chaîne évoque l'enracinement *a posteriori* d'un arbre non raciné utilisé dans certaines méthodes d'inférence phylogénétique.

30. ↑ Cependant cette idée de filiation n'est alors pas complètement associée à celle de transformisme [250, pp. 175–177].

31. ↑ de Queiroz [342] défend cependant l'idée que ce changement de paradigme n'a que peu impacté la taxinomie, qui selon lui ne changea finalement pas beaucoup ses méthodes et ses conceptions.

sur des critères objectifs³². Les critiques formulées par Adolf Naef insistent sur la nécessité d'un rapprochement entre la systématique et la biologie comparée afin d'arriver à une classification qui reflète l'histoire évolutive du vivant³². Ces travaux conduiront Hennig [A189] à proposer en 1950³³ une méthode rigoureuse d'inférence phylogénétique (la méthode **cladistique**) à partir de la comparaison des caractères morfo-anatomiques des espèces [441, pp. 260–262]. La recherche d'une approche plus objective³⁵ d'une part et les problèmes posés par la méthode cladistique pour classer les espèces unicellulaires d'autre part [401, p. 53] conduisent Michener et Sokal [A294]³⁶ et Sneath [A386]³⁶ à proposer indépendamment en 1957 une méthode (alors appelée **taxinomie numérique**) permettant de classer les espèces à partir d'une mesure globale de similarité entre celles-ci et non plus à partir de la comparaison de caractères individuels [A387]³⁷. Cette méthode **phénétique** progresse rapidement à la fin du XX^{ème} siècle avec le développement du séquençage de masse et il existe aujourd'hui une grande variété d'algorithmes phénétiques implémentés dans la plupart des logiciels d'inférence phylogénétique.

Le développement des méthodes automatisées d'inférence. La fin du XX^{ème} siècle est ainsi marquée par la controverse entre les méthodes **cladistique** et **phénétique**. Ces deux types de méthode se développent aussi bien par l'apparition et l'amélioration d'algorithmes permettant l'automatisation des calculs que par des raffinements théoriques. Les méthodes de distance seront ainsi popularisées par les travaux de Fitch et Margoliash [A142]^{38,39} qui introduisent en 1967 le *critère des moindres carrés* donnant plus d'importance aux valeurs faibles dans la matrice de distance lors de la construction de l'arbre, puis par ceux de Jukes et Cantor [A218]⁴¹ qui proposent une mesure plus fine de la distance entre séquences moléculaires⁴². La méthode cladistique connaît une évolution importante avec les travaux de Camin et Sokal [A74]⁴⁴ dans les années 1960 : ils introduisent le critère de **parcimonie** utilisé pour choisir entre les différents arbres possibles et décrivent les premiers algorithmes de parcimonie. À la même époque, Wilson [A442]⁴⁵ puis Le Quesne [A246]⁴⁵ introduisent la **méthode de compatibilité**, basée elle aussi sur le critère de parcimonie mais s'éloignant de la méthode de Camin et Sokal en ce qu'elle refuse le recours à

32. ↑ Williams et Ebach [441, pp. 260–262].

33. ↑ Soit quelques années avant la formalisation de la méthode stématique par Maas [A272]³⁴.

34. Citée par Platnick et Don Cameron [339, p. 381].

35. ↑ Ce qui ne signifie pas pour autant que la méthode phénétique *est* plus objective que la méthode cladistique.

36. ↑ Citées par Stuessy [401, p. 53].

37. ↑ Citée par Stuessy [401, pp. 54–56].

38. ↑ Citée par Felsenstein [136, pp. 131–132].

39. ↑ Même si d'autres travaux, notamment ceux de Cavalli-Sforza et Edwards [A82]⁴⁰, ont grandement contribué au développement de ces méthodes.

40. Citée par Felsenstein [136, pp. 125–129] et par Darlu (communication personnelle).

41. ↑ Citée par Felsenstein [136, p. 134].

42. ↑ Cette dernière ne donne cependant pas de méthode pour construire une représentation à partir des distances calculées⁴³.

43. Pierre Darlu (communication personnelle).

44. ↑ Citée par Felsenstein [136, pp. 129–130].

45. ↑ Citées par Felsenstein [136, pp. 133–134].

toute hypothèse d'**homoplasie**⁴⁶ lors du choix du meilleur arbre⁴⁷.

Cette période voit aussi le développement des **méthodes probabilistes**, intégrant explicitement des modèles d'évolution au processus d'inférence. Parmi ces méthodes, issues des travaux des statisticiens du début du XX^{ème} siècle, on peut distinguer les méthodes de maximum de vraisemblance et les méthodes d'inférence bayésienne⁴⁸. Les méthodes de **maximum de vraisemblance**, introduites en **systématique biologique** par Edwards et Cavalli-Sforza [A130]⁴⁹ en 1964, formulent différentes hypothèses puis cherchent celle selon laquelle la probabilité d'obtenir les données effectivement observées est la plus grande. Elles seront appliquées pour la première fois en 1974 par Neyman [A316]⁴⁹ puis développées par Kashyap et Subas [A220]⁴⁹. Les méthodes d'**inférence bayésienne**, cherchant à trouver le modèle évolutif ayant la plus grande probabilité, en intégrant les informations issues des données observées, ont été utilisées pour la première fois en systématique biologique par Gomberg [A161]⁵⁰.

1.2.3 En linguistique

L'évolution et la filiation des langues. L'idée d'un processus d'**évolution**⁵¹ des langues et de processus de changement associés était déjà répandue dans l'Antiquité, au moins dans les études portant sur les langues de Grèce, de Chine, de Mésopotamie et d'Inde⁵². L'idée que les langues puissent être apparentées⁵³ s'est répandue à partir du V^{ème} siècle et les propositions faites par Augustin d'Hippone⁵⁴. Jusqu'au XVII^{ème} siècle, l'interprétation littérale de la Bible dans le monde chrétien impose l'hébreu comme la langue originelle. Dès lors les études linguistiques cherchent à relier toutes les langues à l'hébreu [A336]⁵⁵.

Premières propositions comparatistes. Au XVI^{ème} siècle, Scaliger distingue quatre groupes de langues européennes (qu'il considère indépendants) en comparant le mot utilisé dans chacune d'entre elles pour *Dieu*⁵⁶. Il pose les premières bases de la **méthode comparative**⁵⁷ et est ainsi peut-être le premier à utiliser concrètement un apparentement pour identifier des groupes de langues européennes. En 1647, van Boxhorn [A58–A60]⁵⁸ propose une première méthodologie d'inférence historique en linguistique. Il remarque que les emprunts ou les mots qui sont semblables mais qui n'ont pas d'étymologie commune peuvent induire des erreurs. Il en conclut la

46. ↑ Dans le vocabulaire de la cladistique, une **homoplasie** est une ressemblance qui n'est pas due à une ascendance commune. Voir l'entrée du glossaire correspondante pour plus de détail.

47. ↑ Voir partie 1.3.5 page 31.

48. ↑ Voir partie 1.3.7 page 36.

49. ↑ Citées par Felsenstein [136, p. 248].

50. ↑ Citée par Felsenstein [136, p. 291].

51. ↑ Voir le glossaire pour une bonne compréhension de ce que j'entends ici par *évolution*.

52. ↑ Atkinson et Gray [27, pp. 514–515].

53. ↑ Cette idée a longtemps été exclue par l'explication créationniste apportée par l'ancien Testament [27, p. 515] selon laquelle les langues trouveraient leur origine dans l'épisode de la Tour de Babel.

54. ↑ Percival [A336], citée par Atkinson et Gray [27, p. 515].

55. ↑ Citée par Atkinson et Gray [27, p. 515].

56. ↑ Pedersen [A335], citée par Atkinson et Gray [27, p. 515].

57. ↑ Atkinson et Gray [27, p. 515].

58. ↑ Citées par van Driem [427, pp. 285–286].

nécessité de différencier les ressemblances linguistiques héritées de celles développées de manière indépendantes⁵⁹. À partir de ses idées de gradualisme et d'**uniformitarisme**⁶⁰, Leibniz [A257]⁶¹ propose en 1710 une « généalogie » reliant les langues d'Europe, d'Asie et d'Égypte.

Jefferson [A211]⁶² suggère en 1782 que la comparaison linguistique peut aussi servir à retrouver les dates de divergence entre les différentes langues. Schlegel [A374]⁶³ introduit le terme de grammaire comparée et utilise pour la première fois en 1808 un diagramme en forme d'arbre pour représenter les relations généalogiques entre les langues. Schleicher publie en 1853 deux travaux incluant des arbres de langues⁶⁴ et propose en 1863 un arbre des langues indo-européennes dans un article faisant explicitement référence aux travaux de Charles Darwin [A375]⁶³.

Le développement de méthodes rigoureuses. Grimm [A178]⁶³ formule en 1822 la première loi de la phonétique, baptisée *loi de Grimm*. Grâce à la mise en évidence de correspondances systématiques de certaines consonnes de plusieurs langues germaniques, celle-ci propose une explication de la transformation de ces langues depuis l'indo-européen. Cette loi est la première des lois phonétiques développées à la fin du XIX^{ème} siècle et au début du XX^{ème} siècle par l'école néogrammatrice. Le manifeste de Osthoff et Brugmann [A329]⁶⁵, emblématique de cette école, considère que les changements phonétiques suivent des lois absolument régulières (principe d'**uniformitarisme**), qu'il s'agit alors de retrouver et de décrire en détail⁶⁶.

La méthode comparative est développée par l'école néogrammatrice [27, p. 518], en particulier par Osthoff et Brugmann [A329]⁶⁷ qui distinguent en 1878 deux types de ressemblances lexicales : les *réentions* qui sont héritées de la forme ancestrale et les *innovations* qui n'existaient pas dans la forme ancestrale⁶⁸. En 1951, Swadesh [A403]⁶⁹ préconise l'utilisation de listes prédéfinies de sens, supposés universels et donc peu sujets à l'emprunt, afin d'éviter d'introduire dans la comparaison des termes empruntés et ne reflétant pas l'histoire verticale des langues^{70,71}.

59. ↑ van Driem [427, p. 286].

60. ↑ Leibniz développe aussi ces concepts en biologie.

61. ↑ Citée par Pedersen [A335], elle-même citée par Atkinson et Gray [27, p. 516].

62. ↑ Citée par Atkinson et Gray [27, p. 516].

63. ↑ Citées par Atkinson et Gray [27, p. 517].

64. ↑ Koerner [A227], citée par Atkinson et Gray [27, p. 517].

65. ↑ Citée par Ben Hamed [41, p. 980].

66. ↑ Atkinson et Gray [27, p. 517] et Ben Hamed [41, p. 980].

67. ↑ Citée par Ben Hamed [41, p. 980] et Atkinson et Gray [27, pp. 517–518].

68. ↑ Ben Hamed [41, p. 980] et Atkinson et Gray [27, p. 518] donnent deux interprétations différentes des apports d'Osthoff et Brugmann. Pour la première, ils différencient « les similarités dues à l'ascendance commune de celles qui sont liées à d'autres processus » (invariants, hasard ou emprunts). Pour le second ils différencient, au sein des similarités dues à l'ascendance commune, celles qui étaient présentes chez l'ancêtre de toutes les langues étudiées, de celles acquises par l'ancêtre commun d'un sous-groupe de langues seulement. En utilisant le vocabulaire de la cladistique, dans le premier cas on différencie les **homologies** des **homoplasies**, alors que dans le second on différencie les **apomorphies** des **plésiomorphies**.

69. ↑ Citées par Ben Hamed [41, p. 981].

70. ↑ Les listes Swadesh de 200 et 100 mots ont été définies par Swadesh respectivement en 1951 [A403]⁶⁹ et en 1955 [A405]⁶⁹. Ces listes, bien que très critiquées, sont encore beaucoup utilisées aujourd'hui, parfois sous des formes légèrement modifiées.

71. ↑ Une autre méthode d'inférence historique, la comparaison de masse, est aussi proposée en 1955 par Greenberg [A170]⁶⁹, mais sera controversée et peu utilisée.

Le développement marginal d'autres méthodes. Swadesh [A404, A405]⁷² propose aussi d'utiliser ces listes pour calculer le pourcentage de mots communs entre deux langues, définissant ainsi une mesure de similarité calculable entre chaque paire de langues. Suivant cette méthode, appelée **lexicostatistique**, il construit ensuite une **matrice de distances** à partir de laquelle il infère les relations entre les langues grâce à des algorithmes de regroupement⁷³. Swadesh [A404]⁷² introduit aussi en 1952 la **glottochronologie**, une méthode qui suppose la vitesse de changement linguistique constante au cours du temps, afin de pouvoir dater les divergences entre les langues⁷⁴. La lexicostatistique et la glottochronologie seront beaucoup critiquées et les méthodes automatisées se développeront peu en linguistique dans la deuxième moitié du XX^{ème} siècle, excepté dans quelques travaux de Sankoff et Kruskal⁷⁵.

1.3 Description de méthodes d'inférence historique

1.3.1 Objectifs

Cette deuxième section sera consacrée à la présentation de certaines des méthodes d'inférence historique basées sur la comparaison formelle développées au cours de l'histoire de la critique textuelle, de la systématique biologique et de la linguistique historique. Ces descriptions ont plusieurs objectifs dans le cadre de ce travail : présenter le travail préparatoire à mes analyses, fournir un outil pédagogique pour faciliter l'appropriation de ces méthodes et fournir une base de connaissances permettant la comparaison de ces méthodes.

Un travail préparatoire à l'analyse. J'ai étudié la plupart de ces méthodes lors de ma première année de thèse, afin de forger des outils analytiques permettant l'analyse de mes données. En effet la question de l'analyse informatisée d'un grand nombre de données d'ordre musical est récente et celle de leur utilisation à des fins d'inférence historique encore plus. Il n'existe donc pas de tradition méthodologique propre à l'ethnomusicologie sur ces questions⁷⁶ et il m'était nécessaire de comprendre les réponses apportées par d'autres disciplines à des questionnements plus ou moins similaires.

Un outil pédagogique au service de l'interdisciplinarité. Cette partie se veut aussi fournir des explications à la fois relativement succinctes et accessibles à des personnes issues de disciplines variées pour comprendre les principes de base de ces différentes méthodes d'inférence historique. Aujourd'hui les méthodes d'inférence historique issues de la systématique biologique

72. ↑ Citées par Atkinson et Gray [27, p. 520].

73. ↑ Atkinson et Gray [27, p. 520].

74. ↑ Cette méthode rappelle celle de l'horloge moléculaire proposée en biologie par Zuckerkandl et Pauling [A448]⁷² en 1962, mais a été vraisemblablement développée de manière convergente dans les deux disciplines [27, p. 520].

75. ↑ Atkinson et Gray [27, p. 520].

76. ↑ Il existe néanmoins des travaux allant dans ce sens, sans que l'on puisse vraiment parler de « tradition » disciplinaire. Ces travaux seront présentés dans la partie 4.3 page 139.

tendent à se répandre dans de nombreuses disciplines différentes. La bonne compréhension de ces méthodes n'est cependant pas aisée et nécessite souvent de longues périodes d'apprentissage. Certaines d'entre elles font appel à des outils (notamment statistiques et mathématiques) qui sont utilisés et enseignés très tôt dans certaines traditions disciplinaires (par exemple en biologie) mais peu ou pas du tout utilisés dans d'autres (par exemple en ethnomusicologie).

Une base de comparaison. Enfin cette partie vise à décrire ces méthodes afin de mieux comprendre leurs similitudes et leurs différences. De nombreux travaux ont noté le développement parallèle de méthodes d'inférence historique en critique textuelle, en systématique biologique et en linguistique historique [27, 41, 325, 339]. Platnick et Don Cameron [339] ont même suggéré, sur la base de cette convergence historique, que ces méthodes seraient toutes des variantes d'une méthode universelle applicable à toutes sortes de données. Cependant l'accent a souvent été mis sur les points communs entre les méthodes, au détriment des différences qui sont loin d'être négligeables. De plus ces comparaisons ont souvent été faites depuis la perspective de la systématique biologique, négligeant une explication approfondie des méthodes issues de la critique textuelle et de la linguistique historique. Cette partie contribuera donc à mieux comprendre les méthodes et je l'espère à poser les bases d'une comparaison un peu plus fine entre les méthodes.

1.3.2 Méthodes et limites

Diversité des méthodes. Le développement des méthodes d'inférence historique, comme l'ensemble du processus de production de connaissances, est un processus qui prend du temps et s'effectue au cours du temps par l'interaction d'un grand nombre de scientifiques. Il existe ainsi de nombreuses variations dans les méthodes présentées ci-après, en fonction de leur élaboration progressive, mais aussi en fonction des conceptions que peuvent en avoir leurs auteur-e-s ou de leur application particulière sur un corpus donné. Les noms utilisés restent pourtant souvent générique, sans se référer à une description particulière (quand elle existe). Ainsi il est difficile de définir « la » *méthode de Lachmann* (formalisée différemment par Maas [A272]⁷⁷, Salemans [366] ou Bordalejo [49]) ou « la » méthode cladistique. J'ai donc choisi de présenter ces méthodes en me référant principalement à une source unique de description. Mais je n'échappe pas au mélange entre différentes méthodes proches ainsi qu'à la déformation des concepts lors de leur reformulation. Je ne peux donc qu'encourager le recours aux textes originels pour une bonne compréhension de ces méthodes.

On comprend que ce problème rend la comparaison entre différentes méthodes encore plus délicate. Ainsi que compare-t-on vraiment lorsque l'on veut mettre en regard « la » cladistique, « la » stématique et « la » méthode comparative de la linguistique historique ? Sans compter que les méthodes s'influencent aussi entre disciplines différentes⁷⁸. Seule une étude poussée de

77. ↑ Citée par Salemans [366, p. 19].

78. ↑ Ben Salemans pourrait par exemple être influencé, lorsqu'il présente certaines méthodes stématiques [366, pp. 22–35], par ses connaissances de méthodes cladistiques, qu'il connaît nécessairement puisqu'il les décrit et les utilise dans la même publication [366, pp. 35–46].

ces différentes méthodes et de leurs variantes permettrait une compréhension fine des similarités entre celles-ci. Un tel travail dépasse largement le cadre de cette thèse et j'espère simplement donner ici les bases générales de compréhension et de comparaison de ces méthodes, en essayant de les trahir le moins possible.

Choix des méthodes présentées. Il m'est bien entendu impossible de décrire ici l'ensemble des méthodes d'*inférence historique* existantes. Les méthodes présentées ici sont pour la plupart des méthodes respectant les critères suivants :

1. **Inférence historique :** Ces méthodes proposent des hypothèses sur le passé.
2. **Comparaison de structures formelles :** Ces hypothèses sont formulées à partir de la comparaison de structures formelles équivalentes entre plusieurs objets⁷⁹.
3. **Systématique :** La méthodologie, explicitée en amont de l'analyse, s'applique de manière systématique à toutes les données étudiées.
4. **Relations de filiation :** Ces hypothèses portent sur les relations de filiation entre ces objets⁸⁰.

Cependant ces critères ne suffisent pas à caractériser l'ensemble de méthodes décrites ici. En effet toutes les méthodes répondant à ces critères ne sont pas présentées, par exemple les méthodes *lexicostatistiques*. On remarquera aussi une prépondérance des méthodes issues de la *biologie de l'évolution*, imputable peut-être au plus grand nombre de types de méthodes différentes dans cette discipline, mais plus sûrement à ma meilleure connaissance de cette discipline et de ses outils.

J'ai de plus choisi de présenter⁸¹ quelques méthodes d'*inférence phylogéographique* qui cherchent à inférer, à partir des relations de filiation entre un ensemble de *taxons* et des répartitions actuelles de ceux-ci, les positions passées des ancêtres de ces taxons ainsi que leur évolution géographique. Ces méthodes ne satisfont donc pas le dernier critère puisqu'elle n'infèrent pas des relations de filiation mais au contraire se servent de telles relations comme données initiales de l'analyse. L'exploration de ces méthodes issues de la biogéographie des taxons afin d'envisager leur application à des données musicales⁸² a constitué une partie importante de mon travail de première année de thèse et j'ai jugé intéressant de présenter les synthèses réalisées à cette occasion, même si elles ne sont pas exploitées dans la suite de ce travail.

1.3.3 La méthode comparative (linguistique historique)

Principe général. La méthode comparative issue de la linguistique historique s'attache à retrouver les prononciations ancestrales d'un ensemble de langues apparentées et par ce biais les

79. ↑ Voir partie 1.1 page 17.

80. ↑ Mais pas forcément de manière exclusive. Ces méthodes peuvent aussi proposer d'autres hypothèses sur l'histoire, comme l'état ancestral des caractères dans la méthode cladistique.

81. ↑ Voir partie 1.3.8 page 40.

82. ↑ La tentative la plus aboutie de cette application est présentée dans le chapitre 5 pages 155 à 185.

relations d'apparentements entre ces langues (voir figure 1.1). Elle se base sur l'identification et la comparaison de mots de différentes langues ayant vraisemblablement une origine commune (les **cognats**) afin d'identifier des correspondances phonologiques systématiques. Les connaissances sur les mécanismes d'évolution des langues permettent d'inférer une prononciation ancestrale pour chacune de ces correspondances, qui servira à proposer une prononciation ancestrale pour chaque ensemble de cognats.

Identification des cognats. La première étape consiste à identifier des mots (plus exactement des **morphèmes**) qui ont probablement une origine évolutive commune, c'est-à-dire tous issus d'un même morphème ancestral. Cette identification est permise par l'expertise linguistique et constitue une hypothèse *a priori* d'apparentement entre ces morphèmes. Ainsi les morphèmes *capra* (italien), *cabra* (espagnol et portugais) et *chèvre* (/ʃɛvr/) sont supposés provenir d'un même morphème et constituent donc ce que l'on appelle un *ensemble de cognats*. On cherche, afin de réaliser une inférence la plus juste possible, à constituer le plus grand nombre d'ensemble de cognats (quatre sont représentés sur la figure 1.1), comprenant le plus grand nombre de phonèmes (sinon tous) de cette langue.

Établissement de correspondances phonologiques. On cherche ensuite, parmi tous les ensembles de cognats constitués, des *correspondances phonologiques systématiques*, c'est-à-dire des phonèmes de différentes langues qui se retrouvent associés dans différents ensembles de cognats. Par exemple les phonèmes /k/ (italien, espagnol et portugais) et /ʃ/ (graphié *ch* en français) se correspondent systématiquement dans le mot signifiant *chèvre* et dans celui signifiant *cher* (et dans bien d'autres mots). Il peut y avoir différentes correspondances pour un même phonème : ainsi le phonème /k/ (italien, espagnol et portugais) peut aussi correspondre au phonème /k/ du français, comme dans le morphème signifiant *cure*.

Reconstruction de la proto-phonologie. Il s'agit ensuite de déduire, pour chacune des correspondances phonologiques identifiées, la phonologie ancestrale, c'est-à-dire la manière dont était prononcé le plus vraisemblablement le phonème à l'origine de ceux de ce groupe de correspondances. Cette étape fait appel aux connaissances de la linguistique et en particulier celles portant sur les mécanismes d'évolution des langues. Je ne rentrerai donc pas dans le détail des critères permettant d'inférer la proto-phonologie, même si certains sont présentés succinctement dans la figure 1.1, me contentant de donner deux exemples. Le phonème ancestral aux phonèmes /k/ (italien, espagnol et portugais) et /ʃ/ (français) est probablement /*k/⁸³ car il est retrouvé dans un plus grand nombre de langues. Le phonème ancestral aux phonèmes /p/ (italien), /b/ (espagnol et portugais) et /v/ (français) est probablement le phonème /*b/ (occlusive bilabiale voisée) car deux des trois consonnes considérés sont des occlusives (/p/ et /b/), deux sont des bilabiales (/p/ et /b/) et deux sont voisées (/b/ et /v/).

83. ↑ Un phonème ou un morphème précédé d'un astérisque '*' indique qu'il s'agit de la reconstruction proposée d'un phonème ou morphème ancestral.

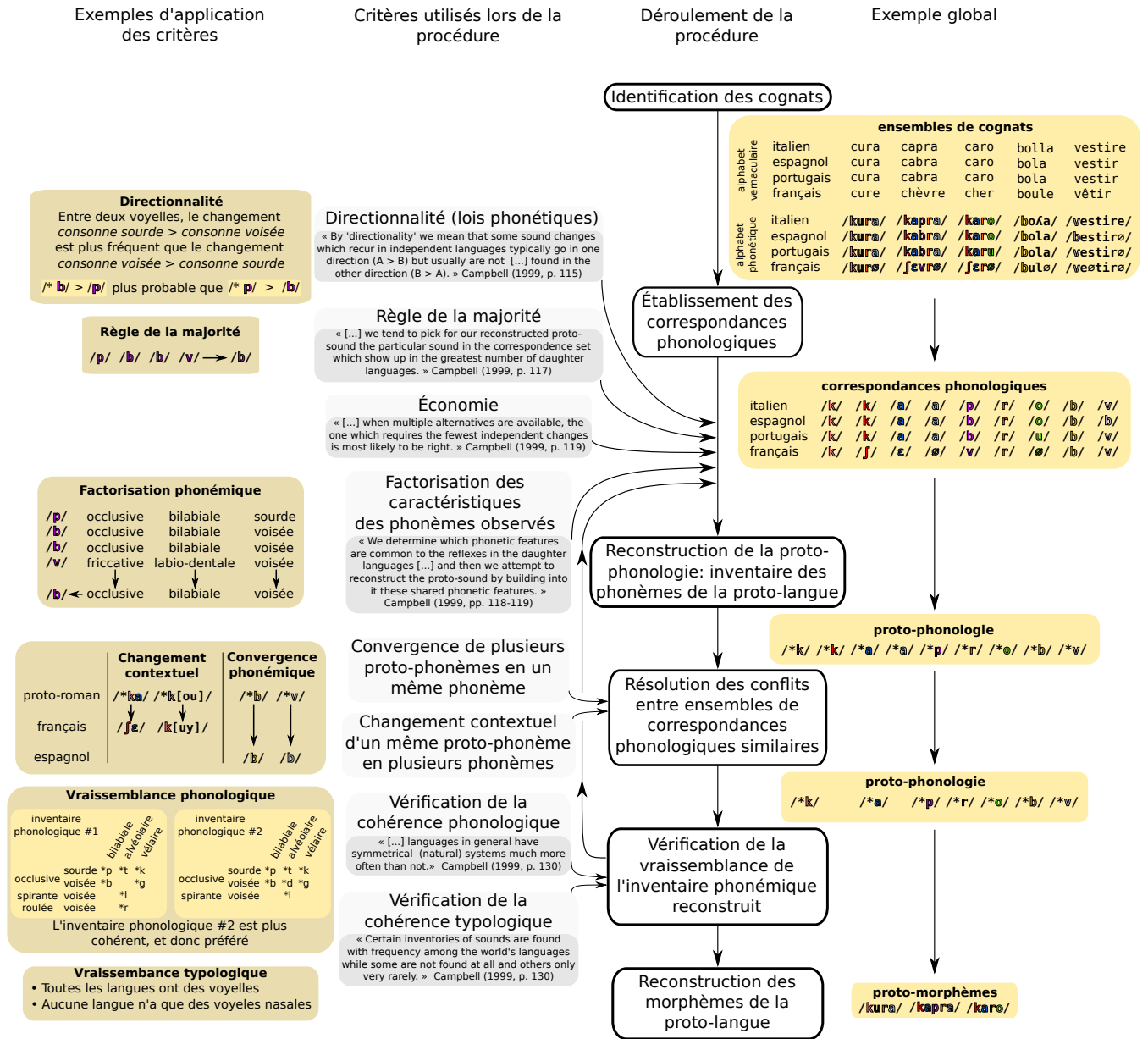


FIG. 1.1 : Déroulement illustré d'une analyse comparative en linguistique historique. Figure réalisée d'après les explications de Campbell (1999) [75, pp. 111-132].

Reconstruction des proto-morphèmes. Les connaissances linguistiques permettent aussi de résoudre les conflits entre différentes correspondances phonologiques similaires ainsi que de vérifier la cohérence de l'inventaire phonologique. Une fois obtenue une proto-phonologie jugée vraisemblable, on peut inférer les morphèmes ancestraux des ensembles de cognats de départ. Ainsi, la proto-phonologie reconstruite pour le proto-roman propose que $/*k/$, $/*u/$ et $/*r/$ soient les proto-phonèmes des correspondances $/k/$, $/u/$ et $/r/$ (pour ces trois correspondances phonologiques les phonèmes sont identiques dans les quatre langues) et que $/*a/$ soit le proto-phonème de la correspondance $/a/$ (italien, espagnol et portugais) et $/\emptyset/$ (français). Ainsi on peut inférer que le proto-morphème de l'ensemble de cognats $/kura/$ (italien, espagnol, portugais) et $/kur\emptyset/$ (français) est $/*kura/$.

1.3.4 Les méthodes stemmatiques

Principe général. Les méthodes **stemmatiques** sont des méthodes de **critique textuelle** visant à inférer les relations d'apparentement entre différents manuscrits et à en reconstruire le texte originel. Je présente ici (et dans la figure 1.2) celle proposée par Salemans [366, pp. 22–35] en 2000, largement inspirée de celles de Greg [A173]⁸⁴ et Dearing [A113]⁸⁴. Après avoir identifié toutes les positions du texte où il existe des variations entre les différents manuscrits, ces positions (appelées *lieux variants*) sont classées en différents types. Seuls certains lieux variants, dits de *type 2*, sont utilisés pour construire une chaîne reliant les différents manuscrits en fonction des **variantes** qu'ils partagent. Si la construction de cette chaîne est possible, on cherche les manuscrits qui pourraient constituer des intermédiaires dans la chaîne grâce à une autre catégorie de lieux variants, dits de *type 1*. Enfin l'identification des variantes originales permet d'inférer quels sont les manuscrits les plus anciens et d'obtenir une hypothèse des relations de filiation entre les manuscrits orientée dans le temps, appelée *stemma*.

Classification des lieux variants. En chaque position du texte où est observée une différence entre les manuscrits, les variantes présentes dans chacun des manuscrits sont notées. Cela permet de catégoriser ce lieu variant dans un des cinq types, présentés dans le tableau de la figure 1.2. Cette typologie est fondée sur le nombre de variantes différentes trouvées en cette position du texte et sur le nombre de manuscrits regroupés par une même variante. Une variante définit ainsi un « vrai » groupe lorsqu'elle est présente dans au moins deux manuscrits. Seuls deux types de lieux variants seront utilisés pour inférer les relations génétiques entre les manuscrits. Les lieux variants de type 2, définissant exactement deux vrais groupes, serviront à construire la chaîne reliant les manuscrits. Les lieux variants de type 1, définissant exactement deux groupes dont l'un est vrai et l'autre ne comprend qu'une variante, serviront une fois la chaîne construite à identifier les manuscrits pouvant être des intermédiaires dans la chaîne.

84. ↑ Citées par Salemans [366, pp. 22–23].

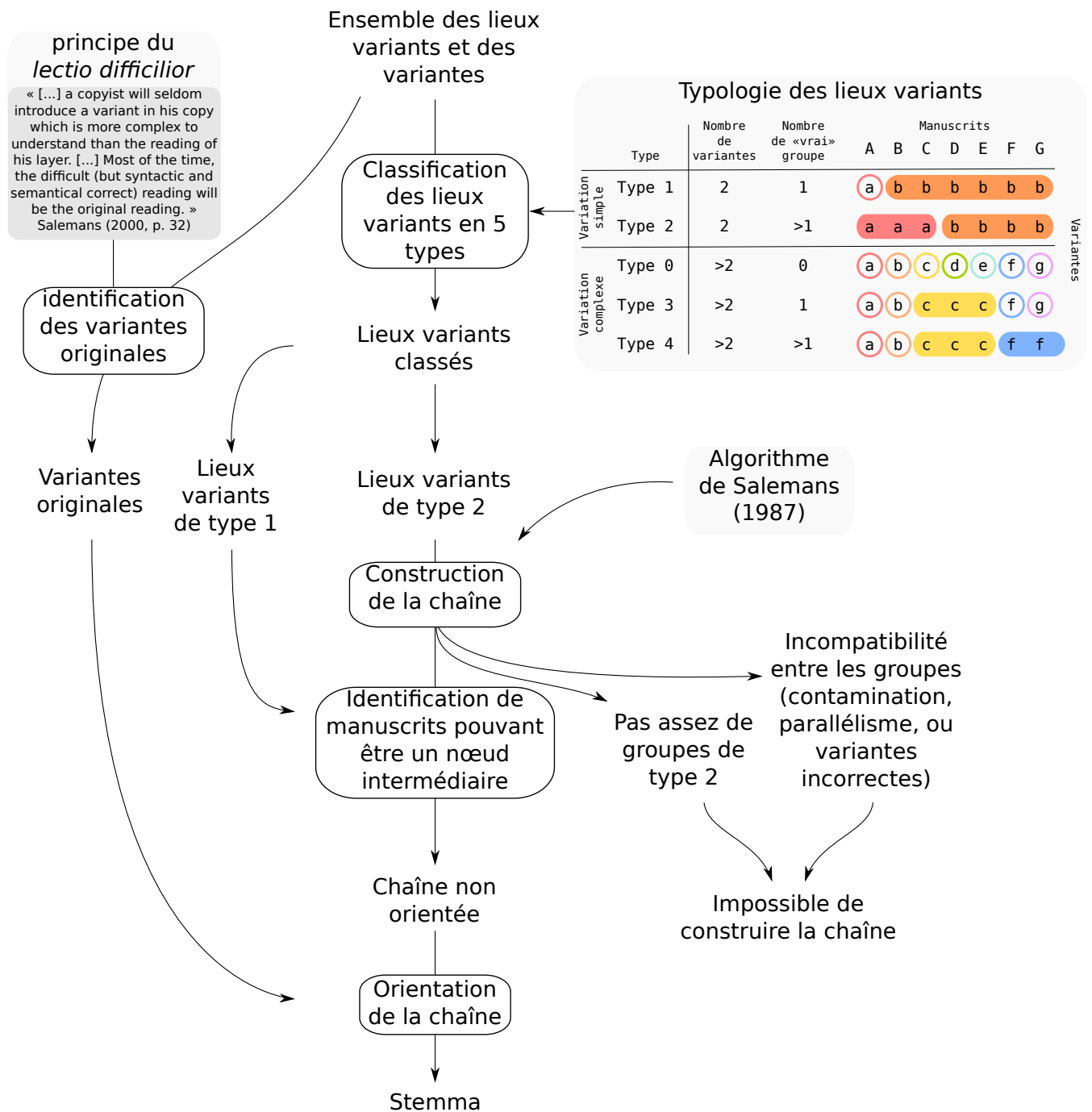


FIG. 1.2 : **Déroulement illustré d'une analyse stemmatique.** Figure réalisée d'après les explications de Salemans [366, pp. 22–35], qui nomme cette technique « méthode de Greg/Dearing » en référence aux travaux de Greg [A173] et Dearing [A113].

Construction de la chaîne reliant les manuscrits. Ben Salemans [366, pp. 27–31] propose un algorithme pour construire une chaîne reliant de proche en proche les manuscrits. Pour chaque lieu variant de type 2 (définissant donc exactement deux groupes d'au moins deux manuscrits), le groupe contenant le moins de manuscrits est sélectionné, constituant une liste de groupes de type 2. L'algorithme relie ensuite les manuscrits présents dans un même groupe (en partant des groupes de plus faible effectif), puis y ajoute progressivement les manuscrits isolés lorsqu'un groupe de type 2 comprend des manuscrits déjà reliés et exactement un manuscrit non relié. Lorsque l'on ne peut plus faire de regroupements, les différents groupes constitués sont reliés les uns aux autres.

Cet algorithme ne permet cependant pas de construire une chaîne dans tous les cas. Lorsqu'un groupe contient des manuscrits déjà reliés et au moins deux autres non reliés, l'algorithme ne permet pas de hiérarchiser les relations au sein de ce groupe et nécessiterait plus de lieux variants de type 2 pour pouvoir continuer. Par ailleurs, lorsque plusieurs lieux variants définissent des groupes incompatibles, l'algorithme ne peut aboutir car il suppose un emboîtement strict des relations entre manuscrits, sans définir de critère permettant de trancher en cas de conflit.

Nœuds intermédiaires et orientation de la chaîne. Les lieux variants de type 1, pour lesquels tous les manuscrits ont la même variante sauf un qui a une variante propre à lui seul, peuvent ensuite être utilisés pour identifier les manuscrits intermédiaires dans la chaîne synthétisant les relations d'apparementement entre manuscrits. En effet, si un manuscrit ne présente aucune variante propre à lui seul, on peut penser (mais sans complète certitude) qu'il est un manuscrit intermédiaire qui a servi comme source pour la copie d'autres manuscrits et que toutes ses variantes ont donc été recopiées. Si en revanche au moins une variante propre à ce seul manuscrit est trouvée, ce manuscrit ne peut être un intermédiaire [366, p. 27].

La chaîne ainsi construite n'est pas orientée et ne permet pas de savoir quelles relations sont les plus récentes et lesquelles sont plus anciennes. Pour cela, il est nécessaire d'identifier, au moins pour certains lieux variants, quelle est la variante présente dans le texte à l'origine de tous les manuscrits (l'*archétype*). Le principe du *lectio difficilior* permet d'apporter une solution à ce problème. Il part du principe que les copistes simplifient le texte et ne le complexifient pas, de telle manière qu'il soit plus facile à comprendre. Ainsi les variantes les plus complexes seront considérées comme celles d'origine, alors que les plus simples seront considérées comme issues d'une modification. Le recoupement de l'information apportée par différents lieux variants quant aux variantes originelles peut permettre d'identifier la position de l'archétype au sein de la chaîne et ainsi de l'*orienter* dans le temps. La chaîne, une fois orientée, est appelée *stemma*.

1.3.5 La compatibilité des caractères

La méthode de compatibilité est une méthode d'inférence phylogénétique, développée dans les années 1970. Comme la méthode *cladistique*⁸⁵, elle se fonde sur l'information fournie par la

85. ↑ Voir partie 1.3.6 page 34.

comparaison des états des caractères observés chez les différents taxons étudiés, mais elle rejette complètement l'homoplasie là où l'analyse cladistique la minimise.

1.3.5.1 Principe

Cette méthode repose sur le principe que seuls les états de caractères apparus une seule fois au cours de l'évolution des taxons étudiés et qui n'ont pas subi de réversion sont porteurs d'information fiable pour l'inférence phylogénétique : eux seuls devraient donc être utilisés. Or les informations fournies par plusieurs caractères peuvent être incompatibles, nécessitant de faire appel à une hypothèse *ad hoc* de transformation de certains caractères, alors considérés comme des homoplasies. La méthode de compatibilité choisit d'éliminer complètement les caractères incompatibles et recherche pour cela le plus grand ensemble de caractères tous mutuellement compatibles, c'est-à-dire ceux pour lesquels on peut proposer un arbre phylogénétique sans aucune homoplasie⁸⁶. Un tel ensemble de caractères est appelé une clique [107, p. 145]. La méthode de compatibilité cherche donc à construire l'arbre phylogénétique à partir de la clique contenant le plus grand nombre de caractères parmi ceux analysés. Il faut donc commencer par rechercher la clique la plus nombreuse.

1.3.5.2 Méthodes de recherche

Le problème de la recherche de la clique la plus nombreuse est un problème NP-complet⁸⁷ [A111]⁸⁸ : il est ainsi difficile de trouver cette clique en un temps raisonnable. En pratique, lorsque les caractères utilisés sont binaires⁸⁹ il existe une procédure de reconstruction assez efficace. En revanche, lorsque les caractères ont des états multiples, il faut soit ramener le problème à une situation où les caractères sont binaires, soit utiliser des méthodes dites de *phylogénie parfaite*⁹⁰ (figure 1.3).

Cas de caractères binaires. Lorsque les caractères utilisés sont tous binaires, la recherche de l'ensemble des cliques est facilitée par l'application du théorème de compatibilité mutuelle⁹¹. En effet, celui-ci précise :

« A set S of characters has all pairs of characters compatible with each other if and only if all of the characters in the set are jointly compatible (in that there exists a tree with which all of them are compatible). »⁹² Felsenstein (2004) [136, p. 90]

86. ↑ « Des caractères sont dits *mutuellement compatibles* quand il existe un arbre qui rend compte des changements d'états de ces caractères sans nécessiter d'hypothèses d'homoplasie. » [107, p. 145]

87. ↑ Les problèmes NP-complets sont une classe de problèmes définie par la théorie de la complexité. À l'heure actuelle, nous ne savons pas trouver efficacement l'ensemble des solutions d'un problème NP-complet : le temps d'exécution des algorithmes résolvant ces problèmes de manière exhaustive augmente généralement de manière exponentielle avec la taille des données.

88. ↑ Citée par Felsenstein [136].

89. ↑ Un caractère est dit *binaire* s'il n'est présent dans l'ensemble des unités évolutives étudiées que sous deux états différents. L'entrée correspondante du glossaire donne quelques exemples.

90. ↑ *Perfect phylogeny methods*, en anglais.

91. ↑ *Pairwise Compatibility Theorem* en anglais, d'après l'appellation de Felsenstein [136, p. 90].

92. ↑ « Un ensemble S de caractères a tous ses caractères compatibles deux à deux si et seulement si tous les

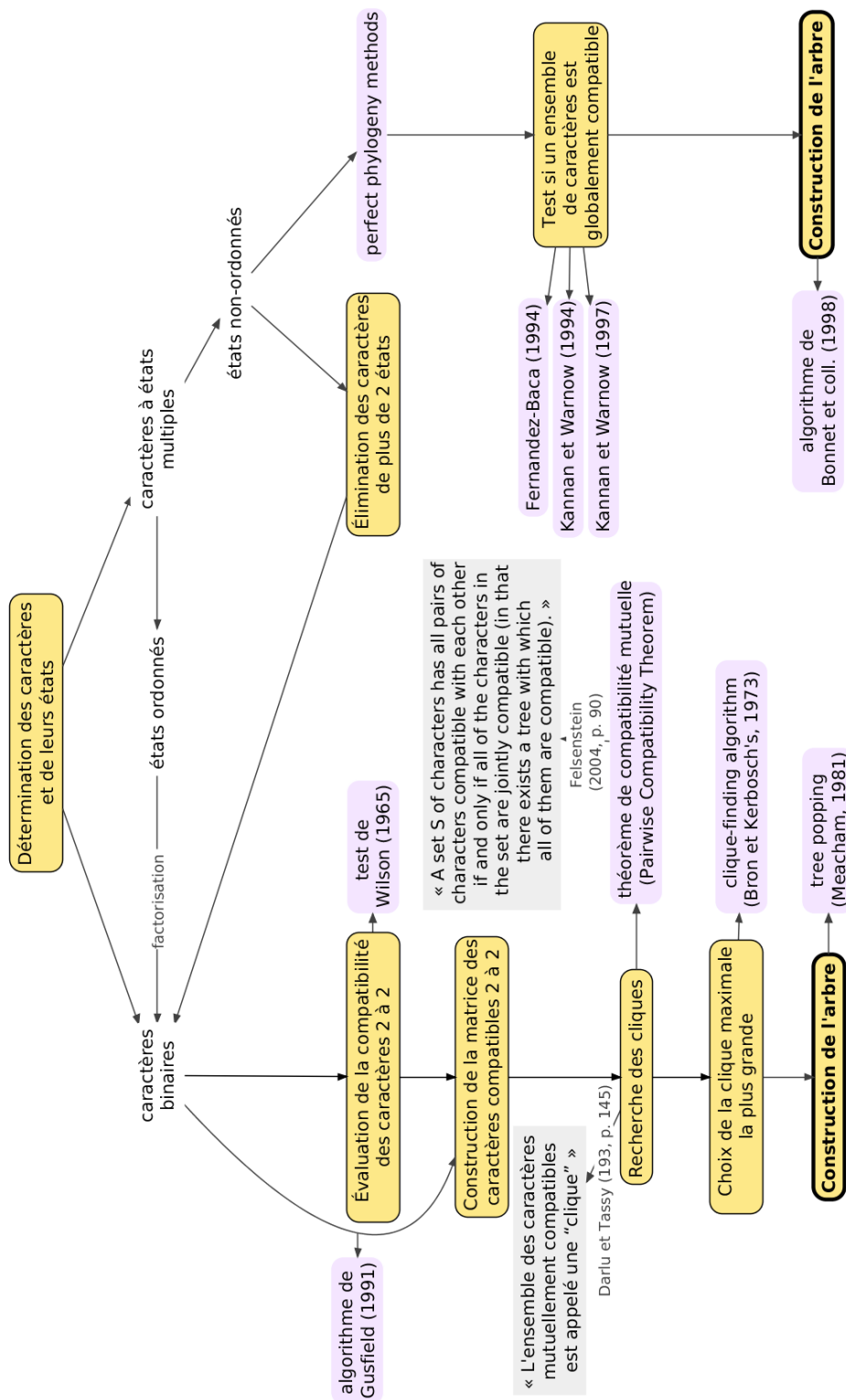


FIG. 1.3 : **Déroulement d'une analyse de compatibilité des caractères.** Les éléments oranges représentent les différentes étapes de la méthode, les éléments violets représentent les outils développés pour réaliser ces étapes associées et les rectangles gris des citations définissant certains termes. Figure réalisée d'après les explications de Daru et Tassy [107, pp. 145–151] et de Felsenstein [136, pp. 87–96].

Il s'agit alors d'évaluer la compatibilité de chaque paire de caractères, grâce au test de Wilson [A442]⁹³, puis de construire la matrice des caractères **mutuellement compatibles**. Le théorème ci-dessus permet ensuite d'identifier la ou les cliques les plus nombreuses. Il suffit enfin de construire l'arbre à partir de cet ensemble de caractères globalement compatibles, par exemple en utilisant l'algorithme de *tree popping* [A288]⁹⁴.

Cas des caractères à états multiples. Lorsque certains caractères ont plus de deux **états**, le théorème de compatibilité mutuelle n'est plus applicable dans tous les cas [136, p. 94]. Si les états des caractères sont ordonnés⁹⁵, il est possible de factoriser le caractère – c'est-à-dire le recoder sous forme de plusieurs caractères binaires [107, p. 92] – et la méthode décrite ci-dessus est applicable. Si les états ne sont pas ordonnés, une telle factorisation n'est pas possible et deux solutions sont alors envisageables. Felsenstein [136, p. 96] propose d'éliminer du jeu de données les caractères possédant plus de deux états, partant du principe que la méthode de compatibilité cherche à ne conserver que les caractères ne s'étant transformés qu'une seule fois. Il est aussi possible d'utiliser les méthodes dites de « phylogénie parfaite », qui sont capables de tester si un ensemble de caractères est **mutuellement compatible** puis de construire l'arbre correspondant. Ces méthodes ne peuvent cependant pas trouver directement la clique la plus nombreuse.

1.3.6 Les méthodes cladistiques

Principe général. La méthode **cladistique** présentée ici est une méthode d'**inférence phylogénétique** issue (entre autres) des travaux de Hennig [A190]⁹⁶, telle qu'exposée par Darlu et Tassy [107]. Elle propose un **arbre phylogénétique** représentant les relations de parentés entre les entités considérées (les **taxons**⁹⁷, par exemple des espèces biologiques) à partir de l'analyse systématique de **caractères**. Cette méthode se distingue d'autres méthodes d'inférence phylogénétique (notamment les **méthodes de distances**) en ce qu'elle utilise toute l'information contenue dans les caractères et en ce qu'elle émet des hypothèses *a priori* sur les formes des caractères (les **états**) les plus anciennes. Un exemple de cette méthode est présentée sur la figure 1.4 où sont comparées quatre espèces végétales, qui constituent chacune un taxon de l'analyse. Les paragraphes suivant expliquent succinctement le principe de cette méthode. Pour plus de détail concernant les variations possibles de l'analyse cladistique on pourra se référer à la figure 1.5 et surtout au livre de Darlu et Tassy [107].

caractères dans l'ensemble sont globalement compatibles (c'est-à-dire qu'il existe un arbre avec lequel ils sont tous compatibles). »

93. ↑ Citée par Felsenstein [136].

94. ↑ Citée par Felsenstein [136].

95. ↑ C'est-à-dire que l'on ne peut passer de n'importe quel état à n'importe quel autre : les relations entre les états sont hiérarchisées (et non cycliques).

96. ↑ Citée par Darlu et Tassy [107, p. 31].

97. ↑ Les méthodes cladistiques font usage d'un abondant vocabulaire spécifique, dont l'explication est plus détaillée dans le glossaire (volume III, pages 327–370).

Choix des caractères. Cette méthode nécessite le choix au préalable des caractères pertinents pour inférer les relations phylogénétiques. Il s'agit d'identifier des caractéristiques observables (on parle de *caractère*) que l'on pense issues d'un même caractère ancestral dans les différents taxons étudiés. Les pattes des rongeurs, des carnivores et des marsupiaux sont héritées des membres⁹⁸ de l'ancêtre commun aux mammifères et sont donc dites *homologues*. En revanche l'aile des oiseaux, celle des insectes ou encore celle des chauve-souris ne provient pas d'une structure existant chez l'ancêtre des bilatériens : ces structures sont donc qualifiées d'*analogues*. Seuls les caractères homologues peuvent nous renseigner sur l'histoire du groupe étudié et doivent donc être utilisés lors de la procédure d'inférence⁹⁹.

Identification des états de caractères. On décrit ensuite les variantes de ces caractères, appelées *états*, au sein du groupe étudié. Dans l'exemple de la figure 1.4, cinq caractères sont analysés : la présence de fleurs, de feuilles, le nombre de cotylédons dans l'embryon, le type de sol où pousse la plante et la présence de bois secondaire. Le caractère « fleurs » peut prendre deux états dans l'échantillon analysé : soit il est présent, soit il est absent. Comme on le voit sur cette figure, les états de caractère ne sont pas nécessairement la présence ou l'absence et il peut y avoir plus ou moins de deux états différents pour un même caractère. On constitue ainsi une *matrice de caractères* donnant, pour chaque taxon et pour chaque caractère analysé, l'état de ce caractère dans ce taxon. La *matrice de caractères* de la figure 1.4 décrit ainsi (caractère C4) le fait que le pin blanc de Provence et le brome dressé poussent sur des sols basiques, alors que le crocus d'automne et la bruyère callune poussent sur des sols acides.

Polarisation des états de caractères. Il s'agit ensuite d'identifier, pour chaque caractère, quel est l'état de caractère qui était présent chez l'ancêtre commun à tous les taxons étudiés, appelé *état ancestral* (par opposition aux états de caractères issus de modifications ultérieures, appelés *états dérivés*). Cette étape, appelée *polarisation*, est essentielle à l'analyse cladistique : les taxons sont regroupés à partir de l'identification de caractères dérivés partagés (appelés *synapomorphies*). En effet le partage de caractères ancestraux ne nous apprend rien sur les relations d'apparentement. Plusieurs critères peuvent être utilisés pour polariser les caractères, le plus courant étant la comparaison avec des *groupes externes* (ou *extra-groupes*) : on choisit des taxons que l'on sait être plus éloignés que tous les autres taxons entre eux, chacun de ces taxons constituant un groupe externe. L'état ancestral de chacun des caractères analysés est déduit de leurs états observés chez les groupes externes. Dans l'exemple de la figure 1.4, on sait que le pin est plus éloigné des trois autres espèces d'angiosperme que l'on cherche à classer : il est choisi

98. ↑ Ces membres sont appelés membres *chiridiens*.

99. ↑ Le résultat de l'analyse cladistique peut cependant nous laisser penser qu'un caractère considéré au départ comme homologue est analogue. Aussi le fait de considérer un caractère comme homologue au départ de l'analyse est toujours une hypothèse *a priori* (parfois appelée *homologie primaire*), susceptible d'être remise en cause *a posteriori*. Une hypothèse d'homologie issue du résultat de l'analyse cladistique est qualifiée d'*homologie secondaire*.

comme groupe externe¹⁰⁰. C'est donc à partir de lui que l'on déterminera l'état ancestral des caractères : comme le pin n'a pas de fleur, l'absence sera l'état ancestral du caractère « fleur ». En revanche, comme le pin a du bois secondaire, l'état ancestral du caractère « bois secondaire » est la présence.

Recherche de l'arbre le plus parcimonieux. La méthode cladistique cherche parmi l'ensemble des arbres que l'on peut possiblement construire, celui qui constitue la meilleure hypothèse concernant les relations d'apparentement entre les taxons étudiés. Pour cela, elle retrace pour chaque arbre les modifications des états de caractères nécessaires pour arriver à l'état observé. C'est une méthode basée sur le principe de la **parcimonie** : le ou les arbres préférés seront ceux qui minimisent le nombre de transformations nécessaires pour expliquer l'état actuel de tous les caractères chez tous les taxons étudiés.

Deux arbres pouvant schématiser les relations d'apparentement entre les quatre espèces comparées sont présentés sur la figure 1.4. Dans le premier, le crocus est plus proche de la bruyère que du brome. Pour expliquer l'observation actuelle des états du caractère C4, le plus simple est de considérer que l'ancêtre commun aux quatre espèces vivait en milieu basique et que ce caractère a changé d'état (passant de « basique » à « acide ») chez l'ancêtre du crocus et du brome : un seul changement d'état est nécessaire. Par contre pour le caractère C5, on voit que deux changements d'états sont nécessaires : le bois secondaire étant présent chez l'ancêtre commun des quatre espèces, il a dû disparaître deux fois, de manière indépendante, chez le brome et chez le crocus¹⁰¹.

Dans le second arbre, le crocus est plus proche du brome que de la bruyère. Dans ce cas, le caractère C5 ne nécessite qu'un changement d'état, alors que le caractère C4 en nécessite deux. De manière globale, on se rend compte qu'il faut sept changements d'états pour expliquer l'évolution de tous les caractères sur l'arbre 1, alors qu'il n'en faut que six sur l'arbre 2 : ce dernier est donc considéré comme plus parcimonieux et donc préféré comme hypothèse concernant les relations évolutives entre ces quatre espèces.

1.3.7 Les méthodes probabilistes

Ces méthodes sont basées sur une conception probabiliste de l'évolution, selon laquelle les modifications observées au cours de l'évolution ne sont ni déterminées, ni complètement aléatoires, mais que leur apparition peut être modélisée par des lois de probabilités. Utilisées principalement dans l'étude de l'évolution des séquences nucléotidiques, elles permettent d'intégrer explicitement des hypothèses évolutives dans l'inférence phylogénétique. On distingue deux catégories principales de méthodes probabilistes. Les méthodes dites de *maximum de vraisemblance*¹⁰² cherchent

100. ↑ Cet exemple ne montre qu'un seul groupe externe, mais il est nécessaire de faire appel à plusieurs groupes externes pour lever toute ambiguïté [107, pp. 46–49].

101. ↑ Une autre possibilité serait que le bois secondaire ait disparu chez l'ancêtre du brome, du crocus et de la bruyère, puis ait réapparu chez la bruyère. Dans les deux cas, deux changements d'états sont nécessaires.

102. ↑ *Maximum likelihood*, en anglais.

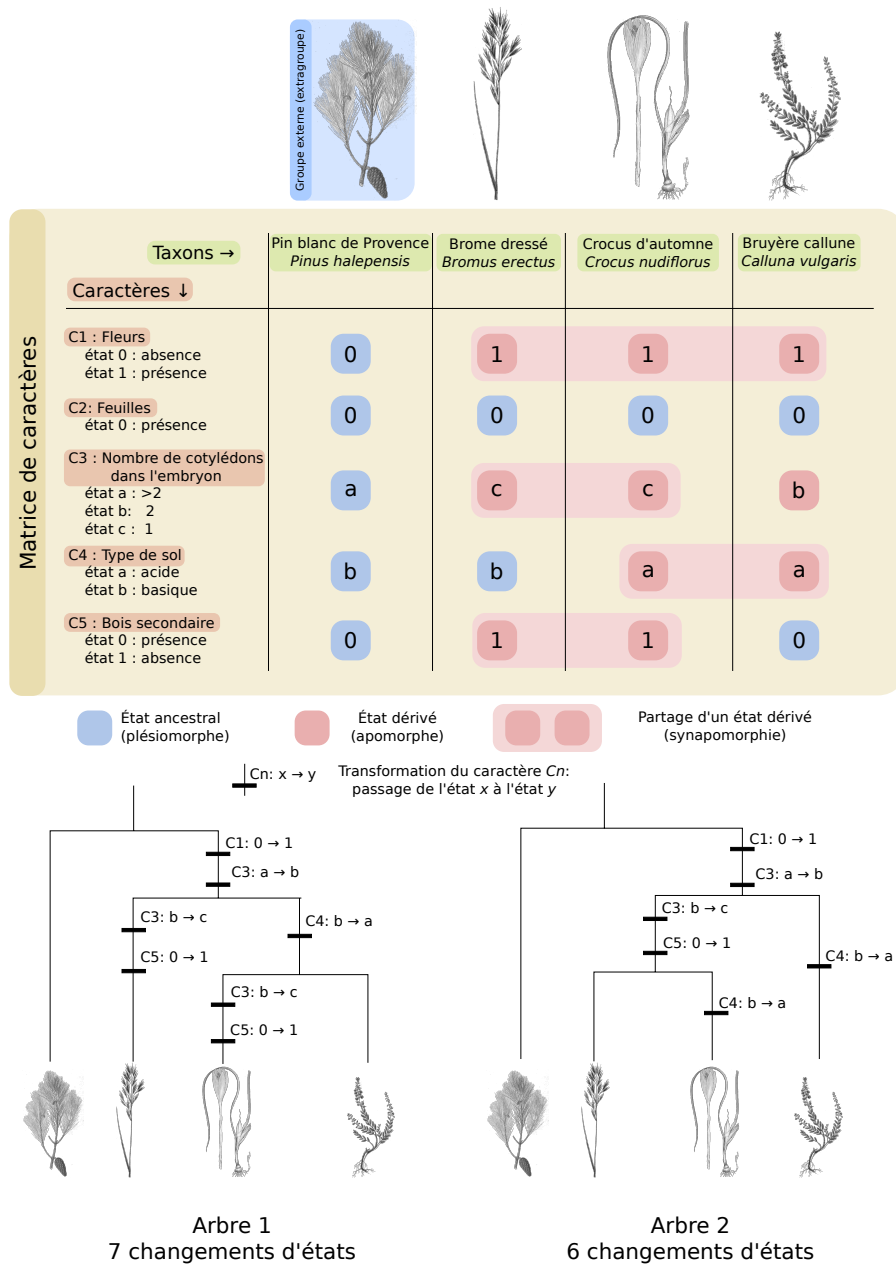


FIG. 1.4 : **Exemple d'analyse cladistique.** La partie supérieure de la figure montre la matrice décrivant les états pris par cinq caractères (C1 à C5) chez quatre taxons (dans ce cas des espèces végétales). La partie inférieure montre deux propositions d'arbre schématisant les relations d'apparentement entre les taxons, ainsi que les changements d'états de caractères nécessaires pour expliquer dans chacun de ces deux cas les observations de la matrice de caractères. Le caractère « type de sol » est un caractère de nature écologique désignant le type de sol sur lequel pousse préférentiellement l'espèce. J'ai choisi de le présenter dans cet exemple car c'est un caractère manifestement **convergent**. En pratique les caractères écologiques sont rarement utilisés dans l'inférence phylogénétique. Les sources des image utilisées dans cete figure sont données dans le tableau de l'annexe D page 457.

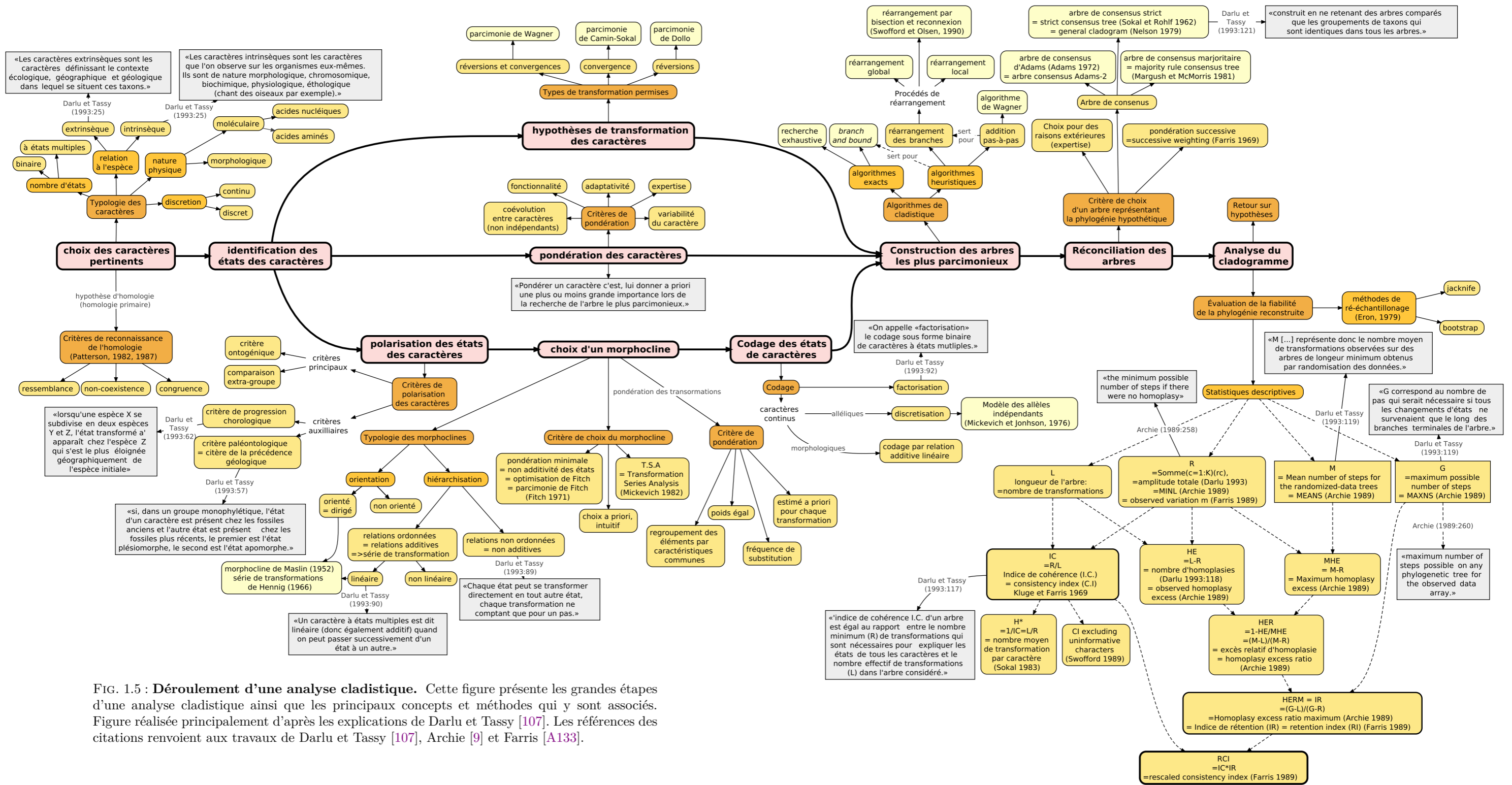


FIG. 1.5 : Déroulement d'une analyse cladistique. Cette figure présente les grandes étapes d'une analyse cladistique ainsi que les principaux concepts et méthodes qui y sont associés. Figure réalisée principalement d'après les explications de Darlu et Tassy [107]. Les références des citations renvoient aux travaux de Darlu et Tassy [107], Archie [9] et Farris [A133].

à trouver l'arbre dont la probabilité d'engendrer (sous un certain modèle d'évolution choisi *a priori*) les données observées est la plus grande. Les méthodes d'inférence bayésienne cherchent en revanche à évaluer la probabilité de différentes hypothèses évolutives à partir des données observées.

1.3.7.1 Les méthodes de maximum de vraisemblance

Ces méthodes [116] définissent une fonction de vraisemblance, qui exprime la probabilité d'obtenir le jeu de données observé en fonction des différentes hypothèses évolutives envisagées. Ces hypothèses portent à la fois sur l'histoire évolutive des taxons considérés (**topologie** et longueur des branches de l'**arbre phylogénétique**) et sur les modalités d'évolution du matériau considéré. Lorsqu'il s'agit, comme c'est souvent le cas, de séquences nucléotidiques, ces modalités incluent par exemple les probabilités de substitution entre les nucléotides, ou encore l'hétérogénéité des taux de substitution selon les régions de l'arbre.

1.3.7.2 L'inférence bayésienne

Principe général. Cette méthode¹⁰³ est fondée sur le théorème de Bayes, qui permet d'exprimer la probabilité d'un événement A_i en sachant qu'un événement B est vrai en fonction de la probabilité que l'événement B soit observé selon les différentes possibilités offertes par des événements A_j mutuellement incompatibles et couvrant l'ensemble des possibles [117] :

$$P(A_i|B) = \frac{P(B|A_i) \times P(A_i)}{\sum_j P(B|A_j) \times P(A_j)} \quad (1.1)$$

où $P(X)$ représente la probabilité de l'événement X et $P(X|Y)$ la probabilité de X sachant Y .

Cette méthode cherche à calculer la probabilité d'un ensemble d'hypothèses évolutives en prenant en compte le fait que les données issues du terrain sont effectivement observées et donc nécessairement « vraies »¹⁰⁴. D'après le théorème de Bayes, cette probabilité *sachant les données* peut être calculée à partir de la probabilité d'observer ces données selon toutes les hypothèses envisageables. Encore une fois, les hypothèses évolutives peuvent inclure non seulement la topologie et les longueurs de branches de l'arbre phylogénétique, mais aussi les modalités d'évolution du matériau étudié. Il s'agit alors de trouver l'ensemble d'hypothèses évolutives dont la probabilité *sachant les données* est la plus grande.

Analyse bayésienne empirique. Dans le cas de l'*analyse bayésienne empirique*, les modalités d'évolution sont choisies *a priori*, par exemple à l'aide d'une estimation par maximum de vraisemblance. Seul l'arbre phylogénétique est inféré par la méthode bayésienne.

103. ↑ Un exemple d'inférence bayésienne est présenté en détail dans le chapitre 5 pages 155 à 185.

104. ↑ On parle alors de la *probabilité sachant les données*.

Analyse bayésienne hiérarchique. En revanche, dans le cas de l'*analyse bayésienne hiérarchique*, les modalités d'évolution du matériau sont aussi estimées par la méthode bayésienne. On cherche donc à inférer à la fois l'arbre phylogénétique et les paramètres évolutifs (vitesse d'évolution, ...). Il faudrait donc calculer la probabilité d'obtenir les données observées pour tous les paramètres possibles du modèle d'évolution. Cependant le nombre de combinaisons possibles des différents paramètres est très grand et le calcul est impossible, même avec un petit jeu de données. La solution la plus courante pour contourner ce problème est d'utiliser la *méthode de Monte-Carlo par chaînes de Markov* (MCMC)¹⁰⁵. Cette méthode repose sur le principe de la marche guidée¹⁰⁶ : les hypothèses sont choisies arbitrairement, puis modifiées à chaque étape selon des critères permettant d'optimiser les chances de trouver l'ensemble d'hypothèses de probabilité maximale¹⁰⁷.

1.3.8 La biogéographie historique des taxons

La *biogéographie historique* est une discipline de la biogéographie s'attachant à analyser les « relations entre la structure et l'histoire des peuplements d'une part, et l'histoire géologique de la surface du globe » [204, p. 29]. La biogéographie des taxons est un sous-ensemble de cette discipline qui s'intéresse à l'histoire de la distribution spatiale d'un taxon particulier¹⁰⁸. On l'oppose à la *biogéographie des aires* qui cherche à inférer l'histoire d'une aire biogéographique, généralement grâce à l'analyse de la distribution et des relations de parentés de nombreuses espèces y cohabitant. La figure 1.6 propose une classification de méthodes de biogéographie des taxons.

Cette partie présente de manière succincte quatre méthodes de biogéographie des taxons : les inférences d'aire ancestrale par optimisation de Camin–Sokal¹⁰⁹, par optimisation de Fitch¹¹⁰, par optimisation pondérée¹¹¹ et l'inférence par analyse de dispersion–vicariance¹¹².

1.3.8.1 Optimisation de Camin–Sokal

Bremer [A63]¹¹³ formalise en 1992 une méthode d'*inférence phylogéographique* s'inspirant des principes de la recherche du centre d'origine et des voies de dispersion, lesquels étaient largement utilisés à la fin du XIX^{ème} siècle et au début du XX^{ème} siècle. La figure 1.7 page 43 illustre le principe de cette méthode. Il s'agit de considérer le *cladogramme* des aires, directement dérivé du cladogramme des espèces inféré par une méthode de reconstruction phylogénétique habituelle, puis d'utiliser un critère de *parcimonie* pour évaluer la probabilité qu'une zone donnée ait ap-

105. ↑ Delsuc et Douzery [116, p. 77].

106. ↑ Le principe de la marche guidée est expliqué plus en détail dans la partie 5.2.3.2 page 162.

107. ↑ Delsuc et Douzery [116, pp. 77–81].

108. ↑ « Taxon biogeography aims to reconstruct only the distributional history of a taxon [...] » [100, p. 138]

109. ↑ Partie 1.3.8.1 page 40.

110. ↑ Partie 1.3.8.2 page 42.

111. ↑ Partie 1.3.8.3 page 42.

112. ↑ Partie 1.3.8.4 page 42.

113. ↑ Citée par Crisci [100, pp. 43–45].

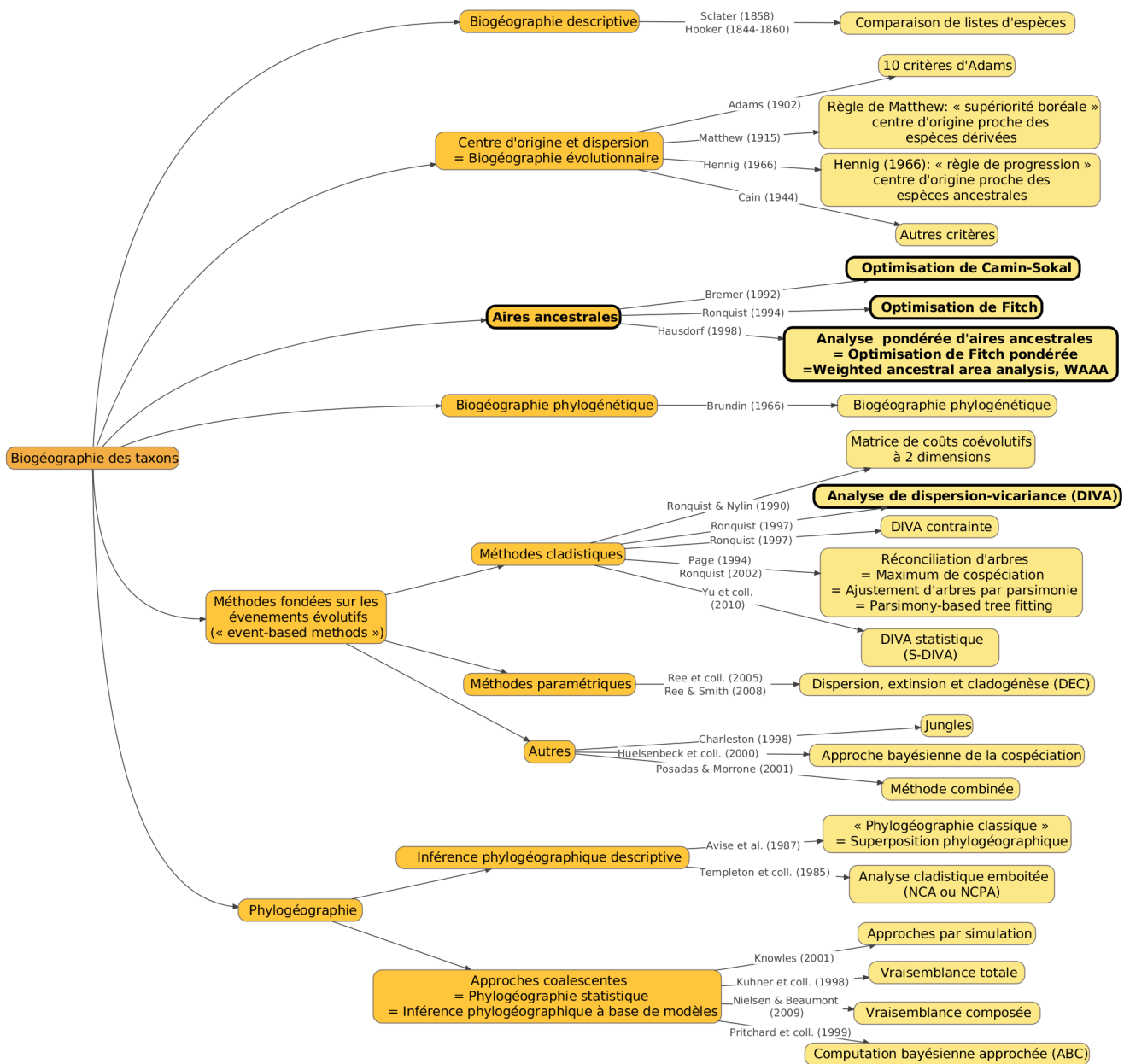


FIG. 1.6 : Classification de méthodes de biogéographie des taxons. Figure réalisée d'après les publications de Crisci [100], Hugot [204], Hickerson et coll. [192] et Lomolino et coll. [266]. On pourra retrouver dans ces travaux les références détaillées des méthodes. Les méthodes en gras sont celles présentées dans cette thèse.

partenu à l'aire ancestrale. On compare ainsi pour chaque aire le nombre de « gains d'espèce » nécessaires si la zone n'était pas dans l'aire ancestrale avec le nombre de « pertes d'espèce » requis dans le cas où la zone était dans la zone ancestrale. S'il faut peu d'hypothèses *ad hoc* en supposant que l'aire ne fait pas partie de l'aire ancestrale, mais beaucoup en supposant qu'elle en fait partie, c'est vraisemblablement qu'elle n'en faisait pas partie (et inversement).

1.3.8.2 Optimisation de Fitch

Ronquist [A359]¹¹⁴ propose en 1995 une méthode inspirée de celle de Bremer [A63]¹¹³, basée sur l'optimisation de Fitch (voir figure 1.8 page 44). Contrairement à cette dernière, il ne considère que le cas où la zone étudiée est supposée appartenir à l'aire ancestrale et regarde le nombre de gains et pertes d'espèces minimal nécessaire pour expliquer la répartition actuelle : plus ce nombre de changement est faible, plus il est vraisemblable que la zone ait appartenu à l'aire ancestrale.

1.3.8.3 Optimisation pondérée

Hausdorf [A185]¹¹⁵ propose en 1998 une modification de la méthode de Bremer [A63]¹¹³ : l'inférence d'aire ancestrale par optimisation pondérée (présentée dans la figure 1.9 page 45). Les transformations sont pondérées par leur emplacement dans l'arbre : plus elles ont lieu sur une branche basale¹¹⁶ plus leur poids est faible. Cette méthode suppose donc que les espèces les moins dérivées procurent plus d'information sur l'emplacement de l'aire ancestrale.

1.3.8.4 Analyse de dispersion–vicariance (DIVA)

L'analyse de dispersion–vicariance (ou DIVA pour l'anglais *Dispersal–Vicariance Analysis*) est une méthode proposée en 1997 par Ronquist [360] permettant d'inférer l'aire occupée par les différents ancêtres hypothétiques d'un ensemble d'espèces, dont la phylogénie et la distribution actuelle sont connues. La figure 1.10 présente les principales étapes de cette méthode. Celle-ci postule que les événements de vicariance (formation de deux espèces dans deux zones distinctes par apparition d'une barrière entre les deux zones) et de duplication (apparition d'une nouvelle espèce dans une seule zone, par spéciation sympatrique ou péripatrique) ont un coût évolutif nul, alors que les événements d'extinction (disparition d'une espèce dans une zone) ou de dispersion (apparition d'une espèce dans une nouvelle zone) ont un coût de 1. La méthode d'optimisation en trois étapes¹¹⁷ permet alors de trouver (en tenant compte des règles d'inférence fixées *a priori*) les aires de distribution aux différents nœuds de l'arbre qui minimise le nombre d'événements d'extinction et de dispersion.

114. ↑ Citée par Crisci [100, pp. 45–47].

115. ↑ Citée par Crisci [100, pp. 47–49].

116. ↑ Une branche est considérée comme basale s'il existe peu de nœuds internes entre cette branche et la racine de l'arbre.

117. ↑ Ou *step-matrix optimization*. Le détail de la procédure n'est pas présenté ici, on pourra se référer à l'article de Ronquist [360] pour plus de précision.

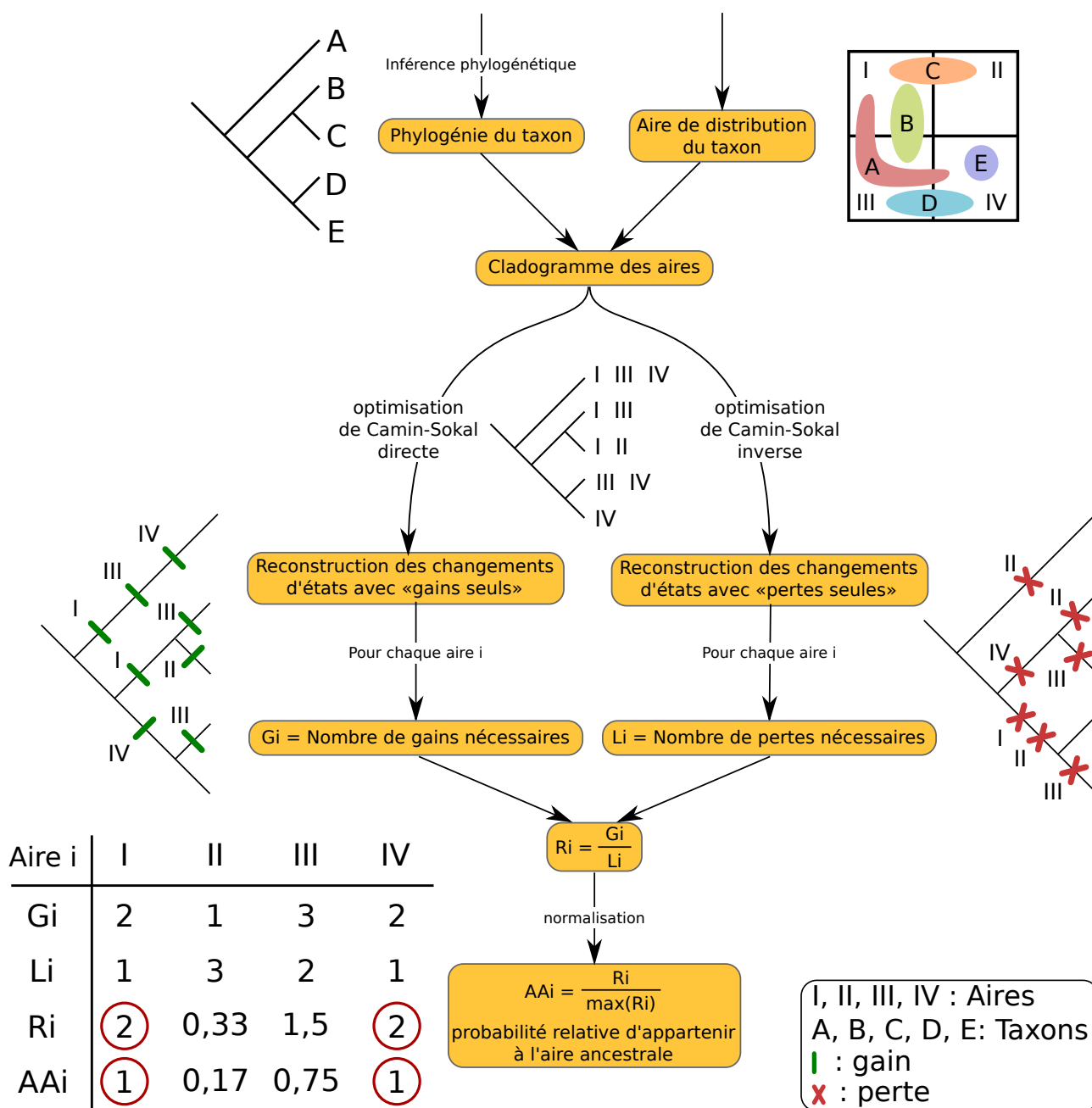


FIG. 1.7 : Déroulement illustré de l'inférence d'aire ancestrale par optimisation de Camin-Sokal. Figure réalisée d'après les explications de Crisci [100, pp. 43-45].

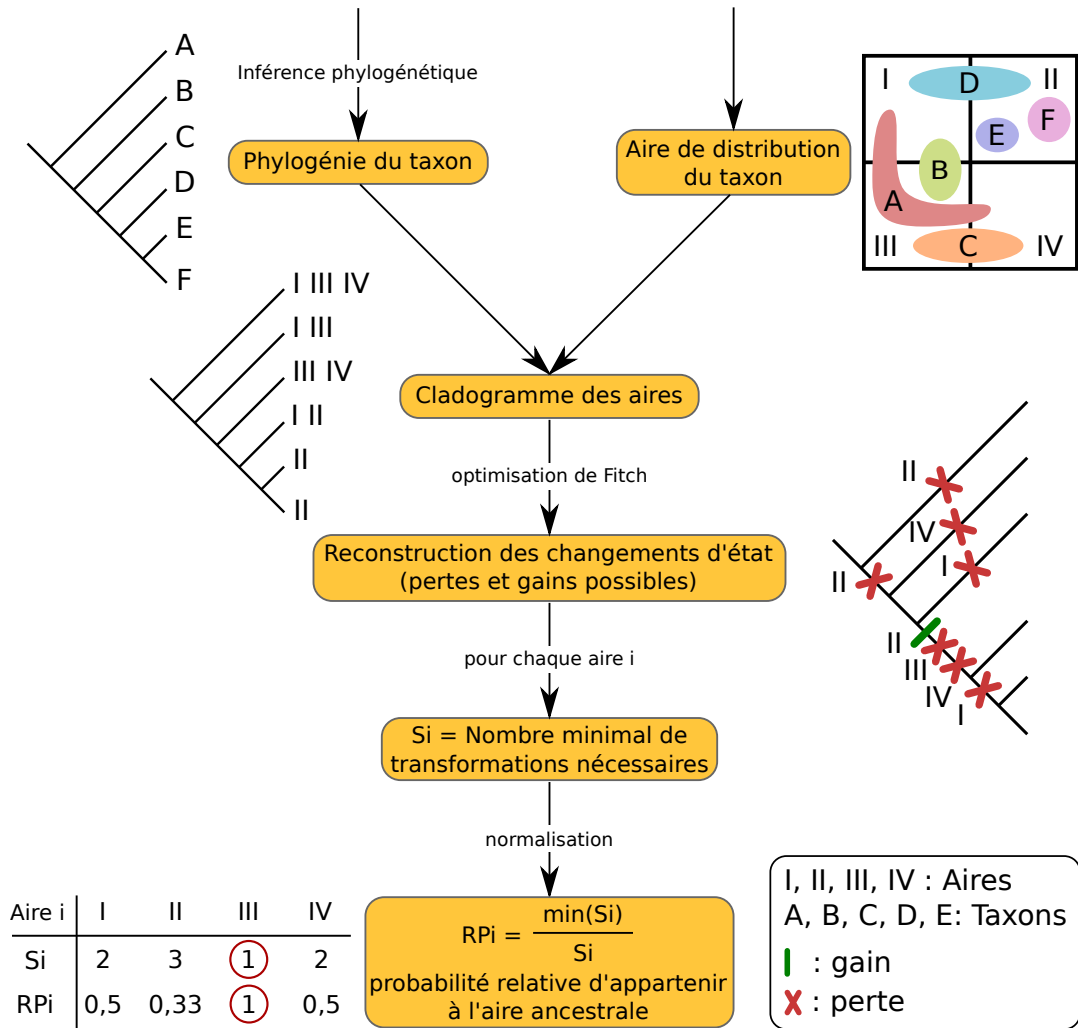


FIG. 1.8 : Déroulement illustré de l'inférence d'aire ancestrale par optimisation de Fitch. Figure réalisée d'après les explications de Crisci [100, pp. 45-47].

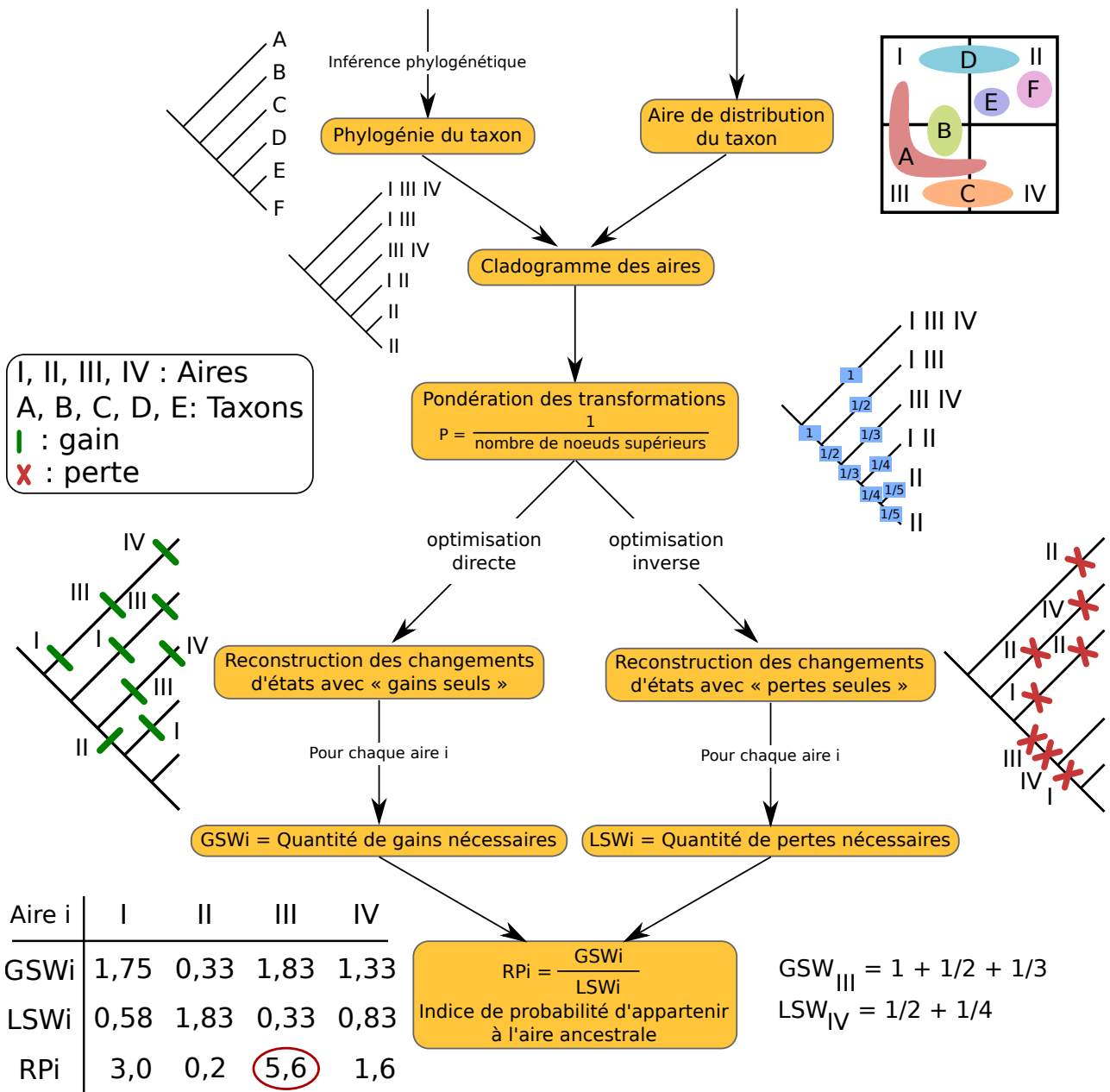


FIG. 1.9 : Déroulement illustré de l'inférence d'aire ancestrale par optimisation pondérée. Figure réalisée d'après les explications de Crisci [100, pp. 47-49].

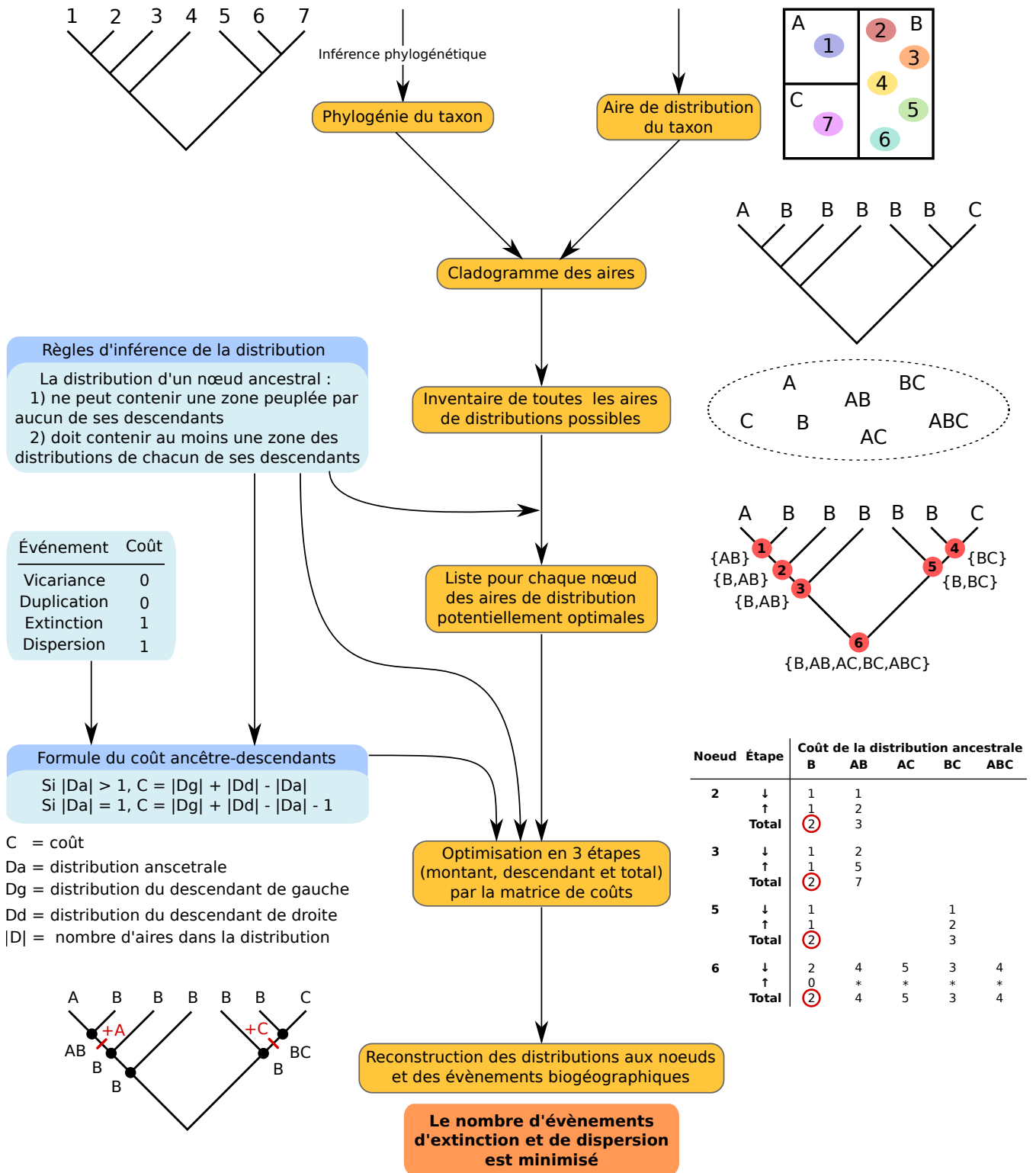


FIG. 1.10 : Déroulement illustré de l'analyse de dispersion–vicariance (DIVA). Figure réalisée d'après les travaux de Ronquist [360].

1.4 Conclusion du chapitre

J'ai retracé à travers ce chapitre le développement dans trois disciplines (la [linguistique historique](#), la [critique textuelle](#) et la [systématique biologique](#)) de méthodes permettant de réaliser une [inférence historique](#) à partir de la comparaison des structures formelles des objets auxquelles elles s'intéressent. Bien que certains outils et concepts soient étonnamment proches, ces développements se sont fait de manière singulière dans chaque discipline et les méthodes qui en sont issues restent substantiellement différentes.

J'ai aussi détaillé un certain nombre de ces méthodes dans le but de mieux comprendre les spécificités de chacune d'entre elles et dans l'espoir de favoriser l'intercompréhension entre des disciplines très éloignées, mais qui rencontrent sur le sujet de l'inférence historique des problématiques communes. Si la grande majorité de ces méthodes n'a pas eu d'application directe dans cette thèse, leur compréhension m'a permis de me familiariser et de me questionner sur les principes généraux de l'inférence historique et de m'ouvrir l'esprit quant à la diversité des approches possibles.

On aura pu remarquer que les méthodes issues de la systématique biologique ont été plus longuement détaillées. Ceci vient sans doute en partie de ce que je les connais mieux, de par ma formation. Mais cela vient aussi de ce que ces méthodes ont été plus fréquemment utilisées pour étudier l'évolution culturelle au cours des dernières décennies. Aussi leur bonne compréhension s'avérait nécessaire, tant pour moi que pour les personnes amenées à lire cette thèse, afin de comprendre le chapitre suivant. En effet, de nombreuses disciplines s'étant appropriées les outils de la systématique biologique pour étudier l'histoire des cultures humaines, il m'a paru intéressant d'enrichir notre étude des musiques du Gabon de ces expériences. Je présenterai donc dans le chapitre suivant¹¹⁸ une revue détaillée de ces utilisations des [méthodes issues de l'inférence phylogénétique](#) pour étudier des données culturelles.

118. ↑ Chapitre 2 pages 49 à 103.

Chapitre 2

Les méthodes issues de l'inférence phylogénétique appliquées aux données culturelles

2.1	État de l'art disciplinaire	53
2.1.1	Limites	53
2.1.2	Critique textuelle	53
2.1.2.1	État des lieux	53
	Les méthodes de distances	53
	Les distances utilisées	54
	La cladistique	55
	Autres méthodes	56
2.1.2.2	Controverses et difficultés	56
	Différences dans les mécanismes d'évolution	56
	Les ancêtres observés	56
	Les multifurcations	56
	Le traitement des homoplasie	57
	La contamination	57
	La co-incidence	57
	L'identification et le traitement des caractères	57
	Le choix des caractères	57
	La pondération des caractères	57
2.1.2.3	Évaluer la pertinence des méthodes	58
	Justifications théoriques	58
	Justifications empiriques indirectes	59
	Comparaison avec les résultats de la stemmatologie	59
	Robustesse de l'analyse	59
	Utilisation d'indices statistiques	60
	Validation empirique directe	60
	Validation sur données artificielles	60

	Utilisation de critères externes de validation	61
2.1.3	Linguistique	61
2.1.3.1	Premières approches	61
	Divergence et convergence de la biologie et de la linguistique	61
	Méthode en réseau de Bandelt	62
2.1.3.2	Les méthodes de caractères	62
	Méthodes cladistiques	62
	Méthodes de compatibilité	63
	Réseau par minimisation des transferts horizontaux (MLN)	63
2.1.3.3	Méthodes probabilistes.	64
	L'inférence bayésienne	64
	Les migrations des populations	64
2.1.4	Anthropologie de la culture matérielle	65
2.1.4.1	Premières approches	65
	L'inférence historique en anthropologie	65
	Des travaux précurseurs, mais sans suite	66
2.1.4.2	Tester des hypothèses historiques	66
	L'étude de Collard et Shennan	66
	Ethnogenèse ou phylogenèse ?	67
	Autres hypothèses testées	68
2.1.4.3	Élargissement des approches	68
	Inférer les relations	68
	Évaluer ces outils d'inférence	70
2.1.5	Musicologie	70
2.1.5.1	Les travaux de Temkin	70
2.1.5.2	Les travaux de Toussaint et coll.	71
	Présentation des analyses	71
	En visant la classification	72
	En visant la reconstruction évolutive	73
2.1.5.3	Les travaux de Le Bomin et coll.	73
2.1.6	Autres applications	73
	Concepts scientifiques	74
	Traditions orales	74
	Économie	75
2.2	Considérations générales sur ces transpositions interdisciplinaires	75
2.2.1	Spécificités disciplinaires	75
	Diversité des utilisations	75
	Particularismes	75
2.2.2	La question de la transmission horizontale	76
2.2.2.1	Influence de la transmission horizontale	76
	Sur la topologie inférée	76
	Sur l'estimation des dates de divergence	78
	Dans les méthodes comparatives phylogénétiquement informées	78
2.2.2.2	Détection et mesure de la transmission horizontale	79
	À partir des arbres phylogénétiques inférés	79
	À partir d'indices statistiques	79
	Indices de cohérence et de rétention	79
	Score δ et score Q-résiduel	80
	Par visualisation des données	80

		51
	Autres méthodes	81
2.3	Tableau comparatif de quelques études	81
2.3.1	Objectifs et limites de ce tableau	81
2.3.2	Présentation générale	82
2.3.2.1	Contenu du tableau	82
	Les lignes : des <i>analyses</i>	82
	Les colonnes : typologie des analyses	83
2.3.2.2	Choix des analyses	83
	1. Des méthodes issues de l'inférence phylogénétique	83
	2. Des données culturelles	84
	3. Des relations comme résultat	84
2.3.3	Description détaillée des champs du tableau	84
2.3.3.1	Objectifs des analyses	84
2.3.3.2	Données analysées	86
2.3.3.3	Typologie des méthodes utilisées	87
2.3.3.4	Éléments de caractérisation de ces méthodes	88
	Méthodes cladistiques	88
	Méthodes phénétiques	90
	Méthodes probabilistes	91
	Autres	91
2.3.3.5	Logiciels utilisés	92
2.3.4	Tableau comparatif	93
2.4	Conclusion du chapitre	103

Comme on l'a vu, les méthodes systématiques d'inférence phylogénétique se sont développées en biologie dans la deuxième moitié du XX^{ème} siècle. Le développement de l'informatique pendant cette même période a permis une automatisation de ces méthodes et sans doute facilité leur développement, permettant aux biologistes d'intégrer une quantité croissante de données dans leurs analyses. La fin du XX^{ème} siècle voit fleurir de nombreux algorithmes d'inférence phylogénétique ainsi que des logiciels qui les implémentent, ce qui les rend de plus en plus faciles d'utilisation.

Depuis le début des années 1990¹ les méthodes développées dans le cadre de l'inférence phylogénétique² sont utilisées pour analyser des données de nature culturelle. Cette transposition méthodologique tient sans doute à la relative facilité d'utilisation de ces méthodes d'une part et leur capacité à analyser un grand nombre de données d'autre part. La manière dont ces méthodes sont utilisées est cependant très différente d'une discipline à l'autre, voire entre différentes études au sein d'une même discipline.

Ce chapitre vise ainsi à faire un état de l'art concernant cette utilisation des méthodes issues de l'inférence phylogénétique pour analyser des données de nature culturelle. Cette recension a pour objectif de mieux comprendre les problématiques auxquelles ont été confrontées les différentes disciplines afin de m'inspirer de cette expérience pour l'étude des musiques du Gabon. Comme tout état de l'art, elle vise aussi à synthétiser les différents travaux effectués sur ce sujet afin d'en donner une vue d'ensemble et faire mieux connaître les recherches engagées, permettant un meilleur échange de savoir et donc un renforcement mutuel entre les disciplines³. Elle cherche enfin à établir une base de comparaison entre les différentes approches qui pourront permettre dans le futur de mieux comprendre et discuter les enjeux de la transposition de méthodes d'une discipline à l'autre⁴.

Je présente donc dans ce chapitre une revue de la littérature concernant l'application des méthodes issues de l'inférence phylogénétique dans différentes disciplines étudiant des objets culturels⁵. Je donne ensuite quelques considérations générales concernant la manière donc cette transposition méthodologique est effectuée⁶, en m'appuyant sur un tableau recensant plus d'une centaine de publications utilisant de telles méthodes⁷.

1. ↑ Même si les premières utilisations de méthodes issues de l'inférence phylogénétique en biologie sur des textes de manuscrits datent de la fin des années 1960 (voir partie 2.1.2).

2. ↑ Je groupe ces méthodes sous le terme de méthodes issues de l'inférence phylogénétique.

3. ↑ N'étant moi-même pas spécialiste des différentes disciplines concernées, cet état de l'art est nécessairement limité par mon peu de recul et de culture dans ces domaines.

4. ↑ Voir par exemple la partie 9.4 page 306.

5. ↑ Partie 2.1 page 53.

6. ↑ Partie 2.2 page 75.

7. ↑ Partie 2.3 page 81.

2.1 État de l'art disciplinaire

2.1.1 Limites

L'ensemble des études appliquant des méthodes issues de l'inférence phylogénétique à des données culturelles est trop vaste pour être étudié ici. Je me concentrerai donc sur les études utilisant des méthodes 1) permettant de tracer des relations entre des groupes, 2) ayant été développées à l'origine dans le cadre de la systématique biologique et 3) qui sont appliquées à l'analyse de données culturelles⁸. Cette revue ne saurait pourtant être exhaustive, même si j'ai tenté d'y inclure, dans la mesure du possible, un maximum d'études publiées sur le sujet. Je n'ai pu par ailleurs rentrer dans les détails des méthodes et des objectifs de tous ces travaux. Il s'agit donc d'en donner un aperçu global, qui pourrait servir de base dans le futur à une étude plus approfondie des spécificités disciplinaires ou au contraire des problématiques communes à différents domaines.

2.1.2 Critique textuelle

2.1.2.1 État des lieux

Pendant la deuxième moitié du XX^{ème} siècle, alors que les méthodes automatisées d'inférence phylogénétique se développent rapidement en biologie, l'attrait pour les méthodes **stemmatiques** diminue, probablement en lien avec le développement d'une pensée post-structuraliste qui remet en cause la possibilité de la reconstruction des textes anciens [353, pp. 104–106]. La conception de l'auteur-e unique d'un texte originel est remplacée par une vision plus construite, où le texte est la convergence d'influences multiples [353, p. 104]. L'image de l'arbre historique est remise en cause, pour lui préférer l'analogie du rhizome [353, p. 104]. Enfin la notion, essentielle chez Lachmann et Maas, d'*erreurs* présentes dans les manuscrits observés qu'il s'agirait d'identifier et de *corriger*, est très critiquée, entre autres pour sa connotation eugéniste [353, pp. 104–106]. Depuis la fin du XX^{ème} siècle, les méthodes développées par les biologistes sont de plus en plus utilisées en **critique textuelle**, notamment parce qu'elles permettent de traiter un plus grand nombre de données et de mieux expliciter les méthodes et les choix d'analyse.

Les méthodes de distances. La première utilisation d'une méthode issue de la biologie de l'évolution pour retrouver l'apparement de manuscrits est due à Griffith [174]. Il utilise dans son analyses de différents manuscrits d'une œuvre de Juvenal une méthode provenant de la **taxinomie numérique** de Sokal et Sneath [A389]⁹, basée sur la mesure, pour chaque couple d'objets comparés, d'une valeur de similarité globale entre ces objets. Après avoir identifié les endroits où le texte varie dans les 16 manuscrits étudiés, il compte pour chaque paire de manuscrits le nombre de similarités sur ces **lieux variants**, obtenant ainsi une **matrice de similarités**. Les

8. ↑ La partie 2.3.2.2 page 83 explique plus en détail les critères délimitant les analyses présentées ici.

9. ↑ Citée par Griffith [174, p. 114].

valeurs obtenues sont alors regroupées en 8 classes de similarité proche, chacune représentée par un symbole. Chaque élément de la matrice est alors remplacé par un des 8 symboles de classe. La matrice est ensuite réorganisée par tâtonnement afin d’obtenir un regroupement des symboles semblables (et donc des distances de même ordre de grandeur). Cette matrice lui permet alors de visualiser des lignes de clivages séparant différentes parties homogènes (du point de vue des symboles contenus) de la matrice et de discuter la proximité des différents manuscrits. Cette méthode permet donc, via une étape très subjective et non répétable, d’obtenir plusieurs groupes de manuscrits proches ordonnés selon une similarité décroissante.

Cette méthode, laborieuse, non reproductible et donnant des résultats limités ne sera pas du tout reprise en critique textuelle¹⁰, et comme on le verra¹¹, les critiques des textes préféreront utiliser la méthode **cladistique**. Il faudra attendre le début des années 2000 pour que des **méthodes de distances** soient de nouveau utilisées, profitant des développements de la **taxinomie numérique**. Macé et coll. [276] utilisent en particulier l’algorithme de **neighbor joining**¹² pour identifier les ressemblances au sein de chacun de deux ensembles de manuscrits¹⁴.

Les méthodes **phénétiqes** sont cependant critiquées (y compris en biologie) car l’utilisation d’une mesure de distance résumant les différences engendre une perte d’information importante, mieux prise en compte par les méthodes basées explicitement sur l’analyse des **états de caractères** et de leur évolution. Bien que les résultats du **neighbor joining** dans les études de critique textuelle comparant différentes méthodes d’inférence historique ne soient pas toujours mauvais¹⁵ [34, 361, 392], cette méthode est surtout utilisée afin de tester l’influence de certains paramètres sur la qualité de l’inférence [390, 393], car sa rapidité permet de répéter de nombreuses fois l’analyse en variant les paramètres¹⁶.

Par ailleurs la popularité croissante des **réseaux** phylogénétiques [209] au cours des années 2000 gagne aussi la critique textuelle, et plusieurs études s’approprient ces méthodes¹⁷, construisant pour la plupart des réseaux à partir d’une **matrice de distances**.

Les distances utilisées. Bien entendu, dans toutes ces analyses, le choix de la mesure de distance utilisée est crucial et a une influence déterminante sur l’inférence historique obtenue. La plupart des travaux utilise une mesure de similarité¹⁸ globale entre les deux textes, définie

10. ↑ Griffith [175] l’utilise cependant une nouvelle fois pour étudier une quinzaine de manuscrits des évangiles.

11. ↑ Voir paragraphe **La cladistique** page 55.

12. ↑ Cet algorithme proposé par Saitou et Nei [A365]¹³ permet de construire un arbre à partir d’une matrice de distances de manière très rapide, grâce à une approximation donnant d’assez bons résultats [136, pp. 166–168].

13. Citée par Felsenstein [136, pp. 166–168].

14. ↑ Regroupant respectivement des manuscrits issus d’un texte de Grégoire de Nazianze et de Basile le Minime.

15. ↑ Voir paragraphe **Validation sur données artificielles** page 60.

16. ↑ Voir aussi l’utilisation de Spencer et coll. [391].

17. ↑ Voir par exemple les travaux suivants : [33, 46, 49, 128, 361, 392, 395, 444]. Ces études explorent d’ailleurs la diversité des méthodes existantes, utilisant aussi bien la décomposition en bipartitions (*split decomposition*) [33, 361, 392], les **réseau de voisins** (*NeighborNet*) [128, 361, 444], les super-réseaux (*supernetworks*) [444] ou les réseaux médians réduits (*reduced median networks* [46, 395]).

18. ↑ Ou de distance, ce qui revient au même.

comme la proportion de mots communs dans les deux séquences¹⁹ [174, 276, 390, 392, 393]. Dans leur étude sur les relations de filiation entre des courriers transmis en chaînes, Bennett et coll. [43] mesurent la similarité entre les textes de deux courriers en leur appliquant en même temps un algorithme de compression²² : plus les textes se ressemblent, plus l'algorithme réussit à compresser l'ensemble des deux fichiers et plus le fichier produit sera de faible taille. Enfin, l'approche de Bordalejo [49] est par ailleurs tout à fait différente : ce n'est pas le contenu des différentes versions de textes des contes de Canterbury qui est comparé, mais leur ordre au sein des différents manuscrits. Elle utilise donc une mesure de distance qui permette de comparer l'ordre dans lequel sont présentés les contes dans chaque manuscrit et de prendre en compte l'ajout ou la suppression d'un conte dans un des manuscrits.

La cladistique. L'analogie entre méthodes *stemmatiques* et *cladistiques* a été remarquée depuis longtemps [339]. Cette proximité méthodologique conduira à l'utilisation de méthodes cladistiques dans le domaine de la critique textuelle. Ainsi Lee [252] utilise deux variantes des méthodes cladistiques pour retrouver les relations de filiations entre 10 manuscrits des *Quaestiones in Heptateuchum* de Saint Augustin. En 1992, Robinson et O'Hara [356] lancent à la communauté scientifique le défi de reconstruire l'histoire de 44 manuscrits du *Svipdagsmál* à partir du contenu textuel uniquement et sans indications contextuelles. Les résultats obtenus par Robinson et O'Hara [357] en utilisant la cladistique, excellents tant en rapidité qu'en qualité, vont susciter un attrait pour la méthode de Hennig, qui se répand dans le domaine de la critique textuelle²³.

Différentes variantes des méthodes cladistiques sont utilisées, comme dans l'étude de Macé et coll. [275] utilisant la stratocladistique, méthode issue de la paléontologie permettant d'intégrer des informations temporelles dans l'analyse, telles que l'âge des manuscrits ou des relations de parenté connues grâce à des données externes. La méthode cladistique comporte par ailleurs des limites importantes²⁴ et elle sera parfois modifiée et adaptée pour y remédier. Ainsi Carlson [78] cherche à permettre qu'un manuscrit observé soit placé à un *noeud interne* de l'arbre (et pas seulement comme *noeud externe*)²⁵ et qu'un manuscrit soit le père de plus de deux autres²⁶.

19. ↑ De nombreuses variantes existent, et diffèrent notamment par :

- les mots pris en comptes et ceux exclus dans le calcul : mots invariants, ponctuations, éléments jugés non significatifs, etc.²⁰ ;
- la pondération éventuelle de certains *lieux variants* [393]²¹ ;
- la correction de la distance (prise en compte des changements multiples de caractères [390], besoin d'une distance ultramétrique, etc.).

20. Voir paragraphe *Le choix des caractères* page 57.

21. Voir paragraphe *La pondération des caractères* page 57.

22. ↑ Cette compression est effectuée par le logiciel GenCompress. Cette méthode est aussi utilisée en biologie pour mesurer la similarité de séquences moléculaires [43, p. 78].

23. ↑ Notamment dans les études suivantes : [7, 233, 275, 354, 355, 357, 366, 367, 392, 394, 444, 447].

24. ↑ Certaines de ces limites sont détaillées dans la partie 2.1.2.2 page 56.

25. ↑ Voir paragraphe *Les ancêtres observés* page 56.

26. ↑ Voir paragraphe *Les multifurcations* page 56.

Autres méthodes. La plupart des autres méthodes d'inférence historique développées en biologie de l'évolution n'ont, à ma connaissance, pas été utilisées en critique textuelle. Les méthodes probabilistes, aujourd'hui très populaires en biologie, ont été presque complètement ignorées. Spencer et coll. [392, pp. 509–510] l'expliquent par le manque de connaissance des mécanismes d'évolution des textes, rendant difficile la proposition d'un modèle d'évolution, nécessaire à l'utilisation de ces méthodes. Cependant Macé et coll. [275] ont récemment proposé une application de la méthode bayésienne pour étudier des manuscrits du *Florilegium Coislinianum*, grâce une typologie minutieuse des variantes, permettant de proposer des hypothèses sur le sens et la probabilité des transformations lors de la copie.

2.1.2.2 Controverses et difficultés

L'application des méthodes phylogénétiques à la critique textuelle pose un certain nombre de difficultés théoriques et pratiques. Howe et coll. [203] en ont récemment fait une revue assez détaillée, et je ne ferai que présenter succinctement les principales d'entre elles.

Différences dans les mécanismes d'évolution.

Les ancêtres observés. En biologie, même si le statut de l'ancêtre²⁷ (et en particulier la relation entre les fossiles observés et les populations ancestrales) reste très débattu, les méthodes d'inférence phylogénétique n'envisagent l'ancêtre qu'hypothétique et l'ensemble des taxons étudiés sont positionnés sur les branches terminales. La question se pose différemment en critique textuelle, où ce sont bien des relations généalogiques entre des objets que nous recherchons et non plus des relations phylogénétiques²⁸. L'approche généralement utilisée est alors de considérer l'ancêtre comme un taxon terminal relié à un nœud interne par une branche de longueur nulle [49, 78] : par ce procédé l'ancêtre est en quelque sorte « assimilé » à ce nœud interne.

Les multifurcations. De même, les logiciels d'inférence phylogénétique visent à produire des arbres dont tous les embranchements sont dichotomiques, les multifurcations ne servant qu'à représenter une incertitude non résolue. Cette conception intégralement dichotomique reste débattue, et est justifiée soit par des arguments biologiques (structure de l'évolution, mécanismes de spéciation, etc.) soit par des arguments épistémologiques. En critique textuelle, un même exemplaire d'un manuscrit peut être à l'origine de multiples copies. Cette spécificité, par rapport à la biologie, encouragent les critiques des textes à s'interroger sur la manière de prendre en compte les multifurcations dans les analyses afin d'améliorer les inférences [326, pp. 60–61] [357, p. 9] [78, p. 13].

27. ↑ Voir par exemple les articles de Dayrat [112] et de Lecointre [250].

28. ↑ Pour Lecointre et Le Guyader [251, pp. 44–46], c'est justement le statut de l'ancêtre qui différencie un arbre généalogique d'un arbre phylogénétique. Voir aussi, à propos de la distinction entre ces deux types d'arbres, le travail de Lecointre [250, pp. 199–201].

Le traitement des homoplasie.

La contamination. La question de l'hybridation entre différentes lignées est un des points les plus fréquemment soulevés lors de l'utilisation de méthodes issues de l'inférence phylogénétiques pour étudier l'histoire évolutive d'objets culturels, et reste importante en biologie dans le cas d'espèces sujettes à l'hybridation ou au **transfert horizontal** de gènes. Dans les études de critique textuelle, où l'on parle plus volontiers de **contamination**²⁹, le problème est fréquemment évoqué³⁰. Plusieurs études proposent des méthodes pour détecter ces contaminations³¹, certaines comme celle de Woerther et Khonsari [447, p. 230] étant là encore inspirées des méthodes utilisées en biologie pour détecter les hybridations ou les transferts latéraux de gènes.

La co-incidence. Le problème de la **convergence** indépendante de certains caractères vers un même état dans différentes lignées est un problème aussi bien en biologie qu'en critique textuelle (où l'on parle plutôt de **co-incidence**), car elle tend à brouiller le signal phylogénétique. Si l'analyse statistique de nombreux caractères (en particulier via une méthode de **parcimonie**) cherche justement à minimiser l'effet de ce signal conflictuel, l'inférence devient néanmoins impossible lorsque celui-ci est trop important. La question est régulièrement abordée en critique textuelle³², il s'agit alors généralement d'exclure ou de pondérer certains caractères, selon leur vraisemblance à porter un signal phylogénétique. Cette pondération peut même être très fine. Ainsi Macé et coll. [275, p. 115] pondèrent chaque **variante** selon la vraisemblance qu'elle puisse être corrigée (irréversibilité), la probabilité qu'elle apparaisse deux fois indépendamment (irreproductibilité) et la longueur du texte varié.

L'identification et le traitement des caractères.

Le choix des caractères. Le choix des caractères, lié au problème des co-incidences, est évidemment crucial dans l'inférence phylogénétique. La problème de savoir si toutes les **variantes**, ou seules celles considérées comme des « erreurs », devraient être prises en compte est un vieux débat en stématique [326, pp. 61–62] [49, pp. 56–65]³³.

La pondération des caractères. Enfin la question de la pondération des caractères, qui consiste à donner *a priori* plus d'importance à certains caractères lors de l'inférence historique, reste un sujet de préoccupation en stématique [275, 354, 357, pp. 72–74]. Spencer et coll. [393]

29. ↑ Terme provenant de la critique textuelle classique où l'objectif était alors de rechercher un hypothétique texte originel « pur ».

30. ↑ Notamment dans les études suivantes : [356, pp. 335–336] [326, p. 57] [355, pp. 86–88] [357, p. 9] [49, p. 68] [78, p. 13] [203, pp. 56–58].

31. ↑ Notamment celles de Robinson [354], Spencer et coll. [394] et Woerther et Khonsari [447, p. 230].

32. ↑ Notamment par Robinson et O'Hara [356, pp. 335–336], Robinson [354], Howe et coll. [203] et O'Hara et Robinson [326, pp. 57–60].

33. ↑ On pourrait faire ici un parallèle avec la **systématique biologique** et le choix de Hennig de n'attester les filiations qu'à partir des caractères dérivés partagés (ou **synapomorphies**).

proposent de tester l'influence de la pondération des caractères sur la topologie de l'arbre obtenu. Les modifications sont classées en 11 types, auxquels il est donné un poids différent, favorisant les changements ayant le moins de chance de survenir de manière convergente. Trois arbres sont construits par *neighbor joining* soit avec cette pondération, soit avec une pondération uniforme, soit avec une pondération aléatoire. La comparaison entre ces arbres leur permet de conclure que le choix de pondérer certains types de transformations n'influe pas sur le résultat obtenu³⁴.

2.1.2.3 Évaluer la pertinence des méthodes

La transposition de nombreuses méthodes issues de la biologie vers la critique textuelle suscite de nombreux problèmes théoriques et pratiques. De plus le manque de recul et la diversité des approches utilisées rendent difficile l'appréciation de l'utilité et de la qualité de ces nouveaux outils. Plusieurs stratégies existent pour évaluer ces méthodes, et si celles-ci sont généralement proposées comme validation ponctuelle d'une analyse, certains articles comme celui de Spencer et coll. [390] cherchent explicitement à développer des moyens d'évaluation fiables des méthodes. On peut distinguer deux types d'arguments utilisés pour évaluer la qualité de ces méthodes : théoriques et empiriques, ces derniers pouvant être soit directs soit indirects.

Justifications théoriques. Comme on l'a vu l'application des *méthodes issues de l'inférence phylogénétique* en critique textuelle suscite un débat. Un grand nombre d'arguments pour discuter la pertinence de cette transposition interdisciplinaire se base sur des considérations purement théoriques. Plusieurs articles, délaissant plus ou moins l'approche proprement empirique, sont d'ailleurs consacrés à cet aspect³⁵. Parmi les arguments théoriques utilisés pour justifier l'application des méthodes issues de l'inférence phylogénétique, les plus fréquents³⁶ sont l'analogie avec la biologie^{37,38} et le succès de l'utilisation de ces méthodes dans d'autres disciplines.

Cependant cette approche ne peut être qu'une heuristique, intéressante dans un premier temps pour justifier l'essai d'une méthode. Les arguments théoriques ne peuvent, *in fine*, suffirent à établir la valeur scientifique d'une approche⁴⁰. Il s'agira alors de chercher des arguments empiriques permettant d'évaluer la pertinence de l'application de la méthode à nos données. Parmi ces arguments empiriques, je sépare un peu artificiellement deux types d'arguments (que je qualifie d'« indirect » ou « direct ») selon la nature et la confiance que l'on peut avoir dans les éléments qui permettent d'évaluer la méthode.

34. ↑ Mais la méthode statistique pour tester cela est contestable : la distance entre les arbres obtenus avec et sans pondération est montrée statistiquement inférieure à celle obtenue entre deux arbres obtenus aléatoirement. Les données étant les mêmes (et les méthodes extrêmement proches), le contraire eut été plus qu'étonnant.

35. ↑ Notamment les travaux suivant : [80, 181, 203, 267, 326, 339].

36. ↑ Que l'on pourrait qualifier, suivant une proposition de Pierre Darlu, de *justifications analogiques*.

37. ↑ Que ce soit au niveau de la structure des données servant à l'analyse, des mécanismes évolutif impliqués, ou des méthodes utilisées.

38. ↑ Voir par exemple les travaux de Don Cameron [A125]³⁹ et Macé et coll. [276, p. 256].

39. Citée par O'Hara et Robinson [326, p. 71].

40. ↑ Cet aspect est détaillé un peu plus dans le chapitre 9.4 et en particulier dans la partie 9.4.1.1 page 307.

Justifications empiriques indirectes.

Comparaison avec les résultats de la stémattologie. La pertinence de l'approche peut-être justifiée par le fait que les résultats obtenus soient proches de ceux proposés dans les études précédentes, en particulier ceux issus des méthodes *stémattiques*. Cet argument est sans doute le plus simple à obtenir car il ne nécessite pas d'analyse supplémentaire et il est utilisé dans de nombreuses études empiriques⁴¹. La justification est cependant faible, et repose complètement sur les résultats auxquels on se réfère. Comme le remarque Maas [274, p. 34] :

« The fact that a phylogenetic analysis of variant readings results in a diagram of the transmission which resembles a “manually” created stemma therefore does not “prove” (and, in fact, not even indicate) that this stemma is the best possible representation of the transmission history. »⁴² Maas (2010) [274, p. 34]

Les différences entre les résultats produits par la méthode stémattique et ceux produits par une *méthode issue de l'inférence phylogénétique* sont d'ailleurs difficiles à interpréter en l'absence d'argument externe permettant de savoir laquelle des deux approches est la plus adéquate. Cette forme de justification, même si elle peut servir à convaincre la communauté scientifique de la validité de l'approche en reproduisant des résultats établis, ne peut donc être suffisante à elle seule pour établir solidement la pertinence scientifique de la démarche. En fait la puissance de cette validation dépend directement du crédit que l'on accorde à la méthode ayant permis d'obtenir les résultats précédents auxquels sont comparés les nouveaux (ici la méthode stémattique « classique »).

Robustesse de l'analyse. Une autre manière de justifier la pertinence d'une méthode est d'évaluer sa *robustesse*, c'est-à-dire de vérifier que la méthode donne des résultats semblables lorsque l'on modifie les paramètres ou les données de l'analyse⁴³. C'est ainsi que Griffith [174] évalue la qualité de sa méthode : l'analyse est répétée sur plusieurs portions différentes du texte, puis des analyses parallèles sont effectuées à partir de chacune de ces portions. Les résultats de ces différentes analyses se révèlent statistiquement corrélés ce qui montre la capacité de la méthode à produire des résultats cohérents à partir de données différentes mais supposées porter la même histoire.

Par un procédé similaire, Spencer et coll. [390] vérifient la constance des résultats de l'analyse lorsque la taille de l'échantillon varie : l'analyse par *neighbor joining* est répétée en ne prenant qu'une partie du jeu de données (de la moitié au 1/32^{ème}) et les différents arbres obtenus sont

41. ↑ Voir par exemple les études suivantes : [128, 252, 273–276, 356, 357, 361, 391, 444].

42. ↑ « Le fait qu'une analyse phylogénétique de témoins variants produise un diagramme de transmission qui ressemble à un stemma produit “manuellement” ne “prouve” donc pas (et en fait n'indique même pas) que ce stemma est la meilleure représentation possible de l'histoire de la transmission. »

43. ↑ La comparaison entre les résultats de différentes méthodes proches (comme les méthodes *stémattiques* et *cladistiques* [275, p. 126]) évoquée ci-dessus⁴⁴ peut d'ailleurs être considérée comme une forme de justification par la robustesse.

44. Voir paragraphe *Comparaison avec les résultats de la stémattologie* page 59.

comparés entre eux. L'analyse de *bootstrap*, une méthode permettant de tester l'influence d'une perturbation du jeu de données sur les résultats⁴⁵, est elle aussi fréquemment utilisée en complément de l'analyse cladistique⁴⁶. Certes la robustesse est une qualité indispensable pour une méthode inférentielle. Cependant elle n'est pas suffisante et ne constitue pas une justification empirique satisfaisante. En effet les résultats produits par une méthode peuvent être à la fois robustes et faux.

Utilisation d'indices statistiques. Des indices calculés à partir des données ou des résultats peuvent aussi être utilisés pour évaluer la pertinence d'une analyse. Ainsi les indices statistiques développés en biologie (*indice de cohérence*, de *rétenion*, ...) sont parfois repris. Ainsi Spencer et coll. [390] proposent d'utiliser l'indice de cohérence réajusté⁴⁷ ainsi que les valeurs de *bootstrap* pour évaluer la fiabilité de leur méthode. Cependant le lien entre de tels indices et la pertinence de l'outil utilisé est indirecte : leurs valeurs sont affectées par de nombreux paramètres et peuvent être interprétées de manières très diverses⁴⁸.

Validation empirique directe.

Validation sur données artificielles. Une autre approche consiste à créer un jeu de données artificiel^{49,50} : on définit des règles régissant la manière dont un texte est recopié et modifié, puis on simule l'évolution de lignées de manuscrits descendant les uns des autres en suivant ces règles. L'ensemble de manuscrits ainsi obtenu, ainsi que les relations (connues) entre eux, est appelé une *tradition artificielle*⁵¹ de copies à partir de la version anglaise de *Perceval*. Des sous-jeux de données, obtenus en enlevant une partie des manuscrits de la tradition artificielle, sont définis aléatoirement, pour simuler le fait qu'on ne retrouve pas tous les manuscrits et qu'il en manque donc certains lorsque l'on cherche à retrouver les relations de parentés entre les textes. À partir de chacun des sous-jeux de données, les relations entre manuscrits sont inférées, permettant de tester la sensibilité de la méthode utilisée à la quantité de manuscrits manquants.

45. ↑ Cette méthode est décrite plus en détail dans l'entrée *bootstrap* du glossaire.

46. ↑ Notamment dans les études suivantes : [7, 275, 390, 444, 447].

47. ↑ *Rescaled consistency index*, en anglais.

48. ↑ Voir le paragraphe *Indices de cohérence et de rétenion* page 79 pour un exemple concernant les multiples interprétations possibles des indices de cohérence et de rétenion.

49. ↑ Cette approche est utilisée depuis longtemps en critique textuelle pour tester d'autres méthodes d'inférence historique, par exemple par Peri [A337] ou Tombeur et coll. [A418].

50. ↑ Ce type de simulation, que l'on pourrait qualifier d'analogique, fait penser à la récente étude de Schillinger et coll. [373] concernant la culture matérielle, décrite dans le paragraphe *Évaluer ces outils d'inférence* page 70.

51. ↑ Une *tradition* de manuscrits peut-être créée artificiellement par une expérience de *transmission en chaîne*. On cherche ensuite à inférer, à partir de tout ou partie des données produites par simulation, l'histoire évolutive de la tradition de manuscrits. Comme on connaît la filiation réelle des manuscrits produits au cours de la simulation, on peut alors comparer avec celle-ci les résultats obtenus par inférence et ainsi évaluer la fiabilité de la méthode d'inférence. On peut relever les travaux de Spencer et coll. [392], Baret et coll. [34] et Roos et Heikkilä [361] utilisant cette approche, la tradition artificielle étant créée grâce à des scribes volontaires recopiant les textes. Cela permet de comparer les mérites de différentes méthodes [34, 361].

De telles données artificielles peuvent aussi être utilisées pour tester l'influence de certaines caractéristiques du jeu de données sur la qualité de l'inférence. Ainsi Spencer et coll. [392] créent une *tradition* artificielle.

Réussir à créer par simulation numérique⁵² des traditions artificielles pourrait permettre d'accélérer la procédure et donc de multiplier ces approches. Cependant il faut d'abord définir un modèle d'évolution des manuscrits, c'est-à-dire un système de règles que le logiciel peut utiliser pour simuler les modifications intervenant lors de la copie d'un manuscrit. L'absence d'un tel modèle d'évolution empêche donc pour le moment l'utilisation de données simulées [392, p. 509].

Cette méthode d'évaluation repose cependant sur la ressemblance supposée des mécanismes d'évolution des manuscrits dans les traditions réelles et dans les traditions artificielles, hypothèse très difficile à vérifier⁵³. La nécessité de pouvoir tester le modèle d'évolution deviendrait d'ailleurs encore plus importante si l'on devait utiliser des données simulées numériquement.

Utilisation de critères externes de validation. Enfin une dernière approche consiste à comparer les résultats avec des données externes attestant certaines relations de filiation entre manuscrits. On peut ainsi évaluer si la méthode est capable d'inférer des relations établies de manière indépendante. Dans l'idéal, il faudrait bien sûr connaître de manière quasi-certaine l'ensemble de relations de filiations entre les manuscrits d'une *tradition*, mais ce cas est très improbable car les informations externes concernant les relations d'apparentement entre manuscrits sont rares. Des indices forts et ponctuels peuvent cependant être retrouvés (par exemple des notes en marge exprimant explicitement la filiation), lesquels ont été plusieurs fois utilisés pour attester la qualité des inférences produites⁵⁴. Cette dernière approche consistant à corroborer les résultats issus de données indépendantes est sans doute la meilleure⁵⁵. La rareté des critères externes permettant de vérifier les résultats fait cependant qu'une telle validation n'est pas souvent possible.

2.1.3 Linguistique

2.1.3.1 Premières approches

Divergence et convergence de la biologie et de la linguistique. L'analogie des méthodes développées en systématique biologique et en linguistique a été remarquée depuis longtemps. Ainsi Platnick et Don Cameron [339], en constatant les similarités conceptuelles et méthodologiques entre les méthodes *cladistiques*, les *méthodes comparatives* utilisées en linguistique historique et les méthodes *stemmatiques*, suggère que ces méthodes sont trois expressions d'une méthode générale d'inférence historique, applicable dès lors qu'existe une relation parent–descendant. O'Hara

52. ↑ Je qualifie de *numérique* la simulation informatique de l'évolution de manuscrits, par opposition à la simulation *analogique* où l'évolution n'est pas reproduite pas ordinateur mais, par exemple, par des personnes jouant le rôle de scribes.

53. ↑ Et vraisemblablement en partie fausse : une partie des facteurs évolutifs intervenant sur de très longues durées sont difficiles à reproduire, comme la *co-incidence* de variants dans des lignées distinctes due à l'évolution globale de la langue utilisée.

54. ↑ Notamment dans les études suivantes : [233, 276, 356, 357, 361].

55. ↑ La validation sur des données artificielles pourrait être tout aussi pertinente. Cependant l'absence de critère fiable pour évaluer la qualité des données créées artificiellement l'affaiblit. J'estime donc que l'utilisation de critères externes de validation est la stratégie de validation d'une méthode la plus fiable parmi celles présentées ici.

[325] montre quant à lui l'essor parallèle du concept d'arbre dans ces trois même disciplines, en partie lié à ces développements méthodologiques.

Cependant, la fin du XX^{ème} siècle est marquée en biologie par une « rupture computationnelle » qui n'a pas d'équivalent en linguistique historique : de nombreuses méthodes d'inférence historique sont formulées de manière systématique et implémentées sous forme d'algorithme⁵⁶ dans des logiciels spécialisés, facilitant l'utilisation de ces méthodes et permettant l'analyse d'une quantité croissante de données. Dès la fin des années 90, cela conduira certain-e-s linguistes à utiliser ces méthodes sur leurs propres données. Une abondante littérature verra le jour dans cette continuité dans les décennies suivantes⁵⁷.

Méthode en réseau de Bandelt. La première⁵⁸ de ces études est peut-être celle de Forster et coll. [146]. Les auteurs s'inspirent d'une méthode en réseau⁶⁰, afin d'étudier les relations entre 17 langues romanes des Alpes. Ils réutiliseront cette méthode quelques années plus tard pour étudier les relations entre 14 langues d'Europe de l'ouest [145].

2.1.3.2 Les méthodes de caractères

Méthodes cladistiques. Les linguistes ne manquent pas de pointer les problèmes que posent les emprunts linguistiques lors de l'utilisation des méthodes issues de l'inférence phylogénétique. Pourtant cette méthode *en réseau* sera peu réutilisée par la suite et ce sont les méthodes représentant les relations sous forme d'arbre qui seront privilégiées au cours des années 2000. Gray et Jordan [169] infèrent les relations entre 77 langues austronésiennes par une méthode cladistique implémentée dans le logiciel Paup. L'arbre ainsi obtenu sert alors à inférer l'emplacement géographique. Pour cela, la longitude et la latitude (décrivant les coordonnées géographiques d'un taxon) sont considérées comme deux caractères de ce taxon : on peut donc inférer les valeurs de ces caractères aux nœuds internes et donc inférer la position géographique correspondant à ces nœuds. Quelques années plus tard, Holden [196] propose une analyse cladistique des langues bantoues, qui lui permet de discuter les routes migratoires qui ont pu être empruntées dans la partie sud de l'Afrique. Rexová et coll. [347] utilise aussi une méthode cladistique pour inférer les relations entre 84 langues indo-européennes, tout comme Fortunato et Mace [147] quelques années plus tard.

Ces méthodes cladistiques seront affinées dans les années suivantes, en particulier en ce qui concerne le type de caractères utilisés et la manière de les coder. Ainsi Dunn et coll. [127] codent

56. ↑ Certaines de ces méthodes sont présentées dans la partie 1.3 pages 24 à 47.

57. ↑ On pourra trouver dans l'article de Darlu [106] une discussion plus approfondie des différentes approches utilisées et des problèmes méthodologiques associés.

58. ↑ L'étude de Nerbonne et coll. [A307]⁵⁹ utilise une méthode très proche de l'*UPGMA*, mais sans y faire explicitement référence. Comme le soulignent les auteurs, cette méthode fait appel à des techniques classiques d'analyse de données. Comme expliqué ci-dessus, je ne m'attarde pas sur de telles études car elles n'empruntent pas explicitement une méthode développée dans le cadre de la systématique biologique.

59. Citée par Darlu [106, p. 65].

60. ↑ Cette méthode en réseau a été proposée quelques années plus tôt par la même équipe [A31]⁶¹ pour inférer des relations phylogénétiques non arborescentes à partir de l'étude de l'ADN.

61. Citée par Forster et coll. [146, p. 174].

des caractéristiques structurelles des langues (systèmes phonologiques et grammaticaux) plutôt que lexicales afin de retracer les relations entre 16 langues austronésiennes d'une part et entre 15 langues papoues d'autre part. Ben Hamed et coll. [A42]⁶² s'intéressent aux relations entre des langues indo-européennes d'une part et à celles entre des langues ouraliques-altaïques d'autre part en se basant sur la description des systèmes vocaliques de ces langues. Gaillard-Corvaglia et coll. [A155]⁶³ proposent un codage très fin des cognats des langues du nord de la France et du sud de l'Italie. Les modifications phonologiques déduites de l'analyse linguistique servent à élaborer un modèle d'évolution des caractères. Celui-ci sert alors à contraindre l'analyse cladistique : chaque transformation de caractère doit être compatible avec ce modèle, c'est-à-dire avec les données issues de l'expertise linguistique [106, pp. 67–68].

Les méthodes probabilistes⁶⁴ se développent progressivement au cours des années 2000 et plusieurs études combinent une analyse cladistique et une analyse **probabiliste** du même jeu de données, que ce soit pour comparer les performances de chaque méthode ou pour conforter les résultats obtenus. Ainsi Holden et coll. [195] comparent les résultats de ces deux types d'analyse pour essayer de comprendre l'histoire et la migration des populations de langue bantoue dans la partie sud de l'Afrique. Rexová et coll. [346] s'appuient de même sur cette combinaison de méthodes pour discuter l'histoire de ces langues, mais en intégrant aussi (comme dans leur précédente étude) des données grammaticales. De la même manière, dans sa thèse de doctorat, Jordan [214, pp. 99–112] construit un arbre issu d'une analyse cladistique et un autre issu d'une analyse bayésienne qui lui serviront de point de départ pour différentes analyses de langues austronésiennes.

Méthodes de compatibilité. Parallèlement, Ringe et coll. [349] s'emparent de la **méthode de compatibilité**⁶⁵ pour retracer les relations entre 24 langues indo-européennes encore en usage ou éteintes. Une autre spécificité de cette étude est qu'elle n'utilise pas seulement des données lexicales : elle inclut aussi des données morphologiques et phonologiques. La méthode de compatibilité nécessite cependant d'exclure tout **emprunt**, pourtant relativement fréquent dans les données linguistiques. Cette équipe développe donc par la suite une méthode pour construire des réseaux phylogénétiques parfaits⁶⁶ s'inspirant des phylogénies parfaites des méthodes de compatibilité et permettant de représenter les emprunts [132, 300, 301]⁶⁷.

Réseau par minimisation des transferts horizontaux (MLN). Afin de mieux prendre en compte les effets de l'emprunt, Nelson-Sathi et coll. [306] proposent d'utiliser la méthode de réseau par minimisation des transferts horizontaux⁶⁸ [A104]⁶⁹ qui permet de construire un

62. ↑ Citée par Darlu [106, p. 67].

63. ↑ Citée par Darlu [106, pp. 67–68].

64. ↑ Voir partie 2.1.3.3 page 64.

65. ↑ Voir la partie 1.3.5 page 31 pour une description de cette méthode.

66. ↑ *Perfect phylogenetic network*, en anglais.

67. ↑ Citées par Darlu [106, pp. 66–67].

68. ↑ *Minimal lateral network (MNL)*, en anglais.

69. ↑ Citée par List et coll. [264, p. 45].

réseau minimisant le nombre d'événements évolutifs et se rapprochant en cela des méthodes **cladistiques**. Un arbre des langues européennes est construit par inférence bayésienne, à partir duquel on recherche le nombre minimal d'événements de transferts horizontaux de vocabulaire entre branches permettant d'expliquer les données. Cette approche est poursuivie, dans leur étude des langues indo-européennes, par List et coll. [264] qui modifient légèrement la méthode de minimisation des transferts horizontaux.

2.1.3.3 Méthodes probabilistes.

L'inférence bayésienne. En 2003, Gray et Atkinson [168]⁷⁰ publient la première d'une longue série d'études basées sur l'analyse **bayésienne** de données linguistiques. À partir d'ensembles de cognats, codés selon leur présence ou leur absence, les auteurs proposent un arbre des langues indo-européennes issu du consensus de 1000 arbres échantillonnés par **inférence bayésienne**. Ils cherchent ensuite à trancher entre deux hypothèses concernant l'origine géographique de ces langues. Ces deux hypothèses impliquant des chronologies différentes, il s'attachent pour cela à estimer l'âge de la racine de cet **arbre de consensus**. Pour tester la robustesse des résultats, cette équipe répète par la suite cette analyse en utilisant d'autres modèles d'évolution, mais aussi en utilisant un jeu de données indépendant⁷¹ [28]. En plus de ces deux analyses, Atkinson [26, pp. 6.1–6.27] présente dans sa thèse une étude similaire s'appuyant sur une méthode bayésienne. Celle-ci lui permet là encore d'inférer les relations entre 35 langues mayas ainsi que la date de divergence entre ces langues (c'est-à-dire l'âge de la racine de l'arbre obtenu). De plus, il utilise l'ensemble d'arbres produit par l'analyse bayésienne ainsi que les distributions actuelles des langues pour inférer l'emplacement géographique du foyer d'origine de ces langues⁷².

L'inférence bayésienne des relations entre les langues devient bientôt un outil pour étudier l'évolution des langues et l'histoire des populations. Ainsi Pagel et Meade [332] construisent un arbre des langues indo-européennes et un autres des langues bantoues et s'en servent pour étudier la vitesse de transformation du vocabulaire dans ces deux familles. En l'absence de données génétiques, les arbres de langues construits par des **méthodes issues de l'inférence phylogénétique** sont utilisés comme modèle de l'évolution des populations. En appliquant une **méthode comparative phylogénétiquement informée**, Jordan et Mace [A213] utilisent par exemple un arbre des relations entre langues austronésiennes (obtenu par inférence bayésienne) pour tester s'il existe une association entre le **système de filiation** et les règles de résidence conjugale⁷³.

Les migrations des populations. Depuis la fin des années 2000, les méthodes bayésiennes ne servent plus seulement à inférer les relations historiques entre les populations (en d'autres termes

70. ↑ Ces auteurs publient en 2006 une explication plus détaillée de cette analyse [25].

71. ↑ Ce jeu de données est issu de l'analyse de Ringe et coll. [349] présentée dans le paragraphe *Méthodes de compatibilité* page 63.

72. ↑ Cette analyse s'inscrit un peu dans la même optique que l'analyse présentée dans le chapitre 5 pages 155 à 185, bien que la méthode utilisée soit légèrement différente.

73. ↑ Les règles de résidence conjugale sont les règles sociales gouvernant le lieu de résidence du couple marié. Dans le cas de la matrilocalité par exemple, le couple s'installe dans le lieu d'origine de l'épouse.

la **topologie** de l'arbre⁷⁴), mais sont utilisées pour étudier d'autres éléments de leur histoire. Ainsi, grâce à une calibration à partir de données archéologiques et historiques, Gray et coll. [166] infèrent non seulement la topologie, mais aussi l'âge de la racine de leur arbre des langues austronésiennes. En faisant appel à un modèle assoupli d'horloge linguistique, Kitchen et coll. [225] estiment l'âge de la racine de leur arbre des langues sémitiques afin de tester différents scénarios évolutifs concernant l'histoire des populations. Lee et Hasegawa [253] utilisent une approche similaire pour étudier les langues japonaises.

Ces travaux montrent que l'étude des dates de divergence par inférence bayésienne est une manière d'aborder l'histoire des populations et en particulier de leurs migrations. Plusieurs études ont cherché récemment à inférer l'histoire des populations en intégrant directement une dimension spatiale. Ainsi Walker et Ribeiro [434] utilisent un modèle de marche aléatoire assouplie⁷⁵ pour inférer l'origine géographique de la dispersion des langues arawaks d'Amérique du Sud. Bouckaert et coll. [54] modifient ce modèle pour le rendre plus réaliste et infèrent l'origine géographique probable de l'arbre des langues indo-européennes. Celle-ci favorise l'hypothèse d'une origine anatolienne de ces langues. Une étude récente de Chang et coll. [A84]⁷⁶ montre cependant (là aussi via une méthode bayésienne) que le choix de la position des langues éteintes impacte l'âge et le lieu d'origine inférés pour l'indo-européen⁷⁶. Par une analyse bayésienne de langues aïnoues⁷⁷, Lee et Hasegawa [254] infèrent les relations entre ces langues, la position géographique de la racine, mais aussi celle de tous les nœuds intermédiaires de l'arbre. En utilisant une méthode très proche de celle décrite et utilisée dans le chapitre 5⁷⁸, Currie et coll. [103] infèrent les relations entre plus de 500 dialectes bantous et s'en servent pour inférer les positions ancestrales d'un certain nombre de nœuds majeurs de l'arbre obtenu, leur permettant de discuter des migrations des populations parlant ces langues. Cette étude de l'histoire géographique des langues bantoues est poursuivie par Grollemund et coll. [179] grâce à une approche similaire et par de Schryver et coll. [379] qui infèrent les relations entre langues bantoues par une méthode bayésienne et les utilisent pour discuter des routes migratoires probables de ces populations.

2.1.4 Anthropologie de la culture matérielle

2.1.4.1 Premières approches

L'inférence historique en anthropologie. Je n'ai pas parlé du passé de l'inférence historique en anthropologie comme j'ai pu le faire pour la linguistique historique et la critique textuelle dans le chapitre 1⁷⁹. García Rivero [156, pp. 58–64] en propose une revue rapide, dont l'angle d'approche est justement celui de l'application des méthodes issues de l'inférence phylogénétique

74. ↑ En effet dans les **méthodes issues de l'inférence phylogénétique**, la topologie d'un arbre (c'est-à-dire les positions relatives des différents groupes définies par cet arbre) modélise généralement les relations de descendance entre ces groupes.

75. ↑ *Relaxed random walk (RRW)*, en anglais.

76. ↑ Pierre Darlu, communication personnelle.

77. ↑ Les langues aïnoues sont parlées au nord du Japon et à l'extrême est de la Russie.

78. ↑ Pages 155–185.

79. ↑ Pages 18–24.

(en particulier de la cladistique) aux données de l’anthropologie culturelle. Au cours du XX^{ème} siècle, cette discipline s’est ainsi attachée à trouver les liens entre les **artéfacts** matériels des cultures passées (par exemple les industries lithiques) par sériation (c’est-à-dire la mise en évidence de séries continues entre les artefacts). Ces séries, pouvant être soit linéaires (anagenèse), soit **arborescentes** (cladogenèse), soit réticulées, sont le plus souvent construites à partir d’une estimation de la similarité globale entre les différents objets [156, p. 62].

Des travaux précurseurs, mais sans suite. D’après García Rivero [156, p. 63], les premières utilisations de **méthodes issues de l’inférence phylogénétique** sont l’œuvre de Chapa au début des années 1980. Elle étudie des sculptures zoomorphes dans la péninsule ibérique [A85]⁸⁰ ainsi que des fibules⁸¹ de la période historique de La Tène [A86]⁸⁰ à l’aide de méthodes **cladistiques**. Cependant ces analyses resteront confidentielles et sans suite [156, p. 63]. Goebel et coll. [160] publient en 1991 une analyse des artéfacts du complexe amérindien du Nenana dans laquelle les objets étudiés sont regroupés en 14 classes dont la similarité deux à deux est ensuite calculée. Ces ressemblances leur servent à construire un arbre (un dendrogramme) représentant les proximités des principales classes par une analyse de regroupement⁸² des centres de gravité⁸³ [160, pp. 63–64, pp. 71–72]. Ceci conduira Buchanan et Collard [70, p. 1684] à qualifier de **phénétique** la méthode de Goebel et coll. mais cette méthode de regroupement, bien que éventuellement utilisée en biologie, a été développée à des fins d’analyse de données et pas spécifiquement pour l’inférence historique.

2.1.4.2 Tester des hypothèses historiques

L’étude de Collard et Shennan. C’est à partir du début des années 2000 que l’utilisation de méthodes issues de l’inférence phylogénétique, en particulier les méthodes **cladistiques**, vont se développer en archéologie et s’étendre plus généralement à l’étude des éléments matériels produits par les cultures. La première étude de ce courant est peut-être celle de Collard et Shennan [93]. Ils étudient un ensemble de poteries de l’époque néolithique issues de fouilles archéologiques en Allemagne afin d’évaluer l’importance relative des mécanismes de **phylogenèse** et d’**ethnogenèse**⁸⁴ dans l’évolution de la culture. Ils s’intéressent pour cela à 4 sites pour lesquels des poteries sont disponibles pour chacune des 10 mêmes phases archéologiques successives. Leur idée est la suivante : si le mécanisme de phylogenèse est prédominant, les regroupements par similarité des 4 sites devraient être identiques pour chacune des 10 phases. Ils effectuent pour chacune des 10 phases une analyse cladistique des 4 sites, caractérisés par les motifs représentés sur les poteries qui en sont issus, puis calculent les valeurs de **bootstrap** associées à chaque groupement. Dans la plupart des cas (6 sur 9), les groupements prépondérants sont les mêmes dans deux

80. ↑ Citée par García Rivero [156, p. 63].

81. ↑ Une fibule est une agrafe, généralement en métal, qui sert à fixer les extrémités d’un vêtement [438].

82. ↑ *clustering*, en anglais.

83. ↑ *Centroid linkage*, en anglais.

84. ↑ Voir paragraphe *Ethnogenèse ou phylogenèse ?* page 67 concernant le débat phylogenèse/ethnogenèse en anthropologie.

phases consécutives, suggérant une prépondérance du mécanisme de phylogénèse. Ils étudient aussi trois cas où un nouveau peuplement apparaît et cherchent à savoir si il a été formé à partir d'une seule ou plusieurs des populations des autres sites. Dans chacun de ces trois cas, ils comparent l'ensemble des peuplements de la phase archéologique précédente ainsi que le peuplement nouvellement constitué par une analyse cladistique et mesurent là encore les valeurs de *bootstrap* des regroupements formés. Dans deux cas sur les trois, le nouveau peuplement est regroupé avec un seul autre groupement suggérant un mécanisme de phylogénèse. Le dernier cas semble lui aussi favoriser l'hypothèse de phylogénèse, même si le résultat est plus ambigu.

Ethnogenèse ou phylogénèse ? L'étude pionnière de Collard et Shennan [93] est représentative, par certains aspects, des utilisations qui vont être faites des méthodes issues de l'inférence phylogénétique au cours des années 2000. Ces approches, fondées le plus souvent sur des analyses cladistiques, sont basées sur une approche strictement hypothético-déductive consistant à formuler une hypothèse et dont la méthode utilisée sera un moyen de la tester. Une large part des efforts est consacrée au débat initié dans les années 1990 [408, p. 444] concernant les importances relatives des mécanismes de phylogénèse et d'ethnogenèse⁸⁵.

Ainsi Tehrani et Collard [408] analysent les ornements de 60 textiles provenant de 5 tribus turkmènes afin d'évaluer si il y a eu une modification de l'influence respective de ces deux mécanismes dans l'évolution des tapis turkmènes avant et après la conquête russe de 1881. À partir de différentes analyses basées sur les méthodes cladistiques (PTP⁸⁶, *bootstrap*, comparaison d'analyses cladistiques) ils suggèrent que le processus de phylogénèse a toujours été prépondérant, même si l'invasion russe a entraîné une augmentation relative du processus d'ethnogenèse. Dans la même idée, Tehrani et Collard [410] cherchent à comprendre le rôle de ces mécanismes dans l'évolution de tissus de 6 tribus iraniennes : ils analysent la présence de techniques de préparation, de fabrication et de motifs décoratifs par une méthode cladistique et concluent de la valeur de l'indice de rétention calculée (0,59) que la transmission est majoritairement verticale.

Grâce au test PTP, Jordan et Shennan [217] examinent eux aussi la présence d'un signal phylogénétique dans leur jeu de données sur les techniques de vannerie des populations nord-américaines de la côte ouest⁸⁷. Jordan et Mace [215]⁸⁷ s'appuient sur la mesure de l'indices de cohérence et de l'indice de rétention pour évaluer l'importance de la transmission verticale. Jordan et O'Neill [216] procèdent de même mais l'analyse cladistique est précédée d'une analyse du réseau de voisins⁸⁸ visant à évaluer si les données sont bien expliquées par un modèle en arbre. Prentiss et coll. [340] se sont récemment impliqués dans ce débat en analysant par des méthodes cladistique et phénétique les caractéristiques de planches de skateboard, proposant une évolution principalement ramifiée dans un premier temps suivie par une période d'évolution avec beaucoup

85. ↑ Respectivement *phylogenesis* et *ethnogenesis* en anglais. On parle aussi parfois de *ramification* (*branching*, en anglais) et de *mélange* (*blending*, en anglais).

86. ↑ *Permutation tail probability*, en anglais. Ce test est décrit plus en détail dans l'entrée *test de longueur par permutation* du glossaire.

87. ↑ Voir aussi le paragraphe *Histoires musicale et culturelle* page 260 concernant ces études.

88. ↑ *NeighborNet*, en anglais.

plus de mélange entre les différents types de planches.

Autres hypothèses testées. Buchanan et Collard [69] s'intéressent aux routes migratoires ayant permis le peuplement de l'Amérique du Nord. Pour cela, ils effectuent une analyse **cladistique** de pointes de projectiles de 25 ensembles archéologiques de la culture paléindienne Clovis. Ils proposent par ailleurs un certain nombre d'hypothèses : 3 hypothèses expliquant la distribution de la variation de ces pointes et 4 hypothèses proposant chacune une voie d'entrée différente pour cette colonisation originelle. Pour chacune de ces hypothèses, les prédictions qui en découlent sont représentées sous la forme d'un **arbre**. Le **cladogramme** est ensuite comparé aux arbres représentant les différentes hypothèses afin d'évaluer la vraisemblance de ces dernières.

Ces méthodes sont utilisées comme moyen de tester des hypothèses portant sur l'histoire culturelle et pas seulement dans le débat phylogénèse/ethnogenèse. Lycett publie une série d'articles dont l'approche est strictement hypothético-déductive et s'appuyant sur l'analyse cladistique : les hypothèses sont formulées, puis les résultats que devrait produire l'analyse cladistique sous chacune de ces hypothèses sont explicités (prédictions). Par exemple, l'hypothèse de Movius-Schick propose que l'absence de l'industrie lithique par débitage Levallois (mode III⁸⁹) en Asie orientale soit due à la faible présence de l'industrie acheuléenne (mode II⁸⁹) dans cette région [268]. Lycett [268] cherche à tester cette hypothèse en étudiant des ensembles de pierres taillées : il formule ses prédictions en décrivant les **topologies** des **arbres** qui seraient compatibles ou incompatibles avec l'hypothèse puis il effectue une analyse **cladistique** de ces ensembles et compare les résultats avec ses prédictions⁹⁰. Dans une logique similaire, Lycett teste l'hypothèse selon laquelle l'apparition de haches de l'industrie lithique acheuléenne hors d'Afrique est issue de la migration de populations humaines depuis l'Afrique [269], ou encore l'hypothèse suggérant que l'industrie lithique « Victoria ouest » est une forme ancestrale de l'industrie Levallois [270].

D'une manière semblable, Tripp [423] effectue une analyse cladistique d'une trentaine de statuettes gynomorphes⁹¹ européennes provenant de la phase préhistorique gravétienne afin de tester deux hypothèses concurrentes concernant les liens entre ressemblance et proximité géographique des statuettes.

2.1.4.3 Élargissement des approches

Inférer les relations. Si l'analyse cladistique est dans un premier temps utilisée comme un outil permettant de tester des hypothèses historiques, elle devient progressivement, au cours des années 2000, un moyen pour inférer directement l'histoire des cultures ou des objets culturels (en particulier des relations d'apparentement) sans formuler d'hypothèse *a priori*. La première étude dans cette perspective est peut-être l'analyse **cladistique** d'un instrument de musique (le psalté-

89. ↑ Les *modes* 1 à 5 sont utilisés pour décrire les techniques de taille des pierres, le mode 1 étant apparu le plus tôt et le mode 5 le plus récemment.

90. ↑ Cependant l'arbre produit est incompatible avec l'hypothèse de Movius-Schick, mais aussi avec d'autres résultats établis, suggérant un problème dans l'analyse ce qui conduit l'auteur à modifier son analyse.

91. ↑ C'est-à-dire des représentations humaines (anthropomorphes) femelles.

rion baltique) faite par Temkin [412] afin de retrouver son histoire, complétée quelques années plus tard par l'étude des relations entre des cornets à piston grâce à différentes méthodes issues de l'inférence phylogénétique [413]⁹². Harmon et coll. [182] étudient via différentes variantes cladistiques les motifs décoratifs de céramiques provenant de 5 traditions culturelles précoloniales du Mexique afin d'inférer les relations entre les *traditions intellectuelles*⁹³ sous-tendant les systèmes symboliques associés à ces motifs.

Contrastant avec leur analyse précédente [69] présentée ci-dessus⁹⁵ fondée sur un test d'hypothèses, Buchanan et Collard [70] cherchent à retracer directement les relations entre 5 ensembles archéologiques nord-américains en réanalysant les données de Goebel et coll. [160]⁹⁶. Ils effectuent 31 analyses phénétiques fondées sur 15 mesures de distance et avec 3 méthodes différentes (*UPGMA*, *neighbor joining* et méthode de Wards) pour montrer la sensibilité de ce type d'analyse au choix de la mesure de distance et de la méthode. Afin de mieux comprendre les relations entre ces sites, ils procèdent par ailleurs à une analyse cladistique à partir de ces informations concernant la présence ou l'absence d'une centaine de types d'outils sur ces ensembles archéologiques. De la même manière, Tehrani et coll. [411] répètent leur analyse des tissus iraniens [410]⁹⁷. mais la méthode cladistique n'est plus utilisée comme un moyen de tester l'importance relative de la phylogénèse et de l'ethnogenèse : l'arbre produit est pensé comme une hypothèse sur les relations historiques entre les cultures. Cet arbre culturel est d'ailleurs comparé, dans la suite de l'étude, à l'arbre des populations⁹⁸ par une méthode dite de cophylogenie [A88]⁹⁹ cherchant à expliquer les différences entre les deux arbres par des événements de transferts horizontaux, de duplications ou d'extinctions culturelles.

Rogers et coll. [358] appliquent cette fois une méthode de distances (le *neighbor joining*) pour inférer les relations historiques entre les populations de 11 groupes d'îles polynésiennes à partir de la description des caractéristiques fonctionnelles et symboliques de leurs canoës : le but final est de comprendre l'origine géographique et les migrations successives de ces populations. Afin d'établir les relations historiques entre certaines cultures du pacifique, Cochrane et Lipo [92] analysent les motifs décorant les poteries de la civilisation Lapita¹⁰⁰ en utilisant deux méthodes issues de l'inférence phylogénétique (cladistique et réseau de voisins), complétées par une nouvelle méthode qu'ils nomment *réseau de distance phénétique*¹⁰¹. Après s'être servi de la méthode du réseau de voisins pour visualiser les relations établies à partir de la description de décorations

92. ↑ Ces deux études sont présentées plus en détail dans la partie 2.1.5.1 page 70.

93. ↑ Une tradition intellectuelle (ou *lignée d'apprentissage*) est définie ainsi : « information transmitted between individuals and their larger groups »⁹⁴ [182, p. 209].

94. « information transmise entre individus et entre groupes plus larges »

95. ↑ Voir paragraphe *Autres hypothèses testées* page 68.

96. ↑ Citée par Buchanan et Collard [70].

97. ↑ Voir paragraphe *Ethnogenèse ou phylogénèse ?* page 67.

98. ↑ La méthode utilisée pour obtenir cet arbre des populations, issu de données ethnohistoriques et linguistiques [411, p. 3869], n'est pas explicite dans l'article, mais l'on peut sans doute en avoir une idée en consultant leur précédente publication sur le sujet [410].

99. ↑ Citée par Tehrani et coll. [411, p. 3870].

100. ↑ La civilisation Lapita est une civilisation des îles océanienne dont les sites les plus anciens ont environ 3500 ans.

101. ↑ *Phenetic distance network*, en anglais.

de tissus d'Ikat de chaîne¹⁰² provenant d'Asie du Sud-Est, Buckley [71] infère les relations entre ces 36 traditions culturelles par une méthode d'inférence bayésienne et en tire des hypothèses concernant les migrations des populations et les formes ancestrales de ces tissus. De même, Le Quellec [245] étudie les relations entre environ 160 peintures rupestres¹⁰³ du Sahara central par *neighbor joining*, par réseau de voisins et par inférence bayésienne.

Évaluer ces outils d'inférence. On constate donc au cours des années 2000 un changement dans l'utilisation des méthodes issues de l'inférence phylogénétique : elles ne servent plus à tester des hypothèses dans une logique hypothético-déductive, mais à proposer directement un schéma des relations historiques entre les groupes étudiés. Ce changement d'approche est accompagné par le développement d'études visant à évaluer la pertinence des méthodes employées lorsqu'elles sont appliquées à des données culturelles. Ainsi Marwick [280], dans son analyse cladistique d'une quarantaine de statuettes de Bouddha de bronze, s'attache à tester les capacités de la méthode à inférer les relations historiques entre les *artéfacts* en utilisant six tests statistiques afin d'estimer la qualité de l'adéquation entre le jeu de données initial et l'arbre produit : le calcul de l'IC et de l'IR, l'analyse de *bootstrap*, le PTP, le test de dissymétrie¹⁰⁴ et la mesure de l'indice de Bremer.

Très récemment, Schillinger et coll. [373] ont publié une étude allant plus loin dans l'évaluation de la pertinence des méthodes cladistiques pour inférer les relations historiques à partir de l'étude d'objets culturels matériels. L'évolution d'une tradition culturelle est simulée par une expérience de transmission en chaîne d'objets sculptés : en sculptant dans une mousse à l'aide soit d'un couteau en plastique soit d'un épluche-légumes, chaque participant-e réplique l'objet reproduit par le ou la participant-e précédent-e. Ceci permet d'obtenir une tradition artificielle dont on connaît l'histoire réelle. Les *artéfacts* obtenus sont ensuite analysés par une méthode cladistique permettant d'intégrer des *caractères continus*. En comparant les résultats de cette analyse avec l'histoire réelle de cette tradition artificielle, les auteur-e-s constatent que l'analyse permet de retrouver une partie de l'histoire, même si un certain nombre d'objets sont mal positionnés dans l'arbre¹⁰⁵.

2.1.5 Musicologie

2.1.5.1 Les travaux de Temkin

Temkin [412] s'intéresse à l'histoire d'un instrument de musique : le psaltérion baltique. Pour cela, il caractérise 26 instruments issus de descriptions ethnographiques ou archéologiques à l'aide de 30 paramètres divisés en deux types : structurels et décoratifs. La *matrice de caractères* ainsi

102. ↑ *Warp ikat*, en anglais. L'Ikat de chaîne est une technique de décoration des tissus où les motifs sont formés par coloration du fil de chaîne, en amont du tissage.

103. ↑ Toutes représentent des *théranthropes*, êtres mythiques à la fois humains et animaux [245, p. 155].

104. ↑ Le test de dissymétrie (*skewness test* en anglais) mesure à quel point les longueurs de la totalité des arbres obtenus est biaisée par rapport à une loi normale : si il existe un signal phylogénétique, on s'attend à trouver un faible nombre d'arbres courts et un grand nombre d'arbres longs [280, p. 167].

105. ↑ Cette étude évoque l'approche utilisée en *critique textuelle* consistant à créer puis étudier des *traditions* artificielles de manuscrits, présentée dans le paragraphe *Validation sur données artificielles* page 60.

constituée est analysée par une méthode cladistique¹⁰⁶. L'arbre des relations entre ces différents psaltériens obtenu est ensuite utilisé pour discuter des routes probables de transmission de cet instrument. Un des intérêts de cette analyse est qu'elle ne suppose pas que l'histoire ainsi inférée soit celle des populations, mais bien celle de l'instrument lui-même, éventuellement différente. Ainsi la verticalité identifiée est celle de l'histoire instrumentale et peut traverser horizontalement la verticalité de l'histoire des populations.

Au cours d'une étude ultérieure Temkin et Eldredge [413] s'intéressent aussi à l'histoire du cornet à pistons. Celle-ci étant bien documentée, ils se proposent de s'en servir pour tester si des méthodes issues de l'inférence phylogénétique peuvent être utilisées pour inférer l'histoire des instruments. Ils construisent un arbre schématisant les relations historiques entre les cornets grâce à trois méthodes : cladistique, phénétique et en réseau. L'ajout des réticulations déduites du réseau sur l'arbre issu de l'analyse phénétique leur permet de proposer des événements de transferts horizontaux. Ils constatent que les résultats ne rejoignent pas l'histoire établie et que la méthode cladistique réussit moins bien que la méthode phénétique [413, p. 148], suggérant le peu d'efficacité de ces méthodes que le développement des méthodes de réticulation pourrait améliorer [413, p. 151].

2.1.5.2 Les travaux de Toussaint et coll.

Présentation des analyses. Depuis le début des années 2000, un groupe de chercheur-e-s utilise les outils d'inférence phylogénétique développés par la biologie de l'évolution pour étudier des objets musicaux. Ces analyses ont été initiées par Toussaint [419] dont l'approche pourrait être reliée à l'ethnomathématique telle que pratiquée par exemple par Marc Chemillier. Alors que les premières études se sont concentrées sur l'analyse comparative de différentes **formules rythmiques**¹⁰⁷, des études plus récentes utilisent ces **méthodes issues de l'inférence phylogénétique** pour analyser des mélodies de flamenco¹⁰⁸.

Les analyses sont toutes basées sur le calcul de la similarité entre chaque paire d'éléments musicaux afin de produire une **matrice de distances**. Celle-ci sera ensuite utilisée pour construire soit un **arbre**¹⁰⁹, soit un **réseau de bipartitions**^{110,111}.

Cependant la plupart de ces études ne précise pas l'algorithme utilisé pour la construction de l'arbre, ce qui rend impossible la bonne compréhension des résultats. Toussaint [419, p. 17] se sert d'une application utilisable sur le site web du *Computational Biochemistry Research Group* du *Swiss Federal Institute of Technology*, mais sans préciser l'adresse de l'outil en question¹¹². La plupart des autres études ([72, 124, 298, 421]) se contente de préciser que le logiciel *SplitsTree* a

106. ↑ Le détail de la méthode n'est malheureusement pas fourni, pas plus que la matrice de données.

107. ↑ Notamment les études suivantes : [124, 210, 414, 419, 421].

108. ↑ Voir par exemple les travaux de Cabrera et coll. [72] et Mora et coll. [298].

109. ↑ Voir par exemple les travaux de Cabrera et coll. [72], Thul et Toussaint [414] et Toussaint [419].

110. ↑ *Splits network*, en anglais.

111. ↑ Voir par exemple les travaux de Díaz-Báñez et coll. [124], Toussaint [419] et Toussaint [421].

112. ↑ Il s'agit vraisemblablement de l'application TreeGen, accessible à cette adresse : <http://www.cbrg.ethz.ch/services/TreeGen>. Cependant cette application, qui ne laisse que peu de choix dans les paramètres de l'analyse, ne précise pas les algorithmes utilisés pour construire l'arbre.

été utilisé, sans préciser la méthode utilisée (réseau ou arbre, type de réseau, méthode de constitution de l'ensemble de bipartitions, méthode de construction du réseau, ...) alors que ce logiciel implémente un grand nombre de méthodes différentes [207]. Seule l'étude de Thul et Toussaint [414] précise que la construction de l'arbre phylogénétique à partir de la matrice de distance a été faite par l'algorithme BioNJ du logiciel *SplitsTree*. Connaître le détail des méthodes utilisées est pourtant indispensable, comme dans toute analyse scientifique, pour pouvoir comprendre et évaluer les résultats d'une étude.

Tous ces travaux, dont les approches sont explicitement plus mathématiques qu'ethnomusicologiques, cherchent avant tout à classer et à catégoriser les objets musicaux étudiés, éventuellement automatiquement par des méthodes informatisées. Cependant des interprétations historiques et évolutives en sont parfois tirées. Selon que l'un ou l'autre de ces objectifs soit visé (classification ou inférence historique) les implications méthodologiques sont différentes et j'en discuterai donc dans les deux paragraphes suivants. De nombreux aspects de ces travaux nécessiteraient des discussions plus détaillées, par exemple concernant le choix des mesures de distances utilisées. Je n'aborderai cependant pas ces aspects ici en me focalisant sur la pertinence de l'utilisation des **méthodes issues de l'inférence phylogénétique** dans ces études.

En visant la classification. Les différentes études cherchent à catégoriser un ensemble de musiques selon des **paramètres musicaux internes** (rythmes et mélodies). Certaines cherchent aussi à identifier des paramètres musicaux coïncidant, dans les populations étudiées, avec des paramètres externes (notamment géographiques). Les méthodes **phénétiques** en particulier, développées dans le cadre de l'inférence phylogénétique à partir de mesures de distances, permettent de visualiser les données de **matrice de distances** et représenter les proximités entre différentes entités, sans hypothèses quant à leur parenté évolutive. On peut cependant regretter que le choix de cette méthode d'organisation et de visualisation des données de distances ne soit pas ou peu justifié, par exemple par des connaissances *a priori* ou par des tests *a posteriori* sur l'adéquation de cette représentation. Par exemple Thul et Toussaint [414] cherchent à évaluer différentes mesures de la similarité de **formules rythmiques**. Pour cela, ils calculent les distances entre les formules rythmiques de deux jeux de données (un de 11 formules et l'autre de 38), contenant des rythmes provenant d'Inde et d'autres d'Afrique. À partir des différentes **matrices de distances** obtenues, ils construisent ensuite un **arbre** grâce à un dérivé de l'algorithme de **neighbor joining**¹¹³. Les mesures de distances utilisées sont alors considérées comme meilleures si l'arbre correspondant réussit bien à séparer les rythmes indiens et les rythmes africains. Cependant les auteurs n'expliquent ni ne justifient leur choix d'une hiérarchisation **dichotomique** en général et d'une méthode phénétique en particulier pour séparer les rythmes de différentes provenances géographiques. Les études fondant la catégorisation sur la construction de réseaux de bipartitions sont d'ailleurs sujettes aux mêmes critiques.

113. ↑ Il s'agit de l'algorithme *BioNJ* du logiciel *SplitsTree* (voir partie 2.3.3.4 page 90).

En visant la reconstruction évolutive. La catégorisation effectuée à l'aide des **méthodes issues de l'inférence phylogénétique** est cependant parfois utilisée pour en tirer des hypothèses sur l'histoire évolutive des musiques étudiées [72, 124, 298, 421]. Ainsi Toussaint [421, p. 15] explique-t-il :

« [...] phylogenetic trees provide a useful tool for visualizing the interrelationships between the rhythms as well as determining economical mechanisms for their generation. These mechanisms may in turn shed light on the evolution of such rhythms. »¹¹⁴

Toussaint (2003) [421, p. 15]

Si cette proposition peut être intéressante, le lien entre similarité et ascendance commune n'est évidemment pas direct, et nécessite encore une fois des discussions qui ne sont malheureusement jamais présentes. Il faudrait en effet *a minima* expliquer en quoi les mesures de distance effectuées sont susceptibles d'être porteuses de traces historiques et comparer les résultats obtenus avec d'autres éléments sur l'histoire des musiques étudiées afin d'en évaluer la pertinence.

2.1.5.3 Les travaux de Le Bomin et coll.

Le Bomin et coll. [240] viennent de publier un article dans lequel une méthode cladistique est utilisée pour étudier les musiques du Gabon. Les patrimoines musicaux de 58 populations réparties dans tout le Gabon sont décrits à l'aide de 322 caractères. Ceux-ci sont divisés en trois catégories : les **répertoires**, les **paramètres performatifs** (procédés polyphoniques, formes vocales, instruments de musique et **techniques vocales**) et les **paramètres intrinsèques** (métrique, **formules rythmiques** et **échelles** musicales) [240, p. 4]. L'analyse cladistique de cette **matrice de caractères** leur permet de construire un arbre dont l'**indice de cohérence** vaut 0,6 et l'**indice de rétention** vaut 0,7, ces fortes valeurs étant interprétées comme un signal de transmission verticale¹¹⁵. De nombreux **clades** révèlent une pertinence selon des critères indépendants aux données utilisées. La séparation la plus marquante intervient au plus haut niveau, puisque l'arbre séparerait au niveau de la racine les patrimoines musicaux des populations **patrilinéaires** d'une part et ceux des populations **matrilinéaires** d'autre part. Par ailleurs, huit clades ont une pertinence géographique et 9 clades ont une pertinence **ethnonymique**.

2.1.6 Autres applications

Comme on l'a vu, les **méthodes issues de l'inférence phylogénétique** ont été réutilisées de manière relativement large en critique textuelle, en linguistique historique et en anthropologie de la culture matérielle, générant pour chacune de ces disciplines des questionnements et des

114. ↑ « [...] les arbres phylogénétiques sont un outil utile pour visualiser les interrelations entre les rythmes ainsi que pour identifier des mécanismes économiques pour leur génération. Ces mécanismes peuvent en retour éclairer l'évolution de ces rythmes. »

115. ↑ Le paragraphe *Indices de cohérence et de rétention* page 79 traite de l'utilisation de ces indices pour évaluer l'importance relative des transmissions horizontale et verticale. La partie 7.2.3.2 page 264 aborde les différences de ces indices selon les catégories de paramètres.

approches plus ou moins spécifiques. Elles ont aussi été adaptées à d'autres études culturelles, comme en musicologie, tout en ayant jusqu'à présent une portée moindre dans les disciplines concernées. Je présente dans cette partie quelques exemples de réutilisation de telles méthodes dans d'autres études culturelles. Par exemple Coward et coll. [97] s'intéressent aux utilisations des plantes faites par des populations humaines européennes du néolithique. La présence ou l'absence de plus d'une centaine d'espèces végétales est décrite dans 250 sites archéologiques. Les relations entre une vingtaine de régions sont ensuite inférées par trois méthodes issues de l'inférence phylogénétique (réseau de voisins, analyse cladistique et inférence bayésienne), nourrissant une discussion sur l'origine et l'évolution de l'agriculture. Skelton [384] s'intéresse à l'histoire d'un système d'écriture utilisé dans une forme de grec ancien, le linéaire B, en comparant les caractéristiques morphologiques des signes retrouvés sur différentes tablettes d'argiles.

Concepts scientifiques. Il n'est peut-être pas étonnant qu'une des premières réutilisations de ces méthodes issues de la biologie soit l'œuvre d'un systématicien. Ainsi, en 1987, Carpenter [79] publie, dans le contexte du débat encore vif entre méthodes cladistiques et phénétiques, une analyse satirique comparant 40 systématicien·ne·s par une méthode cladistique et une méthode phénétique, à partir de l'analyse de 14 caractères aussi divers que les conceptions scientifiques, les méthodes utilisées ou les journaux dans lesquels publient les auteur·e·s. Cette étude, relevant de l'humour scientifique, sera reprise de manière sérieuse quelques décennies plus tard par Ebach et coll. [129] qui établit par une méthode cladistique une classification de 4 écoles de pensées en systématique cladistique, à partir de 7 caractères décrivant les conceptions scientifiques de chaque courant. Fislser et Lecointre [141] poussent beaucoup plus loin les possibilités d'étudier les relations entre concepts scientifiques par les méthodes cladistiques. Une quarantaine de travaux, théoriques ou pédagogiques, portant sur le concept d'*arbre du vivant* sont décrits à l'aide de 91 caractères concernant aussi bien des aspects descriptifs que méthodologiques ou symboliques. L'analyse cladistique de ces données leur permet de retrouver certaines écoles scientifiques reconnues comme telles, mais aussi d'identifier d'autres regroupements moins habituels.

Traditions orales. Les méthodes issues de l'inférence phylogénétique ont plus récemment été utilisées pour étudier la culture orale. Ainsi Tehrani [409] appliquent des méthodes phénétique, cladistique et probabiliste pour étudier les relations entre une soixantaine de variantes de deux types de contes¹¹⁶ parfois difficilement différenciables (*Le petit chaperon rouge* et *Le loup et les enfants*). Par une approche similaire, d'Huy [120–122] s'intéresse à retrouver l'histoire et les relations entre différents mythes partagés par de nombreuses populations à l'échelle continentale voire mondiale¹¹⁷.

116. ↑ D'après la classification Aarne-Thompson-Uther (ATU) des contes-types.

117. ↑ Bortolini [53] s'est récemment intéressé à la distribution de contes populaires à l'échelle de l'Afrique et de l'Europe, mais ces recherches ne sont à ma connaissance pas encore publiées.

Économie. Ces méthodes ont aussi été appliquées, depuis le début des années 2000, pour étudier les organisations et les systèmes de production, par exemple par McCarthy et coll. [A283]¹¹⁸, McCarthy et coll. [A285]¹¹⁸, Leseure [A260]¹¹⁸, Leask [A247]¹¹⁸, Andersen [A6]¹¹⁸ et Baldwin et coll. [A30]¹¹⁸. Ainsi McCarthy [284, pp. 289–290] donne un exemple¹¹⁹ où des caractéristiques des systèmes de production, telles que la présence d’automatisation ou de chaîne d’assemblage, peuvent servir à les classer grâce à une méthode **cladistique**.

2.2 Considérations générales sur ces transpositions interdisciplinaires

Une analyse détaillée des spécificités ou convergences de ces réutilisations méthodologiques dans différentes disciplines est hors de la portée de cette thèse. Je propose cependant dans cette partie quelques réflexions issues de mes lectures plus ou moins partielles des travaux présentés dans la partie 2.1. Elles concernent principalement deux points : les spécificités méthodologiques ou épistémologiques de certaines disciplines et la question de la transmission horizontale.

2.2.1 Spécificités disciplinaires

Diversité des utilisations. Comme on a pu le voir, les utilisations de **méthodes issues de l’inférence phylogénétique** pour étudier des données culturelles prolifèrent (notamment depuis les années 2000) et sont nombreuses et variées. Cette variabilité concerne aussi bien les méthodes utilisées (**phénétique**, **cladistique** ou **probabiliste**, avec une grande diversité au sein même de ces catégories générales) que les objectifs poursuivis, y compris pour une même méthode. Ainsi, l’analyse cladistique peut-être utilisée pour classer un ensemble de données (par exemple chez Ebach et coll. [129]), pour inférer des relations historiques (par exemple dans l’étude de Rexová et coll. [347]), tester une hypothèse formulée *a priori* (Lycett [268]) ou encore tester la présence d’un signal de transmission vertical dans les données étudiées (cas des travaux de Le Bomin et coll. [240]). Cette diversité globale se retrouve au niveau disciplinaire : dans toutes les disciplines où j’ai relevé un nombre relativement important d’études (c’est-à-dire la critique textuelle, la linguistique historique, l’anthropologie de la culture matérielle, l’ethnomusicologie) les trois grandes catégories de **méthodes issues de l’inférence phylogénétique** ont été utilisées. On la retrouve même parfois au sein d’une même analyse, puisque certaines (par exemple Roos et Heikkilä [361]) comparent les mérites respectifs de nombreuses méthodes différentes quand d’autres utilisent une même méthode à différentes fins.

Particularismes. À travers cette diversité, on peut cependant relever des spécificités de certaines disciplines. Ainsi les études issues de l’anthropologie culturelle font beaucoup plus souvent

118. ↑ Citées par McCarthy [284, pp. 288–289].

119. ↑ Cet exemple est à seul but pédagogique.

appel à des méthodes cladistiques que la linguistique historique. Celle-ci fait un usage important des méthodes probabilistes, en particulier ces dernières années. L'intérêt de cette discipline est souvent l'inférence des relations historiques ainsi que des dates de divergence et des localisations géographiques ancestrales des groupes identifiés. Si la question des migrations intéresse aussi l'anthropologie culturelle et l'archéologie, elle l'aborde plus souvent par le biais formel d'un test d'hypothèse dans une logique hypothético-déductive. En critique textuelle, au-delà des relations entre manuscrits c'est bien souvent le contenu du texte des manuscrits anciens qui est recherché, dans la visée de la publication d'une édition critique d'un texte. On voit donc que l'adaptation de ces méthodes aux spécificités de la discipline et leur insertion au sein d'une tradition méthodologique propre¹²⁰ conduit à des tendances particulières dans l'étude de chaque type d'objet culturel.

2.2.2 La question de la transmission horizontale

Une critique récurrente de l'utilisation des **méthodes issues de l'inférence phylogénétique** afin d'étudier l'évolution culturelle réside dans le fait que cette dernière serait beaucoup plus sujette à la **transmission horizontale** que l'évolution biologique, rendant ces méthodes inappropriées [172, p. 2299]. Afin de mieux comprendre les problèmes que cela pose et les manières de les résoudre, je présenterai tout d'abord quelques études cherchant à évaluer l'impact qu'a la transmission horizontale sur les résultats des méthodes phylogénétiques, puis j'évoquerai certains outils utilisés pour détecter et quantifier la transmission horizontale dans les données culturelles.

2.2.2.1 Influence de la transmission horizontale

Si le problème de l'impact de la **transmission horizontale** sur les résultats des méthodes phylogénétiques appliquées à des données culturelles a suscité de nombreux débats [172, p. 2299], ceux-ci sont restés essentiellement théoriques. Les quelques études ayant abordé empiriquement la question sont basées sur la simulation : des données évolutives sont créées par évolution artificielle simulée par ordinateur, puis la méthode testée est appliquée à ces données. Connaissant l'histoire réelle de cette évolution artificielle, l'efficacité de la méthode peut ainsi être évaluée. On peut aussi tester de la même manière sa **robustesse**, c'est-à-dire sa capacité à donner des résultats similaires lorsque l'on modifie légèrement les données ou les paramètres de l'analyse. Cette approche a permis de tester l'influence de la transmission horizontale sur la justesse de la topologie inférée d'une part et sur l'exactitude des dates de divergence estimées d'autre part. Elle a aussi permis de tester son influence lors de l'utilisation des **méthodes comparatives phylogénétiquement informées**. Ce sont ces trois aspects que je détaille ci-dessous.

Sur la topologie inférée. La première étude que j'ai pu retrouver testant l'influence de la transmission horizontale sur la **topologie** de l'**arbre phylogénétique** inféré est celle publiée par

120. ↑ Et parfois avec une logique similaire, comme dans le cas déjà mentionné de la méthode **stemmatiques** en critique textuelle ou de la **méthode comparative** en linguistique historique.

McMahon et McMahon [287, pp. 111–118] dans un ouvrage portant plus généralement sur la classification des langues à partir de méthodes quantitatives. Le but originel de cette analyse est de simuler numériquement des relations phylogénétiques par évolution artificielle puis d’y chercher les motifs caractéristiques d’un emprunt, ce qui permettrait ensuite de repérer les événements de transmission horizontale. Pour cela les auteur·e·s font évoluer 200 unités lexicales (des « mots ») sur une période de 220 générations d’après un arbre dont la topologie a été choisie à l’avance et comprenant 12 langues. Deux modèles d’évolution sont utilisés : le modèle « hihi » plutôt conservateur (une modification par mot en moyenne sur les 220 générations) et le modèle « lolo » dont la vitesse d’évolution moyenne est double. De plus, un emprunt d’intensité variant de 0 à 20% est permis entre deux des langues évoluant (A et B). Une *matrice de distances* est ensuite calculée à partir des fréquences observées des mots apparentés puis un arbre phylogénétique est inféré avec le logiciel *Phylip*. Les auteur·e·s montrent que pour un taux d’emprunt de 10% entre A et B, la topologie inférée est erronée dans 80% des cas, l’effet s’accroissant lorsque le taux d’emprunt augmente.

Cependant, comme le soulignent Greenhill et coll. [172, p. 2304], l’étude comporte de nombreux biais. Tout d’abord la méthode d’inférence phylogénétique utilisée est peu puissante car elle est basée sur le calcul de distances globales, ce qui a pour effet de résumer les différences à un seul chiffre et donc à perdre une partie de l’information historique. Le mode d’emprunt utilisé, autorisé seulement entre A et B (et qualifié de « systématique ») est très particulier et n’affecte pas les principales branches de l’arbre [172, p. 2304]. De plus le choix d’une topologie arbitraire et de deux *taxons* arbitraires autorisés à échanger rend l’étude conditionnée par ces choix sans que la robustesse à une variation de ces choix *a priori* ne soit testée. Enfin la différence entre la topologie inférée et la topologie réelle n’est pas quantifiée, alors que les topologies considérées fausses ne diffèrent que peu de la topologie réelle.

L’étude de Greenhill et coll. [172] a pour but explicite de tester l’influence de la transmission horizontale sur la topologie et les dates de divergences inférées. Les auteurs simulent l’évolution de langues à l’aide d’une variante stochastique du modèle de Dollo¹²¹, où les mots sont considérés apparaître, être empruntés, puis disparaître (*birth–borrow–death*) avec une certaine probabilité fixée *a priori*. L’évolution se fait selon deux topologies prédéfinies, l’une considérée équilibrée et l’autre moins. Le taux d’emprunt est variable dans les différentes simulations et peut se faire de manière globale (un mot pouvant passer de n’importe quelle langue à n’importe quelle autre) ou de manière locale (seules les langues ayant divergé il y a moins de N années peuvent échanger des mots). Les relations phylogénétiques sont ensuite inférées à partir des données produites grâce à une méthode *bayésienne*. La justesse des topologies est enfin quantifiée par la *perturbation* : la mesure de distance de quadruplets (QDM¹²³) est estimée entre la topologie inférée et la topologie réelle, ainsi qu’entre la topologie réelle et celle d’un arbre aléatoire. La perturbation est la différence entre ces deux QDM normalisées. Les auteurs montrent ainsi que la topologie inférée reste

121. ↑ Le modèle utilisé est repris de l’étude de Nicholls et Gray [A317]¹²².

122. Citée par Greenhill et coll. [172, p. 2300].

123. ↑ Pour l’anglais *quartet distance metric*, mesurant la différence entre deux topologies.

toujours plus juste qu'une topologie aléatoire. De manière générale, les résultats sont meilleurs lorsque la topologie réelle est équilibrée. Lorsque l'emprunt est local, la topologie reste très juste même lorsque le taux d'emprunt est fort (jusqu'à 50%). La justesse est moindre lorsque l'emprunt est global et diminue avec le taux d'emprunt. Les auteurs montrent donc qu'un taux d'emprunt réaliste (autour de 20% selon eux) ne diminue que peu la qualité de l'inférence phylogénétique.

Sur l'estimation des dates de divergence. Afin de vérifier les résultats de leur analyse supportant l'hypothèse de l'origine anatolienne de la langue indo-européenne, Atkinson et coll. [28] simulent l'évolution de langues à partir de deux modèles : la version stochastique du modèle de Dollo et un modèle *finite-sites*. Différents taux d'emprunt (10%, 20% et 100%) sont testés ainsi que deux modèles d'emprunt, qu'ils qualifient respectivement de *global* et de *local*. Dans ce dernier cas, seules les langues ayant divergé il y a moins de N années peuvent échanger (la valeur de N étant variable elle aussi). La date de divergence du dernier ancêtre commun est ensuite estimée par inférence bayésienne à partir des données générées, puis comparée à la date de divergence réelle. Les auteurs montrent que la date de divergence est toujours sous-estimée (et jamais surestimée). Un taux d'emprunt global assez élevé (20%) ne fausse que très peu la date. Un emprunt local de 20% pour $N = 4000$ ans ne provoque une sous-estimation de la date que pour le modèle de Dollo stochastique, mais pas pour le modèle *finite-sites*. De plus, en simulant un effet de divergence en chaîne des dialectes (taux d'emprunt de 100% pour $N = 500$ ans), les auteurs montrent que la sous-estimation de la date reste faible¹²⁴.

Par ailleurs, Greenhill et coll. [172] infèrent aussi la date de divergence à la racine de l'arbre et la comparent avec la date réelle. Ils montrent là encore que la date de divergence est systématiquement sous-estimée et ce d'autant plus que le taux emprunt est grand. La sous-estimation est plus importante lorsque la topologie de l'arbre est moins régulière. Lorsque l'emprunt est local la sous-estimation reste assez faible, mais elle devient substantielle lorsque l'emprunt est global.

Dans les méthodes comparatives phylogénétiquement informées. Les *méthodes comparatives phylogénétiquement informées*¹²⁶ (MCPIs) sont des méthodes analysant les données d'un ensemble de populations en prenant en compte l'information concernant leurs relations phylogénétiques établies au préalable¹²⁷. S'il existe de la transmission horizontale au sein du système étudié, les relations phylogénétiques utilisées peuvent être inexactes et ceci peut fausser l'analyse se servant des MCPIs. Il est donc nécessaire d'évaluer à quel point ce biais peut impacter les conclusions de l'analyse.

Deux études ont abordé l'influence de la transmission horizontale sur les résultats produits à l'aide de MCPIs, utilisées pour étudier des tendances dans l'évolution de caractères culturels. Ces

124. ↑ À l'instar de cette étude, d'autres articles semblent analyser l'influence de la transmission horizontale pour vérifier la robustesse de leurs analyses, comme ceux de Atkinson et Gray [25] et de Nicholls et Gray [A317]¹²⁵, mais je n'ai pas eu le temps de les approfondir.

125. Citées par Atkinson et coll. [28].

126. ↑ *Phylogenetic comparative methods (PCMs)*, en anglais.

127. ↑ L'entrée du glossaire *méthodes comparatives phylogénétiquement informées* donne des explications plus détaillées et un exemple d'une telle analyse.

études s'intéressent plus particulièrement à deux MCPIs : la méthode des contrastes indépendants et le **test de Mantel**¹²⁸. Alors que Nunn [319] estime que ces deux méthodes sont très sensibles à des taux même faibles de transmission horizontale, Currie et coll. [102] montrent par une étude plus nuancée que les MCPIs peuvent être dans certaines conditions robustes face aux effets de la transmission horizontale.

2.2.2.2 Détection et mesure de la transmission horizontale

De nombreuses méthodes ont été proposées pour détecter et parfois quantifier l'importance de la **transmission horizontale** lors de l'inférence phylogénétique. Cette partie se propose de présenter succinctement quelques-unes d'entre elles.

À partir des arbres phylogénétiques inférés. Développées dans le cadre de l'étude de la coévolution en biologie [A88], les méthodes dites *de jungle*¹²⁹ permettent de mesurer le nombre d'événements de transferts horizontaux nécessaires pour réconcilier deux **arbres phylogénétiques** non congruents et de donner ainsi une estimation du taux de transmission horizontale [102, p. 3910]. Dans le cadre de l'évolution culturelle, cette méthode a par exemple été utilisée par Temkin et Eldredge [413]¹³⁰ pour étudier l'évolution d'instruments de musique et par Tehrani et coll. [411] pour étudier l'évolution de techniques et de motifs de tissage. Dans une perspective proche, Nelson-Sathi et coll. [306] et List et coll. [264] utilisent la méthode en réseau par minimisation des transferts horizontaux¹³¹ [A104]¹³² qui permet, à partir d'un arbre préétabli, de rechercher le nombre minimal de **transferts horizontaux** permettant d'expliquer les données.

À partir d'indices statistiques.

Indices de cohérence et de rétention. L'**indice de cohérence**¹³³ (IC) et l'**indice de rétention**¹³⁴ (IR) sont deux indices statistiques utilisés en biologie pour quantifier la cohérence entre un **arbre** donné et une **matrice de caractères**. Plusieurs études appliquant des méthodes cladistiques à des données culturelles utilisent ces indices pour avoir une idée de la présence de transmission horizontale dans le système étudié [217, 271, 408]¹³⁵ [240]. La logique sous-jacente est la suivante : si la valeur de ces indices est élevée, les données sont relativement bien expliquées par un arbre (plus que des données aléatoires) ce qui suggère un signal de transmission plutôt vertical. Si elle est faible, c'est qu'il est nécessaire d'expliquer les données par de nombreuses **homoplasies**, lesquelles peuvent indiquer une importante transmission horizontale. Afin d'évaluer

128. ↑ Le test de Mantel est décrit dans l'annexe B.3 page 421.

129. ↑ *Jungle methods*, en anglais.

130. ↑ Voir paragraphe *Les travaux de Temkin* page 71.

131. ↑ Voir paragraphe *Réseau par minimisation des transferts horizontaux (MLN)* page 63.

132. ↑ Citée par List et coll. [264, p. 45].

133. ↑ *Consistency Index (CI)* en anglais.

134. ↑ *Retention Index (RI)* en anglais.

135. ↑ Citées par Nunn et coll. [321, p. 3813].

si le processus de mélange¹³⁶, comparativement à celui de ramification¹³⁷, est plus important lors de l'évolution culturelle que lors de l'évolution biologique¹³⁸, Collard et coll. [94] comparent la valeur des ICs et des IRs obtenus à partir de 21 jeux de données biologiques et 21 jeux de données culturels. Les valeurs obtenues ne se révèlent pas significativement différentes, avec un IR moyen de 0,61 (valeurs entre 0,35 et 0,94) pour les données biologiques et de 0,59 (valeurs entre 0,42 et 0,78) pour les données culturelles. Ceci leur permet de conclure que les données culturelles étudiées jusqu'ici ne sont pas spécialement moins bien expliquées par une structure arborescente que les données biologiques, mais aussi de donner une échelle à laquelle se réfèrent par la suite diverses études pour évaluer la grandeur des IRs calculés¹³⁹.

Cependant Nunn et coll. [321, p. 3813] expriment des réserves quant à l'utilisation de ces indices pour évaluer l'importance de la transmission horizontale. En effet ils sont basés sur la mesure de l'homoplasie qui peut indiquer la présence de transmission horizontale, mais aussi de convergences. La valeur des ICs et IRs peut donc être faible, même en l'absence de transmission horizontale, si la vitesse d'évolution est relativement importante. Par ailleurs, si un assez grand nombre d'éléments sont empruntés conjointement¹⁴⁰, les valeurs de ces indices peuvent être élevées et ne pas révéler ces événements de transfert horizontal peu nombreux mais significatifs de par leur impact. Ainsi des ICs et IRs faibles n'indiquent pas nécessairement une forte transmission horizontale et des ICs et IRs forts ne sont pas incompatibles avec une forte transmission horizontale. De plus ces indices ne peuvent être intégrés dans un cadre statistique et ne peuvent être testés rigoureusement, tout au plus peuvent-ils être comparés à d'autres résultats (tels ceux de Collard et coll. [94]) issus de la littérature.

Score δ et score Q-résiduel. Le score δ ¹⁴¹ est un indice proposé par Holland et coll. [A199]¹⁴² pour évaluer à quel point une matrice de distances peut être représentée sous la forme d'un arbre, avant même d'inférer la topologie de cet arbre. Gray et coll. [167] proposent d'utiliser un dérivé de cet indice, le score Q-résiduel¹⁴³, pour évaluer à quel point un réseau de bipartitions¹⁴⁴ a une topologie proche d'un arbre, et donc indirectement permettre de détecter la présence de transmission horizontale et les unités évolutives impliquées.

Par visualisation des données. Il est aussi possible de représenter les données servant à l'inférence phylogénétique avec des méthodes comme le réseau de voisins¹⁴⁵ [67] ne présupposant pas une organisation en arbre, permettant de visualiser d'éventuelles marques de transmission

136. ↑ *Blending*, en anglais.

137. ↑ *Branching*, en anglais.

138. ↑ Voir aussi à ce sujet le paragraphe *Ethnogenèse ou phylogenèse ?* page 67.

139. ↑ Par exemple les études de Tehrani [409, p. 4] ou Le Bomin et coll. [240, p. 8].

140. ↑ De tels emprunts en « paquets » sont semble-t-il vraisemblables, comme je le discuterai dans la partie 7.2 page 251.

141. ↑ Score delta, ou δ -score en anglais.

142. ↑ Citée par Gray et coll. [167, p. 3925].

143. ↑ *Q-residual score* en anglais.

144. ↑ *Split network*, en anglais.

145. ↑ *NeighborNet*, en anglais.

horizontale [28, p. 196]. Par exemple Buckley [71] représente les données issues de la description de 36 traditions de tissage sous la forme d'un réseau de voisin. Ils observent la présence de nombreuses **réticulations** dans le diagramme obtenu, lequel constitue un **réseau** dont la forme est assez différente de celle d'un **arbre**. Les auteurs déduisent de cette observation (combinée avec d'autres) la présence probable de transmission horizontale [71, pp. 10–12].

Autres méthodes. Une manière d'évaluer le taux de transmission horizontale est évidemment de chercher par une analyse détaillée l'origine de chacun des éléments et d'estimer par expertise si il a été emprunté. C'est par exemple l'approche de Bown et coll. [57] dans leur étude de la variabilité des taux d'emprunt et c'est ce qui est fréquemment fait dans les inférences à partir d'unités lexicales où les mots empruntés sont écartés de l'analyse.

J'ai pu recenser diverses autres méthodes de détection et de mesure de la transmission horizontale au cours de mes lectures, sans avoir le temps de comprendre et discuter chacune d'entre elles. Je me contenterai donc de donner les noms et les références des travaux où je les ai trouvées mentionnées, afin que les personnes intéressées puissent s'y référer et approfondir le sujet. Il s'agit du **test de Mantel** [319, p. 182], de la statistique D [162, p. 3], de l'indice d'**arbritude** de Rambaut¹⁴⁶ [358, p. 3838], du **test de longueur par permutation**¹⁴⁷ [408, p. 449], du biais de distribution des longueurs d'arbres¹⁴⁸ [280, p. 167]¹⁴⁹ ou encore de l'**indice de Bremer** [280, p. 168]¹⁴⁹.

2.3 Tableau comparatif de quelques études

2.3.1 Objectifs et limites de ce tableau

Le tableau de la partie 2.3.4 pages 93 à 102 présente une synthèse d'une centaine de publications scientifiques¹⁵⁰ dans lesquelles des données d'ordre culturel sont analysées à l'aide de **méthodes issues de l'inférence phylogénétique**. Axé sous un angle méthodologique, ce tableau de synthèse a plusieurs objectifs. Tout d'abord fournir une liste des publications incluant de telles analyses, liste certainement pas exhaustive, mais qui se veut aussi complète que possible. Ensuite décrire schématiquement la méthodologie utilisée dans ces analyses afin de pouvoir identifier rapidement le type de méthode dont il s'agit et inversement retrouver les analyses correspondant à un certain type de méthodes. Enfin ce tableau a une visée comparative : il cherche à donner un aperçu des tendances et des différences d'approches au sein d'une discipline, mais aussi entre des disciplines distinctes.

146. ↑ *Rambaut's treenes index*, en anglais.

147. ↑ *Permutation tail probability (PTP)*, en anglais.

148. ↑ *Skewness of the distribution of tree lengths*, en anglais.

149. ↑ Cette étude est présentée succinctement dans le paragraphe *Évaluer ces outils d'inférence* page 70.

150. ↑ Comme expliqué dans la partie 2.3.2.2 page 83, chaque ligne du tableau correspond à une *analyse*, c'est-à-dire à une unité d'analyse et d'interprétation de données dans une publication scientifique. Une même publication scientifique peut ainsi contenir plusieurs analyses.

Ce tableau est cependant insatisfaisant sur plusieurs aspects. Tout d'abord la description des méthodes est assez superficielle et ne comprend pas certains aspects essentiels comme les critères de choix des caractères analysés, les représentations et les transformations de ces caractères pour l'analyse (codage), les valeurs des paramètres choisis *a priori* et les critères de ces choix (par exemple les paramètres d'échantillonnage d'une analyse bayésienne par MCMC ou les groupes externes d'une analyse cladistique). De plus, chacune des colonnes de ce tableau implique une paramétrisation de ces analyses, forcément artificielle et rendant difficilement compte de leur diversité. Par ailleurs, cette paramétrisation n'est souvent pas très aboutie et mériterait des améliorations. Ainsi la catégorisation des analyses selon leurs objectifs nécessiterait au préalable une analyse épistémologique rigoureuse. En outre l'assignation d'une analyse à une certaine catégorie ne peut se résumer à une case dans un tableau et nécessiterait de discuter et d'explicitier les raisons de chacun de ces choix, mais c'est encore une fois au-delà de la synthèse que j'ai choisi de présenter ici.

La réalisation de ce tableau m'a en tout cas permis de clarifier ma compréhension et mes questionnements concernant ce type d'analyses. Elle a aussi alimenté ma réflexion sur la manière dont peuvent être réutilisées les méthodes issues de l'inférence phylogénétique, sur la diversité des méthodes d'inférence historique et sur la manière dont se construit la connaissance historique. Je présente dans la partie suivante les critères sur lesquels je me suis basé afin de choisir, pour chaque analyse, les valeurs de la ligne correspondante dans le tableau 2.1. Ceci permet non seulement de donner les clés de compréhension du tableau, mais surtout de discuter plus en détail les différents paramètres et, ainsi, d'approfondir un peu cette réflexion.

2.3.2 Présentation générale

2.3.2.1 Contenu du tableau

Les lignes : des analyses. Chaque ligne du tableau correspond à ce que je nomme ici une *analyse*, c'est-à-dire à une unité cohérente au sein d'une publication scientifique comprenant l'analyse et l'interprétation d'un ensemble de données. Ainsi il est possible de trouver, dans une même publication, plusieurs analyses distinctes de différents jeux de données, ou d'un même jeu de données par différentes méthodes. Par exemple, Spencer et coll. [392] cherchent à retrouver les relations entre une vingtaine de versions de manuscrits d'une *tradition* artificielle en utilisant quatre *méthodes issues de l'inférence phylogénétique* différentes. L'application de chacune de ces méthodes correspond donc à une analyse différente. Cependant, lorsque plusieurs analyses contenues dans une même publication sont identiques en ce qui concerne les colonnes choisies pour les décrire (par exemple lorsqu'une même méthode est utilisée pour étudier deux jeux de données différents), une seule ligne est notée dans le tableau.

Il arrive aussi qu'une même analyse soit retrouvée plusieurs fois dans différentes publications scientifiques, par exemple lorsqu'elle est publiée à la fois dans une thèse et dans une revue à

comité de lecture. Dans ce cas, seule l'une des deux publications¹⁵¹ est décrite par une ligne du tableau, l'autre étant simplement mentionnée en note de bas de page.

Les colonnes : typologie des analyses. Pour chacune de ces analyses les colonnes du tableau présentent, en plus de la référence bibliographique, certaines caractéristiques que j'ai choisi de décrire : les objectifs de l'analyse¹⁵², le type de données utilisées¹⁵³, le type de méthode d'inférence utilisée¹⁵⁴, certains détails de la méthode utilisée¹⁵⁵ et le logiciel utilisé¹⁵⁶. L'explication du contenu de ces colonnes est détaillée dans la partie 2.3.3¹⁵⁷.

2.3.2.2 Choix des analyses

Le tableau est issu de l'étude de 113 publications scientifiques¹⁵⁸. Pour chacune de celles-ci, j'ai sélectionné les analyses :

1. qui utilisent une **méthode issue de l'inférence phylogénétique**,
2. dont les données analysées par cette méthode sont d'ordre culturel,
3. et dont les résultats s'expriment sous la forme de relations entre les objets étudiés.

1. Des méthodes issues de l'inférence phylogénétique. Ce premier critère exclut les nombreuses méthodes d'**inférence historique** développées en dehors de la biologie de l'évolution, puisque le terme de *phylogénie* est explicitement défini dans cette thèse par référence aux organismes biologiques. En particulier cela exclut la **méthode comparative** linguistique¹⁵⁹, les méthodes **lexicostatistiques**, les méthodes **stemmatiques**¹⁶⁰ et toute méthode d'inférence historique développée expressément pour étudier l'histoire d'un certain type d'objets culturels par les spécialistes étudiant ce type d'objets. Il ne s'agit bien évidemment pas de minimiser l'importance et les apports de ces méthodes, encore moins par comparaison avec celles issues de l'inférence phylogénétique. Mais ces analyses ne rentrent pas dans mon champ de recherche, d'une part parce qu'elles sont beaucoup trop nombreuses pour que je puisse tenter de les recenser, d'autre part parce qu'elles appartiennent à des disciplines qui me sont presque complètement étrangères et enfin parce qu'elles ne rentrent pas dans la problématique de la transposition méthodologique interdisciplinaire qui constitue l'un des axes de cette thèse¹⁶¹.

151. ↑ Souvent la première publiée, mais cela peut aussi être la seule à laquelle j'ai eu accès.

152. ↑ Voir partie 2.3.3.1 page 84.

153. ↑ Voir partie 2.3.3.2 page 86.

154. ↑ Voir partie 2.3.3.3 page 87.

155. ↑ Voir partie 2.3.3.4 page 88.

156. ↑ Voir partie 2.3.3.5 page 92.

157. ↑ Pages 84–92.

158. ↑ J'ai tenté de recenser un nombre aussi complet que possible de publications. Cependant la liste présentée ne pourrait prétendre être exhaustive. D'une part certaines publications me sont très probablement inconnues. D'autre part certaines publications dont j'ai connaissance et qui rentrent dans les critères présentés ci-dessus n'ont pas été incluses tout simplement parce que je n'ai pas eu le temps de les étudier ou que je n'y ai pas eu accès.

159. ↑ Voir partie 1.3.3 pages 26 à 29.

160. ↑ Voir partie 1.3.4 pages 29 à 31.

161. ↑ Voir partie 9.4 pages 306 à 313.

2. Des données culturelles. Ce deuxième critère exclut évidemment les études relevant de la systématique biologique et de la biologie de l'évolution et basées sur les caractéristiques biologiques des individu-e-s, la culture étant pensée ici comme un système collectif de pratiques et de réalisations qui ne sont pas transmises biologiquement. J'ai exclu aussi d'éventuelles études concernant les cultures animales¹⁶² car elles relèvent traditionnellement de la biologie et s'écartent donc de la question de la transposition interdisciplinaire¹⁶³. Une description plus détaillée des différents types de données culturelles utilisées dans ces analyses est donnée dans la partie 2.3.3.2 page 86.

3. Des relations comme résultat. Ce troisième critère vise à ne garder que les analyses dont les résultats dus à l'utilisation de la *méthode issue de l'inférence phylogénétique* s'expriment sous forme de relations entre les objets étudiés. Ces relations sont généralement représentées sous la forme d'un arbre ou d'un réseau. La construction de ces relations ne constitue pas nécessairement le but final de l'analyse¹⁶⁴. Par exemple, l'étude de Le Bomin et coll. [240] cherche à tester l'existence d'un signal de transmission verticale dans les données d'ordre musical analysées. Les relations tracées entre les *patrimoines musicaux* par l'application de la méthode *cladistique* ne constituent donc pas l'objectif de l'analyse, mais justifient néanmoins l'inclusion de cette analyse au tableau 2.1. Les relations ainsi constituées ne sont pas non plus nécessairement des relations historiques¹⁶⁴. Par exemple les arbres et réseaux publiés par Toussaint [419] visent à classer les rythmes étudiés et ne représentent donc pas des hypothèses sur l'histoire de ces rythmes¹⁶⁵. Cette restriction vise en particulier à exclure les *méthodes comparatives phylogénétiquement informées*, qui sont des *méthodes issues de l'inférence phylogénétique* et souvent utilisées pour analyser des données culturelles (satisfaisant donc les deux premiers critères), mais qui représentent une trop grande quantité de publications pour que j'ai pu en analyser un nombre significatif.

2.3.3 Description détaillée des champs du tableau

2.3.3.1 Objectifs des analyses

Pour chaque analyse satisfaisant aux critères énoncés ci-dessus, je me suis intéressé au but qu'elle poursuit. L'application de la *méthode issue de l'inférence phylogénétique* peut n'être qu'une étape de l'analyse et c'est bien le but final de cette analyse (et non seulement celui de l'application de la méthode) qui est décrit ici. De plus, l'objectif d'un travail scientifique est souvent multiple et une analyse peut avoir des apports secondaires qui ne sont pas directement recherchés. J'ai donc identifié le ou les objectifs principaux à partir des buts mis en valeurs dans

162. ↑ Je n'ai pas connaissances d'applications de *méthodes issues de l'inférence phylogénétique* sur des données concernant les cultures animales, mais je n'ai pas exploré la littérature scientifique dans ce but et j'ai donc pu passer à côté de telles études.

163. ↑ Voir supra page 83.

164. ↑ Comme expliqué dans la partie 2.3.3.1 page 84.

165. ↑ Même si les travaux ultérieurs de Toussaint et de ses collègues vont dans ce sens, comme expliqué dans la partie 2.1.5.2 page 71.

le résumé¹⁶⁶ et de l'éventuelle conclusion de la publication. J'ai catégorisé ces objectifs selon la typologie suivante :

Classification. L'analyse cherche à classer les données analysées, à identifier une manière cohérente de les organiser.

Évaluation de méthode. L'analyse cherche à évaluer la capacité d'une méthode à parvenir à un objectif donné (souvent une inférence historique). C'est souvent la **méthode issue de l'inférence phylogénétique** elle-même qui est évaluée. En général la publication comprend plusieurs analyses faites à l'aide de différentes méthodes dont les résultats sont ensuite comparés.

Incertain. Les but de l'analyse ne sont pas clairement exposés ou sont ambigus, ou du moins je n'ai pas été en mesure de les déterminer avec certitude.

Inférence historique. L'analyse cherche à proposer une hypothèse sur une histoire, c'est à dire à effectuer une **inférence historique**. En général le résultat de la **méthode issue de l'inférence phylogénétique** constitue directement l'inférence, à la différence des analyses testant des hypothèses¹⁶⁷, où des hypothèses historiques sont formulées *a priori* et le résultat de la méthode sert à tester ces hypothèses. J'ai distingué deux types d'inférence historique :

- **Relations.** L'analyse cherche à retracer des relations historiques entre les entités comparées. Le plus souvent ces relations prennent la forme d'un arbre, mais parfois aussi celle d'un réseau. Parfois un arbre des relations déjà existant est utilisé et sert de point de départ pour inférer des relations horizontales (**emprunts**).
- **Autre.** Le but final de l'analyse est de faire une inférence historique, mais ne concernant pas (ou pas seulement) les relations entre les entités. Il peut s'agir par exemple de l'inférence de dates de divergence entre groupes, de scénarios évolutifs (migration, ...), de l'évolution de certains **caractères**, de celle de certains paramètres évolutifs (par exemple de vitesses d'évolution [332]) ou encore d'inférence sur la localisation passée de certains groupes.

Pédagogie. L'analyse a une visée pédagogique, notamment pour montrer le fonctionnement et l'application de la **méthode issue de l'inférence phylogénétique**.

Test d'hypothèse. L'analyse cherche à tester une hypothèse ou à comparer les prédictions de plusieurs hypothèses. Ces hypothèses portent généralement sur l'histoire et ces analyses visent donc elles aussi à l'**inférence historique**. Cependant, contrairement aux analyses visant directement l'inférence historique¹⁶⁸, l'inférence n'est pas le résultat immédiat de l'application de la méthode.

166. ↑ Ou de l'introduction s'il n'existe pas résumé (*abstract*, en anglais) dans la publication.

167. ↑ Voir ci-dessous la catégorie *Test d'hypothèse*.

168. ↑ Voir supra la catégorie *Inférence historique*.

Visualisation de données. L'analyse cherche à produire une représentation des données permettant une visualisation de celles-ci porteuse de sens ou ayant une valeur heuristique. Il s'agit généralement d'une représentation sous forme d'**arbre** ou de **réseau**.

2.3.3.2 Données analysées

Il s'agit de la nature des caractères utilisés pour effectuer les analyses. Cette catégorisation est parfois problématique car les différents caractères utilisés dans une même analyse peuvent être de nature très variée. Ainsi, pour ne donner qu'un exemple, l'analyse de Fisler et Lecointre [141] cherchant une classification des notions d'**arbre** décrites dans différents ouvrages de systématique biologique inclut aussi bien des caractères concernant les méthodes utilisées que les significations portées par ces notions, ou encore le contenu de ces arbres. Je groupe grossièrement ces différents caractères sous le terme de « Méthodologie scientifique ». Un des manques de ce tableau est l'absence du type d'objets comparés par ces analyses. Cependant la description du type de caractères utilisés donne bien souvent une indication de la nature de ces objets¹⁶⁹.

Culture matérielle. Les caractères sont des caractéristiques physiques d'objets culturels matériels, comme des mesures des dimensions d'un **artéfact** ou la description de motifs ornementaux.

Lieux variants. Lorsque différentes **variantes** d'une **tradition** de manuscrits sont comparés, les caractères sont des **lieux variants**, c'est-à-dire un emplacement du texte où les différentes variantes présentent des différences.

Méthodologie scientifique. Les caractères sont des éléments de méthodologie ou d'épistémologie des sciences, caractérisant par exemple un-e auteur-e ou une publication scientifique.

Mythème. Les caractères sont des **mythèmes**, c'est-à-dire des unités élémentaires de description d'un mythe ou d'une tradition orale qui peuvent être comparées entre plusieurs versions.

Paramètres musicaux. Les caractères sont des **paramètres** musicaux.

Structures linguistiques. Les caractères sont des caractéristiques structurelles des langues étudiées, relatives en particulier à leur système phonologique ou à leur grammaire.

Structures sociales. Les caractères décrivent des pratiques ou des institutions sociales.

Unités lexicales. Les caractères sont des lemmes, c'est-à-dire des unités du vocabulaire de la langue étudiée¹⁷⁰. En général se sont des lemmes que l'on suppose apparentés dans les différentes langues étudiées, on parle alors de **cognats**.

169. ↑ Sauf peut-être dans le cas de la catégorie « Culture matérielle », relativement vague, où les objets peuvent être de nature très différente (**artéfacts** archéologiques, signes d'un système d'écriture, planches de skateboard, ...).

170. ↑ Des « mots » dans le langage courant.

(simulation) Cette précision est ajoutée lorsque les caractères des objets analysés ont été simulés par *évolution artificielle*, soit par l'application d'un modèle informatique d'évolution (simulation numérique [28, 287]) soit par des expériences de *transmission en chaîne* (simulation analogique [361, 373, 392]).

2.3.3.3 Typologie des méthodes utilisées

La colonne *Type de méthode* précise à quelle grande classe de méthodes se rattache chaque analyse. Je distingue principalement trois types de *méthodes issues de l'inférence phylogénétique* : les méthodes *cladistiques*, *phénétiques* et *probabilistes*. Certaines méthodes ne rentrent dans aucune de ces trois catégories et sont donc placées dans la catégorie *Autres*. Ce champ donne donc une idée générale du type de méthode utilisé, la méthode particulière étant définie dans le champ *Algorithme* (voir page 88).

Cladistique. Une méthode est dite *cladistique* si elle cherche des relations *arborescentes* entre les entités étudiées en minimisant le nombre de changements d'état de caractères induit par ces relations. Je conserve donc dans cette définition trois critères principaux : 1) des relations décrites sous forme d'un *arbre* topologique ; 2) l'utilisation directe des *états* de *caractère* (et non d'une mesure dérivée) ; 3) l'utilisation du critère de *parcimonie*. La partie 1.3.6¹⁷¹ explique le principe général d'une méthode cladistique. Selon cette définition, les *méthodes de compatibilité* (ou phylogénie parfaite) peuvent être considérées comme des méthodes cladistiques où le choix de n'accepter aucune incohérence est une forme extrême du critère de parcimonie.

Subcladistique. Je qualifie de subcladistiques les méthodes validant les deux derniers critères énoncés ci-dessus pour les méthodes cladistiques, mais représentant les données sous forme d'un réseau. Ces méthodes ne sont pas habituellement qualifiées de cladistiques, mais on trouve parfois le terme de *réseau cladistique*.

Phénétique. Une méthode d'inférence phylogénétique est dite *phénétique* si elle se fonde sur une mesure de similarité globale entre chaque couple d'entités étudiées pour inférer les relations d'apparentement.

Probabiliste. Une méthode est dite *probabiliste* si elle est fondée sur l'estimation d'une ou plusieurs fonctions de probabilité. La partie 1.3.7¹⁷² donne plus de détails concernant les méthodes probabilistes.

Autre. Les méthodes ne rentrant dans aucune des trois catégories précédentes, décrites plus en détail ci-après.

171. ↑ Pages 34–36.

172. ↑ Pages 36–40.

2.3.3.4 Éléments de caractérisation de ces méthodes

Je donne dans ce paragraphe des éléments de description un peu plus détaillés des **méthodes issues de l'inférence phylogénétique** utilisées dans les analyses du tableau 2.1. Il s'agit souvent d'algorithmes appliquant ces méthodes, mais j'y inclus parfois d'autres détails, comme le choix de certains paramètres lors du lancement de l'analyse. Ces éléments sont regroupés selon le type de méthode décrit dans la partie 2.3.3.3 page 87.

Méthodes cladistiques.

- **Algorithmes exacts.** Les algorithmes exacts donnent toutes les solutions optimales (donc tous les arbres les plus parcimonieux). Il est certain qu'il n'existe aucun autre arbre également ou plus parcimonieux que les arbres trouvés.
 - **Recherche exhaustive**¹⁷³. Cet algorithme construit tous les arbres possibles et ne garde que le ou les arbres les plus parcimonieux. Il est implémenté notamment dans le logiciel *Paup*¹⁷⁴.
 - **Séparation et évaluation**¹⁷⁵ [A188]¹⁷⁶. L'algorithme de séparation et évaluation permet de trouver de manière certaine le ou les arbres les plus parcimonieux sans construire tous les arbres possibles [107, pp. 82–84]. Il est notamment implémenté dans le logiciel *Paup*¹⁷⁷.
- **Algorithmes heuristiques.** Les méthodes de recherche heuristique essaient d'accéder à l'arbre le plus parcimonieux sans parcourir l'ensemble des arbres possibles, trop long à explorer. Il n'est donc pas certain que les arbres obtenus soient les plus parcimonieux, ni qu'ils aient tous été trouvés. Les méthodes suivantes permettent d'effectuer de telles recherches heuristiques :
 - **SPR**¹⁷⁸ : **Élagage et greffe de sous-arbre**¹⁷⁹. Dans cette méthode, certaines portions des arbres les plus parcimonieux sont déplacées à d'autres endroits de l'arbre à chaque étape de la recherche, afin de voir si cela permet de diminuer le nombre de changements d'état nécessaires. Cet algorithme est notamment implémenté dans le logiciel *Mesquite*¹⁸⁰.
 - **TBR**¹⁸¹ : **Réarrangement par bisection et reconnexion**¹⁸². Dans cette méthode, certains arbres les plus parcimonieux sont coupés en deux et reconnectés entre

173. ↑ Darlu et Tassy [107, p. 82].

174. ↑ Marwick [280, p. 167].

175. ↑ *Branch and bound* en anglais.

176. ↑ Citée par Darlu et Tassy [107, pp. 82–84].

177. ↑ Buchanan et Collard [69, p. 374].

178. ↑ *Subtree pruning and regrafting*, en anglais.

179. ↑ Felsenstein [136, pp. 41–44].

180. ↑ d'Huy [121, p. 95].

181. ↑ Pour l'anglais *Tree Bisection and Reconnection*.

182. ↑ Felsenstein [136, p. 44].

eux en différents points à chaque étape de l'analyse afin de trouver des arbres encore plus parcimonieux. Cet algorithme est notamment implémenté dans le logiciel *Paup*¹⁸³.

- **Autres paramètres des analyses cladistiques.** Il existe un grand nombre de variantes des méthodes et de paramètres utilisables dans l'analyse cladistique. Je ne présenterai pas en détail chacun d'entre eux, mais je donne ici les significations, les traductions et les références des éléments précisés dans le tableau 2.1.

- **Parcimonie pondérée**¹⁸⁴.
- **Réarrangement de branches**^{185,186}.
- **SA**¹⁸⁷ : **Ajouts successifs**¹⁸⁸.
- **RSA**¹⁸⁹ : **Séquence aléatoire d'ajouts**¹⁸⁸.
- **Séquences d'ajouts multiples**¹⁹⁰.
- **Compatibilité**¹⁹¹.
- **Compatibilité pondérée**¹⁹².
- **Phylogénie parfaite**¹⁹³ [A219]¹⁹⁴.

- **Méthodes subcladistiques.** Il s'agit des méthodes de type cladistique représentant les relations sous forme d'un **réseau**. Pour la plupart des méthodes listées ci-dessous, je n'ai trouvé que des utilisations ponctuelles et je ne décrirai donc pas en détail chacune d'entre elles. On se référera directement aux publications les utilisant pour plus de détails. Les différentes méthodes utilisées sont les suivantes :

- **Méthode de Bandelt (1994)** [A31]¹⁹⁵.
- **MLN**¹⁹⁶ : **réseau par minimisation des transferts horizontaux** [A104]¹⁹⁷.
- **Réseau phylogénétique parfait**¹⁹⁸ [300].
- **Réseau phylogénétique minimal presque parfait**¹⁹⁹ [132].

183. ↑ Fislser et Lecointre [141, p. 13].

184. ↑ *Weighted parsimony* en anglais.

185. ↑ *Branch swapping* en anglais.

186. ↑ Darlu et Tassy [107, pp. 84–86].

187. ↑ Pour l'anglais *Stepwise Addition*.

188. ↑ Felsenstein [136, pp. 47–48].

189. ↑ Pour l'anglais *Random Sequential Addition*.

190. ↑ *Multiple Addition Sequence* en anglais.

191. ↑ *Maximum compatibility* en anglais.

192. ↑ *Weighted maximum compatibility* en anglais.

193. ↑ *Perfect phylogeny* en anglais.

194. ↑ Citée par Ringe et coll. [349, p. 77].

195. ↑ Citée par Forster et coll. [146, p. 174].

196. ↑ *Minimal lateral network*, en anglais.

197. ↑ Citée par List et coll. [264, p. 45].

198. ↑ *Perfect phylogenetic network* en anglais.

199. ↑ *Minimal almost perfect phylogenetic network* en anglais.

- **Autres.**

- **DWagner**²⁰⁰. Algorithme *DWagner* implémenté notamment dans le logiciel *Phy-sis*²⁰⁰. N'ayant pas accès à la documentation de ce logiciel, je n'ai pu identifier le type de méthode dont il s'agit, mais Carpenter [79, p. 371] la décrit comme une méthode cladistique.
- **Stratocladistique** [A139]²⁰¹. Méthode issue de la paléontologie permettant d'intégrer des informations temporelles dans l'analyse (typiquement des fossiles).

Méthodes phénétiques. Ces algorithmes tracent des relations entre les groupes étudiés à partir d'une **matrice de distances**. Ces algorithmes ne produisent donc pas nécessairement des arbres, certains traçant les relations sous forme d'un réseau.

- **BioNJ** [A158]²⁰². Cet algorithme est un dérivé de l'algorithme de *neighbor joining*, implémenté notamment dans le logiciel *SplitsTree*²⁰³.
- **Fitch**²⁰⁴ [A142]²⁰⁶. Cette méthode est une forme pondérée de la méthode des moindres carrés [107, p. 182]. Elle est implémentée notamment dans le logiciel *Phylip* [135].
- **Kitsch**²⁰⁴. Cette méthode est implémentée dans le logiciel *Phylip* [135].
- **Moindres carrés**²⁰⁵ [A81]²⁰⁶. La méthode basée sur le critère des moindres carrés cherche à trouver l'arbre minisant la somme des carrés des différences entre les distances présentes dans la **matrice de distances** et celles obtenues sur l'arbre. Elle est implémentée notamment dans le logiciel *Paup*²⁰⁷.
- **Plus proche voisin**²⁰⁸ [A365]²⁰⁹. La méthode du **plus proche voisin** est une méthode heuristique qui regroupe les **taxons** séparés par les distances les plus faibles de manière rapide tout en donnant d'assez bons résultats [136, pp. 166–168]. Cette méthode est implémentée notamment dans le logiciel *Paup*²⁰⁷.
- **Réseau de voisins**²¹⁰ [67]²¹¹. Cette méthode, dont la logique est semblable à celle du **plus proche voisin**, construit un **réseau de bipartitions** à partir d'une **matrice de distances**. Elle est notamment implémentée dans le logiciel *SplitsTree* [A206]²¹¹.

200. ↑ Carpenter [79, p. 371].

201. ↑ Citée par Macé et coll. [275, p. 124].

202. ↑ Citée par Thul et Toussaint [414].

203. ↑ Thul et Toussaint [414].

204. ↑ <http://evolution.genetics.washington.edu/phylip/progs.algs.heur.html>

205. ↑ *Least squares*, en anglais.

206. ↑ Citée par Darlu et Tassy [107, p. 182].

207. ↑ Roos et Heikkilä [361, p. 422].

208. ↑ *Neighbor joining* en anglais [107, p. 153].

209. ↑ Citée par Felsenstein [136, pp. 166–168].

210. ↑ *NeighborNet*, en anglais.

211. ↑ Citée par Spencer et coll. [392, p. 506].

- **NetDraw** [A50]²¹². *NetDraw* (tracé de réseau) est une méthode permettant de tracer un réseau à partir d'une **matrice de distances**, notamment implémentée dans le logiciel *UCINET* [A50]²¹².
- **Décomposition en bipartitions**²¹³ [A206]²¹⁴. Cette méthode construit un **réseau de bipartitions** à partir d'une **matrice de distances**. Elle est notamment implémentée dans le logiciel *SplitsTree* [A206]²¹⁴.
- **UPGMA**²¹⁵ [A388]²¹⁶.

Méthodes probabilistes. Les **méthodes probabilistes** sont des méthodes très souples dans lesquelles on peut spécifier de nombreux paramètres, en particulier des modèles d'évolution assez fins. Je ne détaille pas ici les méthodes utilisées par ces analyses à ce niveau de détail. Je me contenterai de rattacher les analyses aux deux grands types d'inférence probabiliste :

- **Inférence bayésienne par MCMC.** Méthode d'**inférence bayésienne** basée sur la **méthode de Monte-Carlo par chaînes de Markov**. Une explication assez détaillée de cette méthode et un exemple d'application sont donnés dans le chapitre 5 pages 155 à 185.
- **Maximum de vraisemblance.** **Méthode probabiliste** basée sur le calcul de la vraisemblance. Cette méthode est décrite rapidement dans la partie 1.3.7.1 page 39.

Autres.

- **dTHGc : détection des transferts horizontaux de gènes complets**²¹⁷ [47]. Méthode permettant d'inférer des transferts horizontaux de gènes à partir de deux arbres, par exemple celui des gènes et celui des espèces en biologie [47].
- **Bipartitions par parcimonie**²¹⁸. Méthode permettant de construire un **réseau de bipartitions** en utilisant un critère de **parcimonie**²¹⁹.
- **Réseau médian réduit**²²⁰ [A32]²²¹. Méthode consistant à représenter un ensemble d'arbres sous forme d'un réseau multidimensionnel les contenant tous, puis de le réduire afin de diminuer sa complexité²²².

212. ↑ Citée par Cochrane et Lipo [92, p. 3890].

213. ↑ *Split decomposition* en anglais.

214. ↑ Citée par Barbrook et coll. [33, p. 839].

215. ↑ Sigle pour *Unweighted pair group method with arithmetic mean* (méthode de groupement non pondérée par paires à partir de la moyenne arithmétique).

216. ↑ Citée par Felsenstein [136, p. 162].

217. ↑ *Horizontal Gene Transfer (HGT) Detection* en anglais.

218. ↑ *Parsimony splits* en anglais.

219. ↑ Roos et Heikkilä [361, p. 422].

220. ↑ *Reduced median network* en anglais.

221. ↑ Citée par Spencer et coll. [392, pp. 2–3].

222. ↑ Spencer et coll. [392, pp. 2–3].

- **Super-réseau**²²³ [A208]²²⁴. Cette méthode combine un ensemble d'arbre en un réseau, permettant de visualiser les conflits entre les différents arbres [444, p. 446]. Elle est notamment implémentée dans le logiciel *SplitsTree*²²⁴.
- **ad hoc**. Méthode *ad hoc*, imaginée et expliquée pour les besoins de l'analyse correspondante.
- **Non précisé**. Les détails de la méthode ne sont pas précisés dans la publication, ou du moins je n'ai pas été en mesure de les identifier.

2.3.3.5 Logiciels utilisés

Cette colonne décrit le logiciel utilisé pour appliquer la **méthode issue de l'inférence phylogénétique**. Je ne décris pas ici chacun de ces logiciels. On se reportera aux publications correspondantes pour les références détaillées à chacun d'entre eux.

223. ↑ *Supernetwork* en anglais.

224. ↑ Citée par Windram et coll. [444, p. 446].

2.3.4 Tableau comparatif

TAB. 2.1 : Publications utilisant des méthodes issues de l'inférence phylogénétique pour analyser des données culturelles.

Référence	Anal.	Objectifs	Données	Type de méthode	Détails méthode	Logiciel
Apple (2014) [7]		inférence historique : autre	lieux variants	cladistique	non précisé	<i>Paup</i>
Atkinson et coll. (2005) [28]	#1	inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>MrBayes, TraitLab</i>
	#2	évaluation de méthode	unités lexicales (simulation)	probabiliste	inférence bayésienne par MCMC	<i>TraitLab</i>
Atkinson (2006) [26] ²²⁵		inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>TraitLab, MrBayes</i>
Barbrook et coll. (1998) [33]	#1	inférence historique : autre	lieux variants	cladistique	non précisé	<i>Paup</i>
	#2	inférence historique : autre	lieux variants	phénétique	décomposition en bipartitions	<i>SplitsTree</i>
Baret et coll. (2006) [34]	#1	évaluation de méthode	lieux variants	phénétique	plus proche voisin	<i>Paup</i>
	#2	évaluation de méthode	lieux variants	cladistique	non précisé	non précisé
	#3	évaluation de méthode	lieux variants	phénétique	décomposition en bipartitions	<i>SplitsTree</i>
	#4	évaluation de méthode	lieux variants	phénétique	réseau de voisins	non précisé
Ben Hamed et coll. (2005) [A42]		évaluation de méthode	structures linguistiques	cladistique	parcimonie pondérée	<i>Paup</i>
Bennett et coll. (2003) [43]		inférence historique : autre	lieux variants	phénétique	non précisé	non précisé
Bland (2005) [46]		inférence historique : autre	lieux variants	phénétique	non précisé	<i>Paup</i>
Boc (2012) [47]		inférence historique : autre	unités lexicales	autre	détection des THG complets (<i>ad hoc</i>)	non précisé
Bouckaert et coll. (2012) [54]		inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	non précisé
Bowern (2010) [56]		inférence historique : autre	unités lexicales	phénétique	réseau de voisins	<i>SplitsTree</i>

225. ↑ pp. 6.1–6.27

TAB. 2.1 : Publications utilisant des méthodes issues de l'inférence phylogénétique pour analyser des données culturelles.

Référence	Anal.	Objectifs	Données	Type de méthode	Détails méthode	Logiciel
Bryant et coll. (2005) [66]		inférence historique : relations	unités lexicales	phénétique	réseau de voisins	non précisé
Buchanan et Collard (2007) [69]		test d'hypothèse	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
Buchanan et Collard (2008) [70]	#1	inférence historique : relations	culture matérielle	cladistique	UPGMA, plus proche voisin	PAST
	#2	inférence historique : relations	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
Buckley (2012) [71]	#1	visualisation de données	culture matérielle	phénétique	réseau de voisins	<i>SplitsTree</i>
	#2	inférence historique : relations, inférence historique : autre	culture matérielle	probabiliste	inférence bayésienne par MCMC	<i>MrBayes</i>
Cabrera et coll. (2008) [72]		incertain	paramètres musicaux	phénétique	non précisé	<i>SplitsTree</i>
Carlson (2004) [78]		inférence historique : autre	lieux variants	cladistique	<i>ad hoc</i>	non précisé
Carpenter (1987) [79]	#1	classification	méthodologie scientifique	phénétique	UPGMA	<i>Physis</i>
	#2	classification	méthodologie scientifique	cladistique	DWagner	<i>Physis</i>
Cochrane et Lipo (2010) [92]	#1	test d'hypothèse	culture matérielle	cladistique	non précisé	<i>Paup</i>
	#2	test d'hypothèse	culture matérielle	phénétique	réseau de voisins	<i>SplitsTree, Ucinet</i>
	#3	test d'hypothèse	culture matérielle	phénétique	NetDraw	<i>Ucinet</i>
Collard et Shennan (2000) [93]		test d'hypothèse	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
Coward et coll. (2008) [97]	#1	test d'hypothèse	culture matérielle	phénétique	réseau de voisins	non précisé
	#2	test d'hypothèse	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
	#3	test d'hypothèse	culture matérielle	probabiliste	inférence bayésienne par MCMC	<i>MrBayes</i>
Currie et coll. (2013) [103]		inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>BayesPhylogenies</i>
Darwent et O'Brien (2006) [108]		évaluation de méthode	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
de Schryver et coll. (2015) [379]	#1	visualisation de données	unités lexicales	phénétique	réseau de voisins	<i>SplitsTree</i>

TAB. 2.1 : Publications utilisant des méthodes issues de l'inférence phylogénétique pour analyser des données culturelles.

Référence	Anal.	Objectifs	Données	Type de méthode	Détails méthode	Logiciel
	#2	inférence historique : relations	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>BayesPhylogenies</i>
d'Huy (2013) [121]	#1	inférence historique : autre, test d'hypothèse	mythème	cladistique	SPR	<i>Mesquite</i>
	#2	inférence historique : autre, test d'hypothèse	mythème	probabiliste	inférence bayésienne par MCMC	<i>MrBayes</i>
	#3	inférence historique : autre, test d'hypothèse	mythème	phénétique	réseau de voisins	<i>SplitsTree</i>
d'Huy (2013) [122]		inférence historique : autre	mythème	phénétique	BioNJ	<i>SplitsTree</i>
d'Huy (2013) [120]	#1	inférence historique : autre	mythème	cladistique	SPR	<i>Mesquite</i>
	#2	inférence historique : autre	mythème	probabiliste	inférence bayésienne par MCMC	<i>MrBayes</i>
d'Huy et Le Quellec (2014) [123]	#1	inférence historique : autre	mythème	phénétique	réseau de voisins	<i>SplitsTree</i>
	#2	inférence historique : autre	mythème	cladistique	non précisé, SPR	<i>Mesquite</i>
	#3	inférence historique : autre	mythème	probabiliste	inférence bayésienne par MCMC	<i>MrBayes</i>
Díaz-Báñez et coll. (2004) [124]		inférence historique : autre	paramètres musicaux	phénétique	non précisé	<i>SplitsTree</i>
Dunn et coll. (2005) [127]	#1	inférence historique : relations	structures linguistiques	cladistique	non précisé	non précisé
	#2	inférence historique : relations	structures linguistiques	cladistique	non précisé	non précisé
Eagleton et Spencer (2006) [128]		inférence historique : autre	lieux variants	phénétique	réseau de voisins	non précisé
Ebach et coll. (2008) [129]		classification	méthodologie scientifique	cladistique	non précisé	aucun
Erdem et coll. (2003) [132]		évaluation de méthode	structures linguistiques, unités lexicales	subcladistique	réseau phylogénétique minimal presque parfait (<i>ad hoc</i>)	<i>ad hoc</i>
Fisler et Lecointre (2013) [141]		classification	méthodologie scientifique	cladistique	TBR	<i>Paup</i>

TAB. 2.1 : Publications utilisant des méthodes issues de l'inférence phylogénétique pour analyser des données culturelles.

Référence	Anal.	Objectifs	Données	Type de méthode	Détails méthode	Logiciel
Forster et coll. (1998) [146]		inférence historique : relations	unités lexicales	subcladistique	méthode de Bandelt (1994)	non précisé
Forster et Toth (2003) [145]		inférence historique : relations, inférence historique : autre	unités lexicales	subcladistique	méthode de Bandelt (1994)	aucun
Fortunato et Mace (2009) [147]		inférence historique : autre	unités lexicales	cladistique	TBR	<i>Paup</i>
García Rivero et O'Brien (2014) [157]		test d'hypothèse	culture matérielle	cladistique	SA, TBR	<i>Paup</i>
Gray et Jordan (2000) [169]		inférence historique : autre	unités lexicales	cladistique	RSA, TBR	<i>Paup</i>
Gray et Atkinson (2003) [168] ²²⁶		inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>MrBayes</i>
Gray et coll. (2009) [166]		inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>BayesPhylogenies</i>
Gray et coll. (2010) [167] ²²⁷		test d'hypothèse	unités lexicales	probabiliste	non précisé	<i>Beast</i>
Greenhill et coll. (2009) [172]		évaluation de méthode	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>TraitLab</i>
Griffith (1968) [174]		inférence historique : autre	lieux variants	phénétique	<i>ad hoc</i>	aucun
Grollemund et coll. (2015) [179]		inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>BayesPhylogenies</i>
Harmon et coll. (2006) [182]			culture matérielle	cladistique, probabiliste	recherche exhaustive, maximum de vraisemblance, TBR, RSA,	<i>Paup</i>
Holden (2002) [196]		inférence historique : relations	unités lexicales	cladistique	RSA, TBR, parcimonie pondérée	<i>Paup</i>

226. ↑ Voir aussi Atkinson et Gray (2006).

227. ↑ pp. 3928–3931

TAB. 2.1 : Publications utilisant des méthodes issues de l'inférence phylogénétique pour analyser des données culturelles.

Référence	Anal.	Objectifs	Données	Type de méthode	Détails méthode	Logiciel
Holden et coll. (2005) [195]	#1	inférence historique : autre	unités lexicales	cladistique	RSA, TBR, parcimonie pondérée	<i>Paup</i>
	#2	inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	non précisé
Holden et Gray (2006) [194]		inférence historique : relations	unités lexicales	phénétique	réseau de voisins	<i>SplitsTree</i>
Hurles et coll. (2003) [205]	#1	pédagogie	unités lexicales	cladistique	non précisé	non précisé
	#2	pédagogie	unités lexicales	phénétique	décomposition en bipartitions	non précisé
Jordan et Shennan (2003) [217]			culture matérielle	cladistique, phénétique	non précisé, plus proche voisin	<i>Paup</i> , non précisé
Jordan et Mace (2006) [215]		inférence historique : autre	culture matérielle, structures sociales	cladistique	non précisé	<i>Paup</i>
Jordan (2007) [214] ²²⁸		inférence historique : relations	unités lexicales	cladistique	RSA, TBR, parcimonie pondérée	<i>Paup</i>
Jordan et Mace (2007) [A213] ²²⁹		inférence historique : relations	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>BayesPhylogenies</i>
Jordan et O'Neill (2010) [216]	#1	inférence historique : autre	culture matérielle	cladistique	SA, TBR	<i>Paup</i>
	#2	inférence historique : autre	culture matérielle	phénétique	réseau de voisins	<i>SplitsTree</i>
Kitchen et coll. (2009) [225]	#1	inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>Beast</i>
Lantin et coll. (2004) [233]		évaluation de méthode	lieux variants	cladistique	TBR	<i>Paup</i>
Le Bomin et coll. (2016) [240]	#1	inférence historique : autre	paramètres musicaux	phénétique	réseau de voisins	<i>SplitsTree</i>
	#2	inférence historique : autre	paramètres musicaux	cladistique	RSA, TBR	<i>Paup</i>
Lee (1989) [252]	#1	évaluation de méthode	lieux variants	cladistique	non précisé	<i>Phylip</i>

228. ↑ pp. 92–99

229. ↑ Voir aussi Jordan (2007, pp. 99–111)

TAB. 2.1 : Publications utilisant des méthodes issues de l'inférence phylogénétique pour analyser des données culturelles.

Référence	Anal.	Objectifs	Données	Type de méthode	Détails méthode	Logiciel
	#2	évaluation de méthode	lieux variants	cladistique	non précisé	<i>MacClade</i>
Lee et Hasegawa (2011) [253]		inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>Beast</i>
		visualisation de données	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>Beast</i>
Lee et Hasegawa (2013) [254]		inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>Beast</i>
Le Quellec (2013) [245]	#1	inférence historique : relations	culture matérielle	phénétique	plus proche voisin	<i>SplitsTree</i>
	#2	inférence historique : relations	culture matérielle	phénétique	réseau de voisins	<i>SplitsTree</i>
	#3	inférence historique : relations	culture matérielle	probabiliste	inférence bayésienne par MCMC	<i>MrBayes</i>
List et coll. (2014) [264]		inférence historique : relations	unités lexicales	subcladistique	MLN	LingPy
Lycett (2007) [268]		test d'hypothèse	culture matérielle	cladistique	TBR	<i>Paup</i>
Lycett (2009) [269]		inférence historique : autre, évaluation de méthode	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
Lycett (2009) [270]		test d'hypothèse	culture matérielle	cladistique	TBR, RSA	<i>Paup</i>
Maas (2008) [273]		évaluation de méthode	lieux variants	cladistique	non précisé	<i>Paup</i>
Maas (2010) [274]		inférence historique : autre	lieux variants	cladistique	non précisé	<i>Paup</i>
Macé et coll. (2001) [276]		classification	lieux variants	phénétique	Fitch, Kitch	<i>Phylip</i>
Macé et coll. (2012) [275]	#1	évaluation de méthode	lieux variants	cladistique	non précisé	<i>Paup</i>
	#2	évaluation de méthode	lieux variants	probabiliste	inférence bayésienne par MCMC	<i>MrBayes</i>
	#3	évaluation de méthode	lieux variants	cladistique	stratocladistique	non précisé
Marwick (2012) [280]		évaluation de méthode	culture matérielle	cladistique	recherche exhaustive	<i>Paup</i>

TAB. 2.1 : Publications utilisant des méthodes issues de l'inférence phylogénétique pour analyser des données culturelles.

Référence	Anal.	Objectifs	Données	Type de méthode	Détails méthode	Logiciel
McMahon et McMahon (2005) [287]		évaluation de méthode	unités lexicales (simulation)	phénétique	non précisé	<i>Phylip</i>
Mora et coll. (2010) [298]		visualisation de données	paramètres musicaux	phénétique	non précisé	<i>SplitsTree</i>
Nakhleh et coll. (2005) [300]		évaluation de méthode	structures linguistiques, unités lexicales	subcladistique	réseau phylogénétique parfait (<i>ad hoc</i>)	autre
Nakhleh et coll. (2005) [301]	#1	évaluation de méthode	structures linguistiques, unités lexicales	phénétique	UPGMA	<i>Paup</i>
	#2	évaluation de méthode	structures linguistiques, unités lexicales	phénétique	plus proche voisin	<i>Paup</i>
	#3	évaluation de méthode	structures linguistiques, unités lexicales	cladistique	non précisé	<i>Paup</i>
	#4	évaluation de méthode	structures linguistiques, unités lexicales	cladistique	compatibilité (<i>ad hoc</i>)	autre
	#5	évaluation de méthode	structures linguistiques, unités lexicales	cladistique	compatibilité pondérée (<i>ad hoc</i>)	autre
	#6	évaluation de méthode	structures linguistiques, unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>MrBayes</i>
Nelson-Sathi et coll. (2010) [306]		inférence historique : relations	unités lexicales	subcladistique	inférence bayésienne par MCMC, MLN	<i>MrBayes</i> , Matlab
O'Brien et coll. (2001) [322]		inférence historique : relations	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
Pagel et Meade (2006) [332]		inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	non précisé
Pagel et coll. (2007) [331]		inférence historique : autre	unités lexicales	probabiliste	inférence bayésienne par MCMC	<i>BayesPhylogenies</i>
Prentiss et coll. (2011) [340]	#1	inférence historique : autre	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
	#2	inférence historique : autre	culture matérielle	phénétique	plus proche voisin	<i>Paup</i>

TAB. 2.1 : Publications utilisant des méthodes issues de l'inférence phylogénétique pour analyser des données culturelles.

Référence	Anal.	Objectifs	Données	Type de méthode	Détails méthode	Logiciel
	#3	inférence historique : autre	culture matérielle	phénétique	réseau de voisins	<i>Paup</i>
Prentiss et coll. (2016) [341]		inférence historique : autre	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
Rexová et coll. (2003) [347]		inférence historique : relations	unités lexicales	cladistique	RSA	<i>Paup</i>
Rexová et coll. (2006) [346]	#1	inférence historique : relations	unités lexicales, structures linguistiques	cladistique	RSA	<i>Paup</i>
	#2	inférence historique : relations	unités lexicales, structures linguistiques	probabiliste	inférence bayésienne par MCMC	<i>MrBayes</i>
Ringe et coll. (2002) [349]		inférence historique : relations	unités lexicales, structures linguistiques	cladistique	phylogénie parfaite	autre
Robinson et O'Hara (1996) [357]		inférence historique : autre	lieux variants	cladistique	non précisé	<i>Paup</i>
Robinson (1997) [354]		inférence historique : autre	lieux variants	cladistique	non précisé	<i>Paup</i>
Rogers et coll. (2009) [358]		inférence historique : autre	culture matérielle	phénétique	plus proche voisin	non précisé
Roos et Heikkilä (2009) [361]	#1	évaluation de méthode	lieux variants (simulation)	cladistique	non précisé	<i>Paup</i>
	#2	évaluation de méthode	lieux variants (simulation)	phénétique	plus proche voisin	<i>Paup</i>
	#3	évaluation de méthode	lieux variants (simulation)	phénétique	moindres carrés	<i>Paup</i>
	#4	évaluation de méthode	lieux variants (simulation)	phénétique	réseau de voisins	<i>SplitsTree</i>
	#5	évaluation de méthode	lieux variants (simulation)	phénétique	décomposition en bipartitions	<i>SplitsTree</i>
	#6	évaluation de méthode	lieux variants (simulation)	autre	bipartitions par parcimonie	<i>SplitsTree</i>
Salemans (1996) [367]		inférence historique : autre	lieux variants	cladistique	non précisé	<i>Paup</i>
Salemans (2000) [366]		inférence historique : autre	lieux variants	cladistique	séparation et évaluation	<i>Paup</i>
Schillinger et coll. (2016) [373]		évaluation de méthode	culture matérielle (simulation)	cladistique	TBR	TNT
Situngkir (2008) [382]		classification	paramètres musicaux	cladistique	non précisé	non précisé
Situngkir (2008) [383]		visualisation de données	culture matérielle	non précisé	non précisé	non précisé

TAB. 2.1 : Publications utilisant des méthodes issues de l'inférence phylogénétique pour analyser des données culturelles.

Référence	Anal.	Objectifs	Données	Type de méthode	Détails méthode	Logiciel
Skelton (2008) [384] ²³⁰		inférence historique : autre	culture matérielle	cladistique	non précisé	<i>Paup</i>
Spencer et coll. (2002) [394]		inférence historique : autre	lieux variants	cladistique	non précisé	<i>Paup</i>
Spencer et coll. (2003) [391]		évaluation de méthode	lieux variants	phénétique	plus proche voisin	<i>Paup</i>
Spencer et coll. (2003) [390]		évaluation de méthode	lieux variants	phénétique	plus proche voisin	<i>Paup</i>
Spencer et coll. (2004) [393]		évaluation de méthode	lieux variants	phénétique	plus proche voisin	<i>Paup</i>
Spencer et coll. (2004) [392]	#1	évaluation de méthode	lieux variants (simulation)	phénétique	plus proche voisin	<i>Paup</i>
	#2	évaluation de méthode	lieux variants (simulation)	cladistique	non précisé	<i>Paup</i>
	#3	évaluation de méthode	lieux variants (simulation)	phénétique	décomposition en bipartitions	non précisé
	#4	évaluation de méthode	lieux variants (simulation)	phénétique	réseau de voisins	<i>SplitsTree</i>
Spencer et coll. (2004) [395]		inférence historique : autre	lieux variants	autre	réseau médian réduit	<i>Network</i>
Tehrani et Collard (2002) [408]		inférence historique : autre	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
Tehrani et Collard (2009) [410]		inférence historique : autre	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
Tehrani et coll. (2010) [411]		inférence historique : relations	culture matérielle	cladistique	séparation et évaluation	<i>Paup</i>
Tehrani (2013) [409]	#1	inférence historique : autre	mythème	cladistique	TBR	<i>Paup</i>
	#2	inférence historique : autre	mythème	probabiliste	inférence bayésienne par MCMC	<i>MrBayes</i>
	#3	inférence historique : autre	mythème	phénétique	réseau de voisins	<i>SplitsTree</i>
Temkin (2004) [412]		évaluation de méthode	culture matérielle	cladistique	non précisé	<i>Paup</i>
Temkin et Eldredge (2007) [413]	#1	évaluation de méthode	culture matérielle	cladistique, phénétique	non précisé, plus proche voisin	<i>Paup, T-Rex</i>
Thul et Toussaint (2008) [414]		visualisation de données, classification (incertain)	paramètres musicaux	phénétique	BioNJ	<i>SplitsTree</i>
Toussaint (2002) [419]	#1	visualisation de données	paramètres musicaux	phénétique	non précisé	autre
	#2	visualisation de données	paramètres musicaux	phénétique	non précisé	<i>SplitsTree</i>

230. † Les éléments comparés sont des signes d'un système d'écriture syllabique, le linéaire B.

TAB. 2.1 : Publications utilisant des méthodes issues de l'inférence phylogénétique pour analyser des données culturelles.

Référence	Anal.	Objectifs	Données	Type de méthode	Détails méthode	Logiciel
Toussaint (2003) [421]		classification, inférence historique : autre (incertain)	paramètres musicaux	phénétique	non précisé	<i>SplitsTree</i>
Tripp (2016) [423]		test d'hypothèse	culture matérielle	cladistique	réarrangement de branches, séquences d'ajouts multiples	TNT
Walker et Ribeiro (2011) [434]	#1 #2	inférence historique : autre inférence historique : autre	unités lexicales unités lexicales	phénétique probabiliste	réseau de voisins inférence bayésienne par MCMC	<i>SplitsTree</i> <i>Beast</i>
Windram et coll. (2008) [444]	#1 #2 #3	évaluation de méthode évaluation de méthode évaluation de méthode	lieux variants lieux variants lieux variants	cladistique phénétique autre	non précisé réseau de voisins super-réseau	<i>Paup</i> <i>SplitsTree</i> <i>SplitsTree</i>
Woerther et Khonsari (2001) [447]		inférence historique : autre	lieux variants	cladistique	non précisé	<i>Phylip</i>

2.4 Conclusion du chapitre

J'ai présenté dans ce chapitre une revue détaillée de l'utilisation de méthodes d'**inférence historique**, issues de la **systématique biologique**, pour étudier des données culturelles. Cette transposition méthodologique concerne des disciplines nombreuses et variées, telles que la **critique textuelle**, la **linguistique historique** mais aussi l'anthropologie de la culture matérielle ou l'ethnomusicologie. De plus ces études ont exploré une grande partie du large éventail de méthodes développées par la systématique biologique et proposent aujourd'hui des développements méthodologiques qui viennent améliorer nos possibilités de comprendre l'histoire, en particulier lorsque celle-ci s'explique mal par un modèle **arborescent**. Cet état de l'art, terminé par un tableau synthétique listant les différents travaux allant dans ce sens, manque sans doute d'une discussion plus générale sur les enseignements que l'on peut tirer de ces expériences interdisciplinaires. Elle offre cependant une vue un peu plus large sur ce qui s'est fait dans ce domaine et de (re)trouver plus aisément des travaux qui pourraient être inspirants, en particulier lorsqu'il s'agit de disciplines éloignées.

En ce qui concerne l'ethnomusicologie, on a vu que les utilisations des **méthode issue de l'inférence phylogénétique** sont en nombre limité. Les études qui s'y emploient souffrent généralement d'importants problèmes théoriques ou méthodologiques. Si l'on peut penser que les **méthode issue de l'inférence phylogénétique** peuvent être intéressantes pour inférer l'histoire des musiques, cela nécessite de pouvoir comparer différentes musiques, ce qui est loin d'être évident. L'ethnomusicologie structuraliste a précisément développé des outils théoriques et méthodologiques spécifiques qui peuvent être utile dans le cadre de telles comparaisons. Je présenterai donc ces outils dans le chapitre suivant²³¹ et je chercherai à montrer en quoi il sont pertinents dans le cadre de l'inférence historique fondée sur la comparaison des structures musicales.

Chapitre 3

Systématique musicale

3.1	Outils théoriques et méthodologiques	107
3.1.1	Les patrimoines musicaux des musiques du Gabon	107
	Pièces	108
	Répertoires	108
3.1.2	Le structuralisme : de la linguistique à l'ethnomusicologie	108
3.1.2.1	Phases de l'analyse	108
3.1.2.2	Identification des unités émiques	109
3.1.3	La modélisation de la systématique musicale vernaculaire	109
3.1.3.1	Modélisation	111
3.1.3.2	Validation	112
3.1.3.3	Techniques d'enregistrement analytique	112
3.1.4	La catégorisation des patrimoines musicaux	113
3.2	Des outils au service de la comparaison interculturelle	114
3.2.1	Modélisation des pièces musicales	114
3.2.1.1	Rigueur de l'analyse	115
	Objectivation	115
	Répétabilité	115
	Cohérence	116
3.2.1.2	Pertinence et finesse de l'analyse	116
	Pertinence	116
	Finesse	117
	Synthèse	117
3.2.2	Catégorisation musicale	117
3.3	Réflexions autour de la notion de <i>modèle</i>	118
3.3.1	La diversification progressive des concepts de cognition musicale	118
3.3.1.1	Simha Arom et la conceptualisation du modèle	118
3.3.1.2	Le substrat : pièces pluripartites et modèle polyphonique	120
3.3.1.3	Les séries hiérarchisées de modèles d'apprentissage	121
3.3.1.4	Vers une diversité des modèles cognitifs	122
3.3.2	Problèmes et perspectives des concepts utilisés	123
3.3.2.1	Pour une ethnomusicologie cognitive	123
3.3.2.2	Modèles instauratif et scientifique	124

3.3.2.3	Perspectives : une approche scientifique, matérialiste et populationnelle	125
3.3.3	Tableau comparatif des concepts de cognition musicale	127
3.4	Conclusion du chapitre	128

L'ethnomusicologie a mis en place au cours du XX^{ème} siècle un certain nombre de méthodes permettant une analyse plus rigoureuse du matériel musical et donnant accès à une connaissance plus fines sur les cultures étudiées. Celles issues de l'étude de la **systematique** et des **systemes categoriels** des patrimoines musicaux se sont révélées particulièrement fécondes. Après avoir présenté rapidement les méthodes permettant de modéliser les règles sous-tendant les réalisation musicale dans une culture donnée¹, je montrerai en quoi ces méthodes sont utiles et nécessaires à la comparaison interculturelle². Je finirai en partageant quelques réflexions sur la notion de *modèle*, concept clé de la pensée structuraliste de Simha Arom³.

3.1 Outils théoriques et méthodologiques

Comparer les patrimoines musicaux de différentes populations nécessite avant toute chose de pouvoir les décrire d'une manière qui les rende intelligible à l'ethnomusicologue et dans une forme qui permette d'établir des équivalences entre cultures différentes. Je ne peux rentrer ici dans le détails de tous les outils théoriques et méthodologiques de l'ethnomusicologie. Aussi je me contenterai de présenter ceux qui me semblent particulièrement importants dans le cadre de ce travail, puisqu'il posent les fondements de la comparaison interculturelle.

J'exposerai d'abord succinctement la manière dont sont organisés les patrimoines musicaux dans les cultures étudiées, en introduisant les notions de *pièce* et de *répertoire*, fondamentales pour comprendre ces musiques⁴. J'évoquerai ensuite quelques principes venus de la linguistique structurale qui ont été transposés à l'analyse structuraliste en ethnomusicologie⁵. Je présenterai ensuite les outils ethnomusicologiques permettant de mettre en évidence les règles présidant à la réalisation des *pièce* musicale (c'est-à-dire leur **systematique musicale**)⁶. Je montrerai enfin comment les outils de la **categorisation musicale** servent à mettre en évidence la manière dont les pièces musicales sont organisées en répertoires⁷.

3.1.1 Les patrimoines musicaux des musiques du Gabon

Je ne rentrerai pas ici dans le détail de la description générale des musiques du Gabon, puisque celles-ci sont présentées dans le chapitre 6⁸. Je me contenterai donc de présenter rapidement la manière dont sont organisés formellement les **patrimoines musicaux** des populations du Gabon. On pourra se référer au chapitre 6 ainsi qu'aux termes du glossaire pour plus de détails.

1. ↑ Partie 3.1 pages 107 à 114.
 2. ↑ Partie 3.2 pages 114 à 118.
 3. ↑ Partie 3.3 pages 118 à 128
 4. ↑ Partie 3.1.1 page 107.
 5. ↑ Partie 3.1.2 page 108.
 6. ↑ Partie 3.1.3 page 109.
 7. ↑ Partie 3.1.4 page 113.
 8. ↑ Pages 187–245.

Pièces. Au Gabon, les **pièces** musicales constituent l'unité musicale fondamentale des **patrimoines musicaux**. Elles sont caractérisées par un certain nombre de **paramètres** musicaux ou extra-musicaux. Ainsi la pièce *Mebounabame*, exécutée par les Ngom du village de Dienga, est caractérisée entre autres par une **métricité** de 12 **pulsations** de subdivision binaire et par la **formule rythmique** « ♪ ♪ ♪ ♪ ♪ »⁹, ainsi que par le texte des chants (d'où provient le nom endogène de cette pièce). Elle est aussi caractérisée par le contexte dans lequel elle est exécutée, à savoir lors de la cérémonie du *Lisimbu* au cours de laquelle sont initiées les jeunes filles [243, p. 283]. Toutes ces caractéristiques confèrent à l'ensemble des réalisations de cette pièce une identité propre, reconnaissable par l'ensemble des interprètes.

Répertoires. Ces pièces sont organisées en ensembles, appelés **répertoires**. La pertinence culturelle de ces ensembles est attestée par l'existence d'un nom endogène ainsi que bien souvent par des **paramètres** musicaux distinctifs. L'ensemble des pièces réalisées au cours de la cérémonie du *Lisimbu* chez les Ngom de Dienga constitue par exemple un répertoire, que l'on nommera lui aussi, par métonymie, le répertoire du *Lisimbu*.

L'analyse structuraliste en ethnomusicologie permet de comprendre l'organisation des pièces et des répertoires au sein d'un patrimoine musical. D'une part en explicitant ce qui fonde l'identité d'une pièce à travers un **modèle** de celle-ci, comme je l'expliquerai dans la partie 3.1.3¹⁰. D'autre part en analysant comment sont organisées les différentes pièces au sein des **répertoires**, ce que j'expliquerai dans la partie 3.1.4¹¹.

3.1.2 Le structuralisme : de la linguistique à l'ethnomusicologie

En introduisant l'analyse structuraliste en ethnomusicologie, Simha Arom a beaucoup emprunté à la linguistique structurale. Je retrace ici succinctement certains de ces concepts issus de la linguistique et leur adaptation à l'ethnomusicologie.

3.1.2.1 Phases de l'analyse

Dans son ouvrage *Polyphonies et polyrythmies instrumentales d'Afrique Centrale*, Arom [13, pp. 215–286] propose en 1984 une théorie structuraliste de l'analyse musicale. Il se fonde pour cela en grande partie sur les travaux de Kenneth Pike et de sa distinction entre les niveaux *étiq*ue et *émiq*ue de l'analyse. Pike [A338]¹² s'inspire de la dialectique des concepts de *phonétique* et *phonémique* en linguistique pour proposer de l'élargir hors du domaine musical. On peut ainsi distinguer plusieurs phases successives de l'analyse structuraliste. Dans un premier temps la description fine des éléments étudiés, par des caractéristiques qui ne sont pas nécessairement

9. ↑ Voir le jeu de données associé au chapitre 6, dans le tableau page 228.

10. ↑ Pages 109–113.

11. ↑ Page 113.

12. ↑ Citée par Arom [13, pp. 243–247].

culturellement pertinentes (niveau *étique*). Dans un second temps l'identification des éléments qui ont une pertinence culturelle qui permet une description simplifiée et utile (niveau *émique*).

3.1.2.2 Identification des unités émiques

L'analyse musicale structuraliste s'appuie ainsi sur les outils théoriques de la linguistique structurale. Le matériau sonore doit tout d'abord être divisé, par *segmentation*, en *unités distinctives pertinentes*. Pour cela, il s'agit d'identifier et de regrouper dans une même *classe paradigmatisée* l'ensemble des éléments qui, dans un même contexte, peuvent se substituer les uns aux autres (c'est-à-dire qui peuvent *commuter*). Les règles mises au point par Chenoweth [A91]¹³, représentées sur la figure 3.1, permettent de déterminer si deux unités étiques doivent être ou non regroupées au sein d'une même unité émique. Par ailleurs les unités et éléments pertinents doivent autant que possible être attestés par des éléments contextuels extérieurs au matériau musical étudié, confirmant leur validité culturelle. Un élément contextuel est d'autant plus susceptible d'attester la pertinence d'un élément musical qu'il a une relation proche à celui-ci, comme le montrent les cercles concentriques de la figure 3.2.

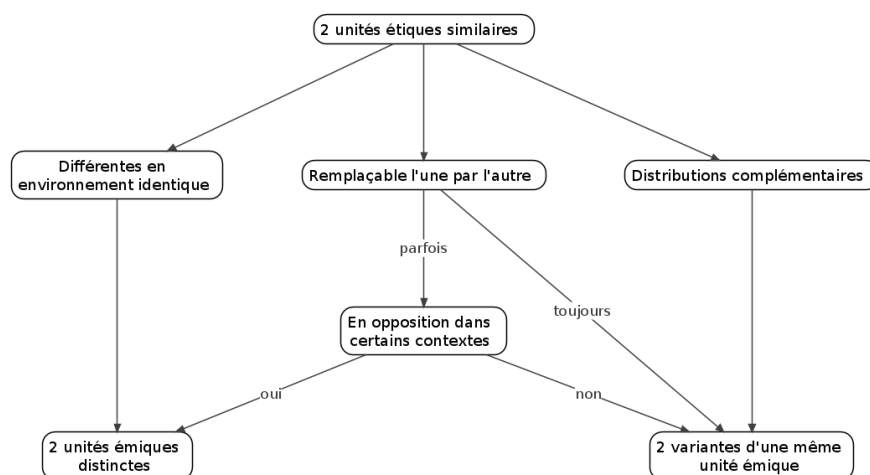


FIG. 3.1 : Règles de regroupement des unités étiques en unités émiques. Ce schéma est réalisé d'après les propositions de Chenoweth [A91]¹³.

3.1.3 La modélisation de la systématique musicale vernaculaire

L'analyse d'une *pièce* musicale vise à extraire le système de règles, explicite ou implicite, qui préside l'ensemble des réalisations de cette pièce, c'est-à-dire à en dégager la *systématique musicale*. La première étape de l'analyse, nécessairement étique, doit servir de base à une analyse

13. ↑ Citée par Nattiez [A302, pp. 228–230], elle-même citée par Alvarez-Pereyre et Arom [3, p. 93].

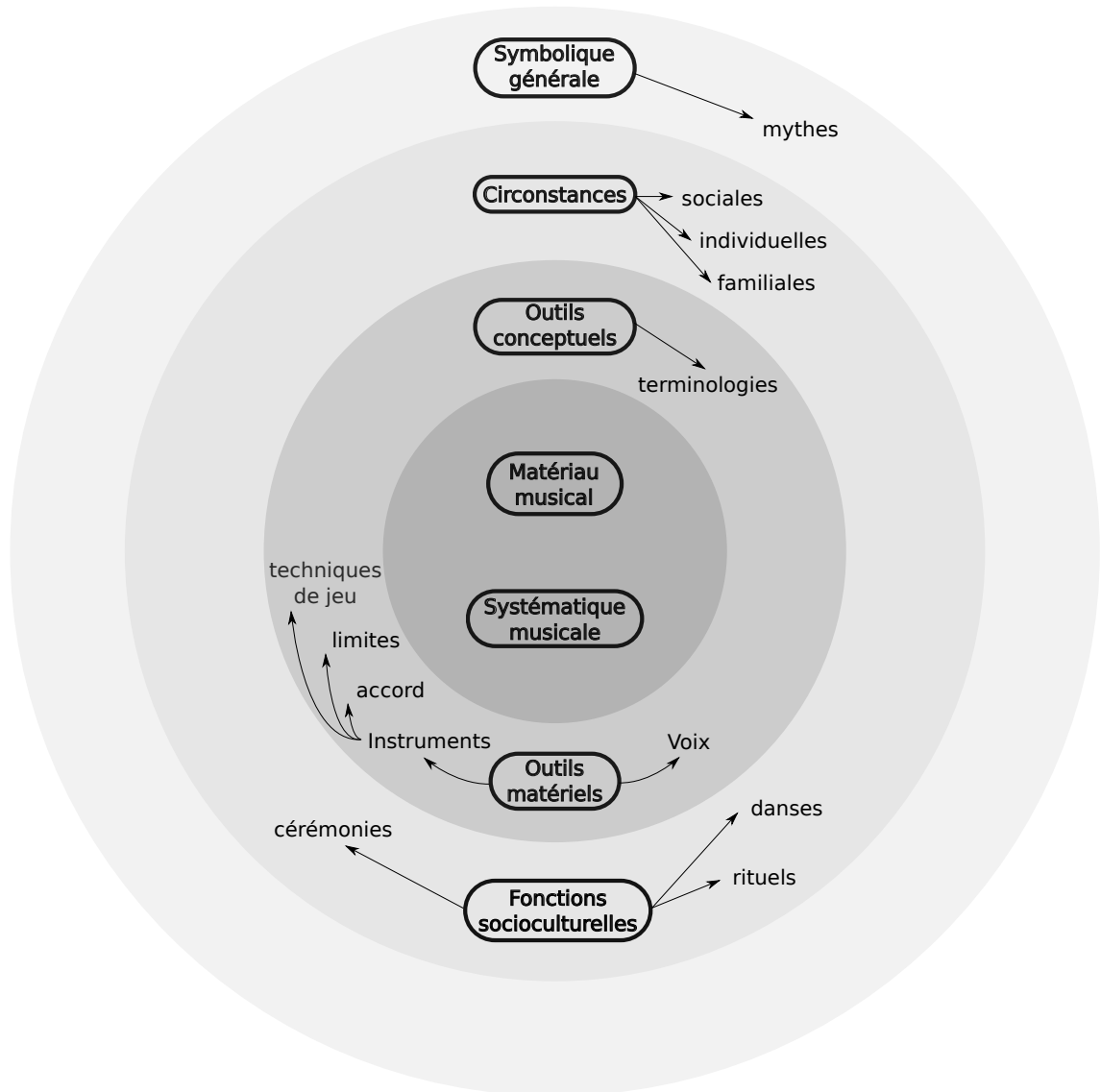


FIG. 3.2 : **Données contextuelles susceptibles d'attester la pertinence de données musicales.** Des données de différents ordres peuvent être plus ou moins utiles lorsqu'il s'agit d'attester la pertinence de données d'ordre musical (matériau musical ou systématique musicale). Dans ce schéma, un certain type de données est d'autant plus utile pour valider un autre type de données que les cercles auxquels ils appartiennent sont proches. La figure a été réalisée à partir des descriptions de Alvarez-Pereyre et Arom [3, pp. 72–74].

émique capable de dégager les unités culturellement pertinentes, aboutissant à une modélisation de cette pièce.

L'analyse structuraliste d'une pièce musicale consiste donc à modéliser la *systématique musicale* de cette pièce, c'est-à-dire l'ensemble de règles présidant à sa réalisation. Cette activité de modélisation de la systématique musicale vernaculaire est parfois elle aussi appelée *systématique musicale*. Afin de différencier les deux sens de ce terme, je parlerai ici de *systématisation musicale* pour parler de l'activité scientifique de modélisation, par analogie avec celui de *catégorisation musicale*¹⁴.

3.1.3.1 Modélisation

Il s'agit donc de modéliser l'ensemble des réalisations d'une *pièce* en une représentation permettant d'en appréhender les principales caractéristiques culturellement pertinentes. Cette représentation est appelée le *modèle* de la pièce¹⁵ :

« Par *modèle*, il faut entendre ici une *représentation sonore à la fois globale et simplifiée d'une entité musicale*; le modèle en accuse la singularité *en ce qu'il condense, en une épure, l'ensemble de ses traits pertinents et eux seuls.* »

Arom (1991) [15, p. 70]

Les stratégies pour modéliser une pièce musicale dépendent des musiques étudiées :

« Dans les musiques de tradition orale, la modélisation peut être obtenue soit par l'émergence — c'est-à-dire la concrétisation — du modèle, soit par sa déduction. »

Arom (1991) [15, p. 70]

Ainsi l'ethnomusicologue peut parfois demander aux interprètes la réalisation d'une *version minimale* de la pièce, simplifiée et épurée par leurs soins, qui peut servir de modèle ethnomusicologique de cette pièce :

« Le premier de ces procédés est naturellement le plus simple et de loin le plus fiable. En effet, dans les populations où la notion de modèle a cours, les musiciens sont le plus souvent à même de concrétiser l'épure sonore qui sert de référence à une pièce ou à l'un de ses niveaux d'apprentissage. On dispose alors de modèles *immanents, concrets.* »

Arom (1991) [15, p. 70]

Dans d'autres cas, les pièces peuvent être réalisées sous des formes spécifiques lors de l'apprentissage, révélant les paramètres jugés définir et caractériser cette pièce pour les tenant-e-s de la culture. Ces réalisations destinées à l'apprentissage sont parfois les mêmes que celles qui sont proposées lorsque que l'ethnomusicologue demande une version minimale :

14. ↑ Voir les entrées *systématique musicale* et *systématisation musicale* du glossaire, ainsi que l'entrée *catégorisation musicale* et le tableau 9.1 page 333 qui y est associé.

15. ↑ La partie 3.3 pages 118 à 128 présente plus en détail la notion de *modèle* et propose une discussion des différents acceptations que peut avoir ce terme.

« Ces modèles constituent bien souvent la base même de la transmission de la connaissance musicale : ainsi, c'est par l'acquisition directe de ces formes simplifiées que les enfants sont familiarisés avec le répertoire traditionnel ; »

Arom (1991) [15, p. 70]

Cependant il n'est parfois pas possible d'obtenir des interprètes une *version minimale*. Dans ce cas seule l'analyse musicologique, via la comparaison des différentes réalisations de la pièce, pourra permettre de proposer un modèle de la pièce :

« Ailleurs, là où une telle procédure est inapplicable, la modélisation ne peut s'appuyer que sur une suite de déductions formelles, aboutissant à l'élaboration d'un modèle théorique. »

Arom (1987) [14, p. 70]

L'analyse musicale structuraliste est déterminante pour comparer les différentes réalisations. À partir des *transcriptions musicales émiques* (n'intégrant donc que les *paramètres* musicaux culturellement pertinents)¹⁶ de chacune des réalisations de la pièce, elle permet de proposer une *partition modélisée* de la pièce.

3.1.3.2 Validation

Mais une fois élaboré ce modèle ethnomusicologique de la pièce, il est nécessaire de vérifier sa pertinence. Puisqu'il s'agit de modéliser la *systématique musicale* vernaculaire, c'est auprès des tenant·e·s de la culture musicale étudiée que l'on pourra tester sa validité. En effet, puisqu'il est censé représenter et condenser les règles présidant à la réalisation musicale, il doit être possible de générer, à partir du modèle obtenu, des réalisations musicales culturellement pertinentes, c'est-à-dire des réalisations que les tenant·e·s de la culture jugent être « la même pièce » :

« Mais il faut alors, afin de vérifier la validité des analyses elles-mêmes, qu'il soit possible, à partir du modèle, de reproduire de nouveaux objets, autrement dit, d'en élaborer de nouvelles réalisations qui, d'un point de vue culturel, soient identiques à l'objet initial.

[...] À la phase déductive succèdera donc une phase inductive, puisque, après être passé de l'objet à son modèle, on doit maintenant procéder en sens inverse : réaliser l'objet à partir du modèle, afin de vérifier que ce dernier a été correctement déduit. »

Arom (1985) [13, pp. 272–273]

3.1.3.3 Techniques d'enregistrement analytique

Les musiques d'Afrique centrale sont remarquables par les relations complexes qu'entretiennent les différentes *parties* vocales et instrumentales, tant du point de vue mélodique que rythmique. Ainsi l'analyse musicale d'une *pièce* à partir d'un *enregistrement conventionnel* est

16. ↑ La *partition étiq*ue, se voulant exhaustive, étant à proscrire.

extrêmement difficile car il est presque impossible de différencier à la seule écoute les différentes parties constitutives.

Simha Arom développe à partir de 1967 une nouvelle méthode d'enregistrement adaptée à l'analyse ethnomusicologique et en particulier à l'étude des polyphonies et des polyrythmies d'Afrique centrale : le *re-recording* [11]¹⁷. Cette technique d'enregistrement analytique permet de séparer l'enregistrement des différentes parties de la pièce sur des pistes audio distinctes, tout en conservant l'information quant à la synchronisation des différentes parties. Le travail de transcription en est alors facilité, en particulier pour les pièces dont l'imbrication des parties est complexe.

La technique consiste à enregistrer d'abord la pièce telle qu'elle est jouée habituellement, appelée « *tutti* ». La première partie est ensuite rejouée et enregistrée seule alors que l'interprète écoute l'enregistrement *tutti* avec un casque audio. Puis la seconde partie est à son tour jouée et enregistrée seule alors que son interprète écoute l'enregistrement de la première partie et ainsi de suite jusqu'à ce que toutes les parties soient enregistrées. Enfin, afin d'établir un étalon temporel commun, la pulsation est enregistrée à partir de l'écoute de chacune des parties enregistrées. Cette technique d'enregistrement sera par la suite améliorée par Susanne Fürniss et Sylvie Le Bomin¹⁸ par une méthode que j'appelle ici l'*enregistrement continu des parties séparées*¹⁹.

3.1.4 La catégorisation des patrimoines musicaux

La compréhension d'une culture musicale passe aussi par l'étude de la place de chaque pièce musicale au sein de l'ensemble du patrimoine musical. En effet les productions musicales ne sont pas autonomes puisqu'elles s'inscrivent dans un contexte social plus global : une pièce donnée est réalisée dans un contexte spécifique, caractérisé par d'autres éléments contextuels, musicaux ou non (chants, danses, éléments symboliques, ...). Ainsi chez les Nkomi du village Amendje, la pièce *Awiri na wiri* est réalisée lors d'une cérémonie d'initiation appelée *Imbwiri yi ighono*, caractérisée par la formule rythmique $\overset{3}{\text{J}}\text{J}\text{J}\text{J}\text{J}$ ainsi que par des symboliques spécifiques données aux objets de la cérémonie [238, pp. 54–57].

Plusieurs pièces sont réalisées lors d'un même contexte d'exécution et constituent un répertoire. C'est un ensemble de pièces pertinent d'un point de vue culturel, puisqu'un répertoire est généralement identifié par un nom endogène et associé à un contexte spécifique de réalisation. Ainsi, chez les Benga du Cap Estérias (Gabon) les pièces nommées *Ibembe renguna*, *Seti aduro oronga* et *Olatano* sont jouées pendant une cérémonie particulière, le *Ndjembe*. Ces trois pièces font donc partie du répertoire du *Ndjembe*.

Arom et coll. [21] proposent un ensemble d'outils théoriques et méthodologiques pour modéliser la manière dont le patrimoine musical d'une population donnée est structuré. Ces outils ont été élaborés à partir de leur expérience de l'analyse des musiques d'Afrique centrale. Il s'agit,

17. ↑ La technique est décrite plus en détail dans l'ouvrage de Arom (1985) [13, pp. 190–213].

18. ↑ Citées par Lechaux [249, p. 82].

19. ↑ Cette technique est décrite dans la partie *Supporting information* (page 221) du chapitre 6, en particulier dans la figure page 223.

via l'analyse musicale des pièces, de mettre en évidence les paramètres musicaux permettant de définir des **catégories** musicales.

Chaque pièce est analysée afin de mettre au jour ses caractéristiques musicales et contextuelles, selon les méthodes présentées ci-dessus²⁰. Ainsi sont définis un certain nombre de **paramètres**, c'est-à-dire des caractéristiques que l'on peut décrire pour chacune des pièces en donnant leur **valeur**. Ainsi la pièce *Mebounabame* jouée par les Ngom du village de Dienga (Gabon) lors de la cérémonie du *Lisimbu* a une forme **responsoriale** : on dit que le **paramètre** « **forme vocale** » prend la **valeur** « responsoriale » dans cette pièce. Il existe une infinité de paramètres possibles et certains ne seront pas pertinents ou opératoires lors de la classification. On sélectionne donc les paramètres adaptés à l'objectif de l'analyse : ce sont les **critères**.

Il s'agit alors de mettre en regard les valeurs de paramètres identifiées dans les différentes pièces des différents répertoires. On cherche à identifier quelle propriété permet de caractériser complètement un répertoire, c'est-à-dire qu'elle permet de regrouper tous les pièces du répertoire, et elles seules. Une telle propriété est appelée un **trait distinctif**. Ainsi, dans l'ensemble des pièces analysées chez les Ngom du village de Moualo, les deux **pièces** du **répertoire** *Makuma* ont une valeur de « huit temps de subdivision binaire » pour le **paramètre** « **métricité** », alors qu'aucune autre pièce ne prend cette valeur pour ce même paramètre. L'organisation de la période en 8 temps de subdivision binaire est donc un **trait distinctif** de la catégorie *Makuma* chez les Ngom du village de Moualo.

3.2 Des outils au service de la comparaison interculturelle

L'intérêt des méthodes d'analyse des **patrimoines musicaux** présentées ci-dessus ne s'arrête pas aux seules description et compréhension d'une culture musicale. En effet elles ouvrent la voie à la comparaison pertinente de différentes cultures musicales. L'étude de Fürniss et Olivier [153] me permettra d'illustrer tout au long de cette partie comment la mise en évidence de la **systématique musicale** de différentes cultures musicale peut servir à la comparaison interculturelle. Les auteures se servent de ces outils pour comparer les musiques des Pygmées Aka (République centrafricaine) et des Bochimans Ju|'hoansi (Namibie) entre lesquelles des scientifiques avaient depuis longtemps noté des ressemblances, mais qui n'avaient jamais été analysées de manière rigoureuse.

3.2.1 Modélisation des pièces musicales

Je détaille dans cette partie en quoi l'analyse structuraliste des **pièces** musicales est un pré-requis à la comparaison interculturelle. D'une part car elle permet une analyse plus rigoureuse, quête de tout travail scientifique sérieux, parce qu'elle lui confère une certaine objectivation, une certaine répétabilité et une certaine cohérence²¹. D'autre part car elle permet d'approfondir

20. ↑ Voir **partie 3.1.3** pages 109 à 113.

21. ↑ **Partie 3.2.1.1** page 115.

l'analyse comparative via une pertinence et une finesse accrues et par sa dimension synthétique²².

3.2.1.1 Rigueur de l'analyse

Objectivation. L'analyse structuraliste permet tout d'abord de définir des critères clairs sur lesquels se fonde la comparaison. C'est un élément indispensable du processus de rationalisation de l'analyse, sans lequel la comparaison ne peut dépasser la dimension au mieux heuristique et au pire impressionniste et non scientifique. Aussi Fürniss et Olivier [153] partent de ce constat comme point de départ de leur analyse :

« [...] le degré de ressemblance attribué à ces deux patrimoines musicaux varie d'un chercheur à l'autre : Charlotte Frisbie (1972 : 287)²³ y trouve "d'immenses similitudes"; Gilbert Rouget affirme que "la ressemblance des musiques bochimán et pygmée était si frappante qu'elle méritait d'être soulignée" (1956 : 1)²⁴, tandis que pour Yvette Grimaud, bien que "la musique des Bochimans et celle des Pygmées ont en commun certains traits généraux, [leur] ressemblance [...] est allée s'estompant à mesure que l'on acquiert une meilleure connaissance des deux musique" (1960 : 1)²⁵. »²⁷

Fürniss et Olivier (1997) [153, p. 10]

Elles proposent, par la suite, l'analyse de la systématique musicale des Aka et des Ju|'hoansi à partir de la description détaillée pour chacune de ces cultures de 11 paramètres (métrique, rythmique, segmentation du cycle, échelles musicales, intervalles mélodiques, plurivocalité, parties constitutives, distribution des parties, références mentales, techniques plurivocales, procédés de variation). Ceci leur permet d'explicitier les ressemblances et différences constatées entre ces deux patrimoines musicaux.

Répétabilité. La rigueur de la procédure de l'analyse structuraliste permet d'objectiver et de minimiser les biais induits par l'analyse, à défaut de pouvoir les éliminer complètement. En effet, les réalisations d'une même pièce au sein d'une même culture peuvent être très variées, selon qu'elle soit produite en contexte ou hors contexte, ou encore qu'elle soit simplifiée à des fins de transmission. Lorsque l'ethnomusicologue demande expressément la réalisation d'une pièce, celle-ci intervient donc bien entendu dans un contexte très particulier. C'est d'autant plus vrai lorsque cette séance a pour finalité l'analyse ethnomusicologique et que l'enregistrement est réalisé via une technique d'enregistrement analytique. Cette modélisation, qui suit une procédure rigoureuse, permet de s'assurer de comparer des objets obtenus au cours d'un processus d'analyse scientifique similaire. C'est tout l'intérêt de l'étude de Fürniss et Olivier dans laquelle, selon leurs mots, « [l]es

22. ↑ Partie 3.2.1.2 page 116.

23. ↑ Frisbie (1971) [A148]²⁶

24. ↑ Rouget (1980) [A362]²⁶

25. ↑ Grimaud (1960) [A177]²⁶

26. Citées par Fürniss et Olivier [153, p. 10].

27. ↑ Les modifications entre crochets sont l'effet de Fürniss et Olivier [153, p. 10]. J'ai rajouté en note de bas de page les références mentionnées par les auteurs.

corpus ont [...] l'avantage d'avoir été constitués d'après la même méthode » [153, p. 10]. En effet si des méthodes d'analyse différentes avaient été suivies dans l'étude des musiques des Aka d'une part et des Ju|'hoansi d'autre part, les paramètres comparés n'auraient pas nécessairement pu être mis en regard.

Cohérence. Le dernier point ne concerne pas directement l'analyse structuraliste, mais y est de fait intimement lié en raison de la tradition de rigueur logique et méthodologique initiée par Simha Arom. Celle-ci veut qu'il existe une cohérence au sein des paramètres analysés. Ainsi lors d'une analyse comparative on s'attend à ce que tous les paramètres de même niveau appartenant au champ de l'analyse soient évalués. Si certains paramètres sont écartés de l'analyse, ce doit être justifié et le choix ne peut être arbitraire. Grauer [165, p. 8] analyse brièvement la ressemblance entre les musiques bochimanes et pygmées en se basant sur les travaux de Lomax [265]²⁸ :

« [T]he Bushmen of the Kalahari Desert and the Pygmies of the tropical forests [...] vocalize in remarkably similar ways. In Cantometric terms, their group performances can be described as interlocked, with maximal vocal blend, polyphony, precisely coordinated rhythms, yodeled, with open, relaxed throats, no embellishment, short phrases, meaningless vocables, etc.. Particularly interesting is the production of a continuous flow of sound, based on the dovetailing of repeated or slightly varied, motives. »²⁹ Grauer (2006) [165, p. 8]

On voit cependant que seuls les paramètres attestant de la ressemblance entre ces musiques sont présentés, à l'exclusion des nombreuses différences pourtant relevées presque 10 ans plus tôt par Fürniss et Olivier [153].

3.2.1.2 Pertinence et finesse de l'analyse

Pertinence. L'analyse de la systématique musicale endogène est fondée sur la notion de pertinence. Seuls les paramètres reconnus comme pertinents par les tenant-e-s de la culture sont conservés comme critères d'analyse. Les paramètres choisis pour la comparaison le sont en fonction des objectifs de l'analyse. Les méthodes d'analyse de la systématique musicale proposent ainsi de définir les paramètres à partir de l'analyse des données de terrain, qui permettront d'identifier les éléments pertinents. Cette approche s'oppose par exemple à celle de Lomax [265] qui définit *a priori* un ensemble de paramètres dont il va chercher à identifier les valeurs pour des pièces provenant de cultures musicales du monde entier.

28. ↑ Cette étude est par ailleurs présentée et discutée plus en détail dans la partie 462 page 142.

29. ↑ « [L]es Bochimans du désert du Kalahari et les Pygmées des forêts tropicales [...] vocalisent de manière remarquablement similaire. Dans le vocabulaire de Cantometric, leurs performances collectives peuvent être décrites comme entremêlées, avec un mélange maximal des voix, de la polyphonie, des rythmes coordonnés avec précision, produites avec la gorge ouverte et relâchée, sans ornementation, des phrases courtes, des syllabes non signifiantes, etc. Il est particulièrement intéressant de noter la production d'un flot sonore continu, basé sur l'imbrication de motifs répétés ou légèrement variés. »

Finesse. La mise au jour de la **systématique musicale** des populations étudiées permet aussi une description plus fine du matériel musical, ouvrant la voie à une comparaison elle aussi plus précise. La technique du *re-recording*, par exemple, permet d'analyser et de comprendre l'imbrication des différentes parties constitutives d'une pièce. Ainsi Fürniss et Olivier [153, p. 23] peuvent-elles remarquer que chez les Aka, les pièces polyphoniques sont issues de quatre parties bien distinctes³⁰, quand chez les Ju|'hoansi le contrepoint est généré par une seule partie. Cette observation leur est fondamentale pour comprendre la différence de nature qui existe entre le contrepoint de ces deux cultures, apparaissant similaire à première vue, mais correspondant à deux conceptions radicalement différentes de la plurivocalité [153, p. 28].

Par ailleurs la multiplication des paramètres d'analyse (et donc des angles d'approche) permet aussi une comparaison plus précise. Loin d'une dichotomie binaire entre des objets semblables et d'autres différents, on peut trouver des convergences sur certains paramètres alors que d'autres divergent. Fürniss et Olivier [153] montrent par exemple que le sentiment de ressemblance entre les musiques Aka et Ju|'hoansi correspond à certaines similarités réelles de ces musiques (pièces majoritairement contrapuntiques, plusieurs procédés de variation similaires, etc.), mais que ces musiques présentent aussi de nombreuses différences, en particulier au niveau conceptuel.

Synthèse. L'analyse structuraliste permet par ailleurs de synthétiser l'information pertinente en une représentation modélisée. Cette représentation permet de s'affranchir de la variation liée à la réalisation de la pièce musicale et qui peut compliquer fortement le travail de comparaison. Cela ne vaut bien entendu que si la comparaison ne porte pas sur la variabilité ou les procédés de variation des cultures considérées. Mais, même dans ce cas, la modélisation est à la base de la comparaison, puisque c'est à partir du modèle que sont établies les variations. C'est ce qui permet à Fürniss et Olivier [153, p. 25] de comparer les procédés de variation des Aka et des Ju|'hoansi.

3.2.2 Catégorisation musicale

On l'a vu, la comparaison des différentes **pièces** d'un **répertoire** est à la base de la **catégorisation musicale**. Mais l'activité de catégorisation peut servir en retour à la comparaison, que celle-ci soit intra- ou inter-culturelle :

« La catégorisation peut être envisagée selon deux acceptions, suivant que l'on considère l'acte de catégoriser en lui-même ou bien le résultat qui en découle. Dans le premier cas, on s'intéresse à l'opération qui consiste à classer les éléments relevant d'un même domaine, qu'il soit technique, linguistique, artistique ou de toute autre nature ; l'accent est alors mis sur les processus qui aboutissent à la classification. Dans le second, c'est le résultat qui tient une place centrale, résultat à partir duquel on peut opérer des comparaisons, que ce soit au sein d'un même domaine comme entre

30. ↑ Ces quatre parties avaient d'ailleurs déjà été identifiées par Simha Arom³¹.

31. Sylvie Le Bomin, communication personnelle.

différents domaines d'une même culture ou, pour un domaine donné, entre différentes cultures. »
Arom et coll. (2008) [21, p. 279]

On voit donc que le résultat de la catégorisation peut être utilisé comme base de comparaison. Les outils de la catégorisation musicale donnent accès à des informations d'un niveau supérieur, concernant la manière dont une culture conçoit et organise ses productions musicales. Cette modélisation scientifique de la catégorisation endogène permet une comparaison de ces conceptions entre plusieurs cultures. Les représentations graphiques des « mondes musicaux » des Aka d'une part et des Ju'hoansi d'autre part proposées par Fürniss et Olivier [153, pp. 11–12] révèlent ainsi les différences profondes dans la manière dont sont organisées ces cultures musicales.

3.3 Réflexions autour de la notion de *modèle*

Dans la lignée des travaux de Simha Arom, l'ethnomusicologie francophone s'intéressant aux musiques d'Afrique centrale consacre une partie importante de ses travaux à modéliser la manière dont les populations étudiées conceptualisent les *pièces* musicales. Les termes consacrés pour désigner ces concepts sont cependant nombreux et variés, chaque ethnomusicologue ayant un terme de prédilection : *modèle* chez Simha Arom, *substrat* chez Susanne Fürniss et Hugo Ferran, *matrice* chez Sylvie Le Bomin, *conception* chez Marie-France Mifune.

Je propose ici une comparaison de ces différents concepts permettant de mieux comprendre les points communs et les différences entre ceux-ci, avant de proposer quelques pistes de réflexions pour affiner et préciser encore les notions utilisées. Ces réflexions s'éloignent un peu du sujet principal de cette thèse. Elles forment cependant une partie non négligeable de mon travail de doctorant et sont sans aucun doute nées de la tension interdisciplinaire qui a orienté en permanence mon travail. En outre la question du *modèle* est, comme on l'a vu, au cœur de la comparaison interculturelle aussi n'est-elle pas dénuée de propos ici.

Comme on le verra, ces différents concepts dérivés de la notion de *modèle* comportent une certaine ambiguïté, à laquelle je n'échapperai pas au cours de cette étude. J'utilise ainsi fréquemment le terme de *modèle* pour désigner l'ensemble de ces concepts qui recourent pourtant des réalités parfois très différentes. Je lui préfère parfois les termes d'*objet cognitif* ou d'*entité cognitive* qui, bien que n'étant pas très satisfaisants non plus, mettent l'accent sur les aspects cognitif et émique, afin de les distinguer du *modèle scientifique* qui est une abstraction relevant du point de vue étique³².

3.3.1 La diversification progressive des concepts de cognition musicale

3.3.1.1 Simha Arom et la conceptualisation du modèle

C'est sans doute Simha Arom qui a le premier exploré les fondements cognitifs de l'interprétation musicale dans l'ethnomusicologie africaniste francophone, en développant la notion de

32. ↑ Voir [partie 3.3.2](#) pages 123 à 126.

modèle ainsi que la méthodologie permettant de le mettre au jour. Dans son ouvrage majeur *Polyphonies et polyrythmies instrumentales d'Afrique Centrale : Structure et méthodologie* [13], il se fonde tout d'abord sur le concept d'*archétype* développé par Constantin Brăiloiu [13, p. 226], ensuite sur les travaux de linguistique structurale, en particulier sur la notion de *code des équivalences* de Roman Jakobson [13, p. 242], enfin sur sa propre expérience des musiques de tradition orale et en particulier des musiques centrafricaines [15, pp. 71–72].

La notion de modèle a tout d'abord permis à Arom dans les années 1970 de traduire un concept vernaculaire des Banda Linda, *àkónè*, utilisé pour désigner une formule musicale très épurée existant pour chaque partie de chaque pièce, et jouée telle quelle uniquement dans un contexte d'apprentissage [15, p. 71]. Simha Arom montre par ailleurs que cette formule est aussi utilisée dans toutes les réalisations de la pièce comme point de départ des variations, mais sans être jamais jouée dans sa forme épurée d'*àkónè* [15, p. 72]. Comme on le verra, les caractéristiques de cette musique auront une certaine influence sur la manière dont sont définis le modèle ainsi que les différents concepts qui en sont issus.

Dans un article reprenant ce concept tel qu'élaboré par la linguistique structurale pour l'appliquer aux musiques d'Afrique centrale qu'il étudie alors, Simha Arom définit ainsi le modèle :

« Dans la perspective ainsi définie, le modèle est l'énoncé minimal, la référence ultime de chaque entité musicale — fut-elle strictement rythmique, mélodico-rythmique ou polyphonique — qui est à l'origine de toutes les réalisations qui en sont culturellement admises. Il constitue la réalisation — attestée ou virtuelle — minimale, la plus épurée, d'une entité musicale et à ce titre, contient à l'état implicite, comme en filigrane, toutes ses virtualités.

Par *modèle*, il faut entendre ici une *représentation sonore à la fois globale et simplifiée d'une entité musicale*; le modèle en accuse la singularité *en ce qu'il condense, en une épure, l'ensemble de ses traits pertinents et eux seuls.*^{C6}

Le modèle équivaut à la réalisation la plus dépouillée d'un morceau, qui puisse être identifié comme tel par les détenteurs de la tradition à laquelle le morceau considéré appartient.^{C7}

Dans les musiques de tradition orale, la modélisation peut être obtenue soit par l'émergence — c'est-à-dire la concrétisation — du modèle, soit par sa déduction. Le premier de ces procédés est naturellement le plus simple et de loin le plus fiable. En effet, dans les populations où la notion de modèle a cours, les musiciens sont le plus souvent à même de concrétiser l'épure sonore qui sert de référence à une pièce ou à l'un de ses niveaux d'apprentissage. On dispose alors de modèles *immanents, concrets.*^{C7}

Ces modèles constituent bien souvent la base même de la transmission de la connaissance musicale : ainsi, c'est par l'acquisition directe de ces formes simplifiées que les enfants sont familiarisés avec le répertoire traditionnel.

Ailleurs, là où une telle procédure est inapplicable, la modélisation ne peut s'appuyer que sur une suite de déductions formelles, aboutissant à l'élaboration d'un modèle théorique.^{C7} »³³ Arom (1991) [15, p. 70]

Cet extrait regroupe bien les différents sens que donne Simha Arom à la notion de modèle : pour lui, c'est à la fois :

- un matériau musical effectivement *réalisé* et *simplifié* utilisé pour la *transmission*,
- un « objet cognitif » possédé par les tenant-e-s de la culture permettant d'une part l'*identification* d'une pièce et servant d'autre part de *référence* lors de la réalisation,
- et le produit d'une analyse scientifique de *modélisation* visant à condenser les traits *pertinents* de la pièce étudiée.

Cette notion est valable pour toute entité musicale³⁴, ce modèle est supposé être immuable et universel dans la population étudiée. On peut remarquer que cette conception première du *modèle* est très générale et fortement influencée par la nature de la musique étudiée par Simha Arom. Comme on le verra, ses successeur-e-s ont repris ce concept tout en l'affinant grâce à l'étude d'autres musiques centrafricaines et éthiopiennes. Simha Arom propose aussi une méthodologie pour mettre au jour ce modèle. Parfois la *pièce* ou la partie musicale étudiée peut être réalisée par les interprètes sous la forme d'une *version minimale* : c'est alors cette version que Simha Arom propose d'utiliser comme modèle. Si en revanche il n'existe pas de réalisation qui puisse constituer un modèle, l'ethnomusicologue doit dégager les caractères culturellement pertinents qui seuls devront être intégrés au modèle. Ces caractères sont identifiés par l'analyse rationnelle du matériau musical, notamment via l'analyse paradigmatique³⁵. Leur pertinence est vérifiée par recoupement avec d'autres données culturelles, endogènes ou exogènes au matériau musical [18]. Le modèle théorique alors construit à partir des caractères pertinents doit ensuite être validé par les tenant-e-s de la culture, par exemple par génération de réalisations artificielles à partir du modèle théorique, lesquelles pourront être acceptées ou invalidées, confortant ou infirmant le modèle proposé³⁶.

3.3.1.2 Le substrat : pièces pluripartites et modèle polyphonique

Comme on l'a vu, Simha Arom parle le plus souvent de *modèle* pour une *pièce* musicale. Il différencie occasionnellement le modèle d'une pièce de celui de ses parties [13, p. 255], mais sans introduire de différence de vocabulaire. De même, la majorité des concepts de cognition musicale

33. ↑ Dans ce passage, Simha Arom compile plusieurs définitions légèrement modifiées issues de divers endroits de deux de ses précédentes publications [13, 14]. Les notes à la fin de chaque paragraphe renvoient vers les extraits correspondant dans ces publications. De manière générale, toutes les notes commençant par C (C1, C2, ...) renvoient aux citations données dans la partie B.1, pages 378 et suivantes.

34. ↑ Même si il distingue déjà en 1985 le modèle des différentes parties et celui de la pièce [13, p. 255], il n'introduit pas de nouveau terme pour les différencier.

35. ↑ Mais parfois complétée par l'analyse acoustique, voir par exemple l'article de Arom et coll. [23, p. 384386].

36. ↑ Voir paragraphe *Validation* page 112.

développés par l’ethnomusicologie africaniste française se réfèrent à la pièce (voir tableau 3.1 page 127).

Dans un article de 1993 concernant les polyphonies vocales des Pygmées Aka, Susanne Fürniss distingue le *patron* de chaque *partie*, « la version la plus épurée et non-variée de cette partie » et qui en est « la référence pour toutes ses réalisations »^{C19} [150, p. 104], du *substrat*, constitué de « l’imbrication des quatre parties constitutives » c’est-à-dire le « patron polyphonique » qui constitue « l’essence du chant polyphonique »³⁷ [150, pp. 116–117]. Le substrat est ainsi l’essence d’une pièce polyphonique possédée par chacun-e des interprètes, obtenu par l’imbrication des *patrons* de chaque partie et qui sert de référence à la réalisation de chaque partie. Ici encore la notion de substrat, définie à partir d’un cas particulier, intègre à la fois la description d’une entité (l’essence de la pièce) la fonction de cette entité (la réalisation de la pièce) et la méthode pour y accéder (l’obtention et l’imbrication des patrons).

Hugo Ferran, dans sa thèse de doctorat, reprend tout en se l’appropriant la notion de substrat de Susanne Fürniss. Pour lui, « [c]ontrairement à l’épure qui est mélodique, le substrat est polyphonique »^{C13} [138, p. 43]. Il reprend donc l’idée d’un modèle qui puisse être polyphonique (et pas seulement mélodique), mais pour l’appliquer à une seule partie d’une pièce polyphonique et non plus à l’ensemble des parties. C’est alors la notion de *canevas* qui prend cette place, « modèle d’une pièce pluripartite » résultant « de l’imbrication des substrats sous-jacents à chacune des parties constitutives »^{C16} [138, p. 483].

3.3.1.3 Les séries hiérarchisées de modèles d’apprentissage

Chez Simha Arom, la notion de modèle a une dimension immuable : pour une *pièce* donnée, il est unique et stable au cours du temps. Vincent Dehoux s’intéressera le premier à la diversité possible des modèles, ouvrant ainsi la voie à l’étude de l’extension et de la pluralité des modèles musicaux. Il montre au cours de son étude des chants à penser accompagnés à la *sanza* chez les Gbaya de République centrafricaine [114] qu’il existe pour chaque pièce quatre modalités d’interprétation différentes de celle-ci, correspondant à différentes techniques de jeu acquises successivement au cours de l’apprentissage de la *sanza*.

Simha Arom, alors qu’il commente le travail de Vincent Dehoux, n’hésite pas à considérer ces modalités comme différents modèles successifs d’une même pièce [15, p. 75]. Les modalités d’interprétation³⁸ mises en évidence par Dehoux se distinguent pourtant de la notion de modèle développée par Arom. En effet les quatre modalités sont communes à toutes les pièces du répertoire et ne sont pas caractéristiques d’une pièce : elles correspondent à un type de jeu plutôt qu’à un modèle de pièce.

Sylvie Le Bomin reprend cette idée au cours de sa thèse de doctorat sur la musique pour

37. ↑ Sauf précision, les mises en forme particulières des citations sont toujours l’effet de leur auteur-e, comme indiqué dans l’annexe A.1 page 373. C’est donc ici Susanne Fürniss qui souligne.

38. ↑ Vincent Dehoux distingue des *modalités d’exécution* qui sont chacune des réalisations particulières d’une pièce, et des *modalités d’interprétation* qui sont les quatre types de jeu communs à toutes les pièces du répertoire qui peuvent être utilisées pour exécuter une même pièce.

orchestre de xylophones des Banda Gbambiya de République centrafricaine [234]. Elle met en évidence trois *modalités de réalisation*³⁹ correspondant à « trois manières de jouer du xylophone soliste » [234, p. 141] communes à l'ensemble du répertoire considéré. Pour chaque pièce, ces trois modalités définissent alors trois *matrices*, « structures sonores strictement répétitives » issues de l'« utilisation stricte des ces modalités » pour la pièce en question [234, p. 141]. Sylvie Le Bomin [234, p. 141] combine donc la notion de modèle de Simha Arom et celle de modalité d'interprétation de Vincent Dehoux pour produire le concept de *matrice*, modèle d'une pièce selon une modalité de jeu particulière auquel se réfère l'interprète au cours de la réalisation^{C24}. Sylvie Le Bomin introduit ainsi explicitement l'idée non seulement que le modèle peut ne pas être unique⁴⁰, mais aussi qu'il peut être modifié au cours de la vie de l'individu.

3.3.1.4 Vers une diversité des modèles cognitifs

Sylvie Le Bomin [237, p. 182] introduit aussi la distinction entre les *références culturelles* (qu'elle assimile aux *matrices*) qui sont collectives et partagées via la transmission et les *références individuelles* ou *énoncés types* qui sont propres à un interprète et sur lesquelles celui-ci se base lors de la réalisation^{C29}. Par ce nouveau raffinement sémantique et conceptuel, elle introduit en fait deux distinctions importantes. D'une part la séparation des fonctions qu'ont ces références pour les tenant-e-s de la culture, en l'occurrence transmission et réalisation⁴¹. D'autre par la séparation entre des références propres soit à *un* individu, soit à *un groupe* d'individu, ce qui comme on le verra ouvre la voie à une approche populationnelle de la notion de modèle.

Dans ce prolongement, et après avoir montré la possibilité d'une modification de la référence individuelle au cours de la vie d'un individu, Sylvie Le Bomin envisage la possibilité d'une évolution culturelle de la référence collective : les matrices de niveau supérieur pourraient être des « énoncés types d'innovations techniques s'étant cristallisés »^{C25} [234, p. 278]. Une référence de nature individuelle, appropriation d'une référence culturelle initiale (la matrice 1), serait ainsi consacrée et établie comme référence de nature culturelle, partagée par toute une population. En proposant un mécanisme d'évolution des références musicales culturelles par interaction entre les niveaux individuels et collectifs, Sylvie Le Bomin s'engage ainsi dans une compréhension évolutive des réalisations musicales observées et des grammaires musicales modélisées.

On voit que le concept de *modèle* initialement proposé par Simha Arom s'est largement diversifié et raffiné au niveau conceptuel mais aussi au niveau sémantique, chaque chercheur-e approfondissant la notion et introduisant de nouveaux termes. Comme on le verra cet enrichissement est fréquemment porteur de confusions, d'une part car le sens des termes utilisés varie entre les ethnomusicologues et, d'autre part, car les notions considérées par un-e même ethnomusicologue englobent de manière plus ou moins implicite des concepts éclectiques voire contradictoires. On ne peut donc que saluer l'effort collectif ayant abouti à des définitions distinctives des termes

39. ↑ Ce terme correspond à celui de *modalité d'interprétation* de Vincent Dehoux.

40. ↑ Validant ainsi l'idée de *série de modèles successifs* de Simha Arom [15, p. 75].

41. ↑ Simha Arom reconnaît aussi occasionnellement au modèle la fonction d'identification des pièces, conjointement à celle de réalisation plus fréquemment évoquée (cf. tableau 3.1 page 127).

*modèle, stéréotype, archétype, prototype, matrice et épure*⁴² [51, p. 61], tout en regrettant la concision de ces définitions et l'absence des discussions qu'elles ne peuvent qu'engendrer.

3.3.2 Problèmes et perspectives des concepts utilisés

3.3.2.1 Pour une ethnomusicologie cognitive

La question des fondements matériels et cognitifs des phénomènes dont les notions dérivées de celle de *concept* cherchent à rendre compte a longtemps été ignorée, ou plutôt évitée, par les ethnomusicologues. En effet, considérant impossible de déterminer si les tenant-e-s de la culture ont conscience d'un *modèle*, les ethnomusicologues préfèrent se contenter de savoir si le modèle en question est explicite, c'est-à-dire qu'il est effectivement matérialisé (ou du moins matérialisable) en une réalisation sonore concrète ou s'il est au contraire implicite et n'est alors jamais matérialisé. Comme le remarque cependant Jean-Jacques Nattiez^{C32} [303, p. 79], l'utilisation chez Simha Arom, (« [t]oute musique ethnique [...] [est] sanctionnée par des règles que sous-tend une théorie. »^{C1}), ou chez Sylvie Le Bomin [237, p. 187] du mot *théorie* suggère que les interprètes auraient une conception consciente bien qu'implicite des objets cognitifs considérés, mais il est difficile d'être certain du sens donné au *théorie*.

Seule Susanne Fürniss donne son avis sur la conscience qu'ont les pygmées Aka d'un tel modèle, lequel se situerait « à un niveau d'abstraction hors conceptualisation verbalisée », suggérant que le substrat conceptualisé (donc pas conscient), même si la notion de « conceptualisation verbalisée » sème le doute, laissant penser que c'est la verbalisation qui permet d'identifier la conceptualisation et mélangeant ainsi la *conscience* et la *réalisation* du modèle. Hugo Ferran, sans se prononcer, aborde frontalement la question de la nature cognitive du modèle en considérant que même si « nous sommes loin de savoir comment ce substrat est inscrit dans l'esprit des musiciens maale, quelle espèce d'objet mental [...] il constitue », « [i]l n'en reste pas moins que ce substrat relève des capacités cognitives des musiciens étudiés »^{C15}.

Cette question, longtemps délaissée, est de plus en plus présente dans les travaux récents, qui se tournent vers une ethnomusicologie cognitive⁴³. Je n'ai pas eu le temps d'approfondir plus cet aspect, aussi je citerai simplement certains travaux qui me paraissent s'engager dans cette voie. Tout d'abord ceux d'ethnomusicologues ou de linguistes travaillant dans la continuité des recherches de Simha Arom tout en approfondissant les aspects cognitifs et psychologiques des pratiques musicales, et en particulier ceux de Marie-Christine Bornes Varol et Susanne Fürniss [52, pp. 475–483] et d'Auréli Helmlinger [187]⁴⁴. Ensuite une partie de l'ethnomusicologie cognitive vient de l'extension de la psychologie de la musique (qui s'est essentiellement consacrée au cours du XX^{ème} siècle à l'étude de sujets de culture occidentale) à d'autres cultures musicales, ces recherches étant portées en particulier par l'université finlandaise de Jyväskylä [131, 229, 417].

42. ↑ Cette synthèse est le fruit des réflexions menées collectivement par les membres du groupe de travail « Processus d'Identification en situation de contact » de l'ex-UMR 8099 Langues-Musiques-Sociétés.

43. ↑ Dont Simha Arom aurait posé les fondements si l'on en croit Jean Molino [A297, p. 198], cité par Hugo Ferran [138, p. 20].

44. ↑ Voir aussi Molino [A297] et Will [A440].

Enfin une dernière approche, initiée par James Kippen [223], vient de l'informatique théorique et de l'ethnomusicologie computationnelle et cherche à représenter les processus cognitifs par des modèles computationnels [40, 89, 223].

3.3.2.2 Modèles instauratif et scientifique

Une confusion traverse cependant la plupart de ces concepts ethnomusicologiques de cognition musicale, portée en particulier par le choix originel de Simha Arom du terme de *modèle*. Comme le montre Bernard Vecchione [428] dès 1991, on peut distinguer deux acceptations complètement différentes du mot. Le modèle *instauratif*, qui a cours dans le domaine des *pratiques* a une fonction *musicale* qui est celle de susciter, de produire, d'« instaurer de la musique » [428, p. 14]. C'est le modèle qui est généralement construit et utilisé par l'interprète en vue de la production musicale. Le modèle *scientifique*, qui est pour sa part utilisé lors de l'activité scientifique, a une fonction *musicologique* d'objectivation et de représentation de la réalité : il sert de fondement à la réfutation. C'est à l'inverse le ou la chercheur-e qui le construit et s'en sert au cours de son activité scientifique.

L'ambiguïté du terme de modèle est déjà présente chez Simha Arom. Si pour lui « le modèle constitue une représentation à la fois globale et simplifiée de l'objet »^{C6} [13, p. 273], on ne sait pas vraiment qui, de l'interprète ou de l'ethnomusicologue, est à l'origine de cette représentation. Lorsque « la modélisation peut être obtenue, soit par l'émergence [...] du modèle, soit par une série de déductions successives. »^{C7} [14, p. 119], c'est bien l'universitaire qui a recours à l'activité scientifique de modélisation. Et le modèle, qui « contient l'ensemble des traits pertinents » et qui sert à « vérifier la validité des analyses elles-mêmes »^{C5} [13, p. 272], remplit bien les fonctions scientifiques (*musicologiques*) de représentation et réfutation. Mais lorsque « [le modèle constitue la] référence mentale sur laquelle est fondée l'exécution de chaque pièce »^{C10} [20, p. 65] c'est bien d'un objet cognitif propre aux interprètes et à fonction *musicale* qu'il s'agit. Et s'il est possible d'obtenir « l'émergence — c'est-à-dire la matérialisation — du modèle » c'est bien qu'il ne s'agit pas d'un construit de l'ethnomusicologue, mais d'un objet cognitif et matériel qui a une existence indépendamment de l'activité scientifique.

Cette confusion perdure dans les termes hérités de la notion de modèle de Simha Arom. Chez Hugo Ferran, le substrat est un outil pour le chercheur, puisqu'il « permet de visualiser en un seul schéma l'ensemble des réalisations possibles »^{C14} [138, p. 214] et, simultanément, « une référence mentale »^{C18} [138, p. 484] des tenant-e-s de la culture « partagé[e] par l'ensemble des habitants d'une même région »^{C14} [138, p. 214]. Pour Susanne Fürniss, le patron d'une partie est « la référence pour toutes ses réalisations »^{C19} [150, p. 104], il s'agit donc bien d'un modèle instauratif. Pourtant dans la pratique de l'analyse le patron désigne la modélisation par la chercheuse (sous forme d'une partition modélisée) de l'objet cognitif servant de référence à l'interprétation [150, partition 6 p. 117].

C'est sans doute dans une formule de Simha Arom que l'on comprend le mieux l'ambiguïté : « la modélisation peut être obtenue [...] par l'émergence — c'est-à-dire la matérialisation —

du modèle »^{C7} [14, p. 119]. En proposant d'utiliser une réalisation particulière (la plus épurée) pour modéliser scientifiquement (et produire un modèle *scientifique*) l'objet cognitif servant de fondement à la réalisation musicale (qui constitue pour l'interprète un modèle *instauratif*) sans distinguer les deux types de modèles par le vocabulaire, il conduit à la confusion voire l'assimilation non seulement entre le modèle scientifique et l'objet modélisé, mais aussi entre l'objet cognitif aux fondements matériels largement inaccessibles et une réalisation sonore particulière issue de celui-ci.

3.3.2.3 Perspectives : une approche scientifique, matérialiste et populationnelle

Différencier le modèle scientifique de l'objet modélisé est essentiel pour plusieurs raisons. Tout d'abord cela permet d'aborder la question du modèle dans une perspective matérialiste. Je pars ainsi d'un postulat selon lequel il existe effectivement des « objets cognitifs » qui, même si nous connaissons fort mal leur nature, sont au fondement de la pratique musicale. Dès lors ces « objets cognitifs » ne peuvent exister qu'au niveau individuel (et non au niveau collectif). Les références collectives⁴⁵ ne peuvent donc plus être considérées comme un objet réel existant indépendamment de l'activité scientifique⁴⁶, mais seulement comme un modèle scientifique construit par l'ethnomusicologue. Cette distinction incite donc à l'étude des objets cognitifs individuels et de leurs relations afin de comprendre le musical dans toute sa complexité, c'est-à-dire à aborder la question de la représentation cognitive d'une manière matérialiste⁴⁷ (et donc excluant tout vision essentialiste), comme interaction des représentations individuelles.

Reconnaître l'utilisation d'un énoncé musical particulier comme modèle scientifique d'un objet cognitif (et non pas comme l'objet cognitif lui-même.) permet aussi d'affiner nos conceptions non seulement des objets cognitifs impliqués dans le musical, mais aussi des modèles scientifiques utilisés pour les modéliser. Concernant les fondements cognitifs de l'activité musicale, il conviendrait de distinguer les différentes fonctions possibles. Ainsi, comme on l'a vu, il y a sans nul doute des fondements cognitifs présidant aussi bien à la *réalisation* musicale, à l'*apprentissage* musical ou encore de l'*identification* des différentes pièces de musique. Si il est probable que les processus cognitifs impliqués dans ces différentes fonctions se recoupent, encore faudrait-il le montrer et il est tout aussi vraisemblable qu'ils ne se superposent pas complètement. Du côté des ethnomusicologues, il faudrait distinguer les différentes méthodes aboutissant à la modélisation (utilisation d'une réalisation particulière comme modèle, construction d'un modèle à partir de plusieurs énoncés, etc.)..

Il revient ensuite à l'ethnomusicologue d'évaluer si tel modèle scientifique modélise bien tel objet cognitif et si plusieurs objets cognitifs aux fonctions différentes peuvent être modélisés par un même modèle. Les écrits de Simha Arom suggèrent par exemple que dans les populations

45. ↑ Par exemple le « substrat, partagé par l'ensemble des habitants d'une même région » chez Hugo Ferran^{C14}, ou la notion de « référence culturelle » chez Sylvie Le Bomin^{C30}.

46. ↑ Dépourvues de substrat matériel, ces références culturelles ne pourraient correspondre qu'à une conception essentialiste et non pas matérialiste de la culture.

47. ↑ Telle que l'a initiée Sylvie Le Bomin dans sa conception de la matrice comme cristallisation d'un énoncé type.

qu'il a étudiées, une réalisation particulière d'une pièce produite par un interprète avec le moins de variations possibles peut être un bon modèle scientifique pour modéliser l'entité cognitive servant de base à cet interprète lors de la réalisation de la pièce, mais aussi pour modéliser l'entité cognitive permettant à cet interprète d'identifier la pièce⁴⁸, suggérant une proximité de ces deux entités cognitives.

48. ↑ Voire, pourquoi pas, pour modéliser l'entité cognitive de chacun-e des membres de la population considérée leur permettant d'identifier la pièce.

3.3.3 Tableau comparatif des concepts de cognition musicale

paramètre	modalité	Modèle Arom	Entité Arom et coll.	Entité Fürniss	Patron Fürniss	Substrat Fürniss	Conception Mifune	Substrat Ferran	Épure Ferran	Canevas Ferran	Entité Ferran	Matrice Le Bomin	Version minimale Le Bomin	Référence Le Bomin	Énoncé type Le Bomin	Épure Le Bomin	Modèle Bornes Varol et coll.	Matrice Bornes Varol et coll.	Épure Bornes Varol et coll.
nature	individuel	?			?		?		?	?				?	x ^{C31}	?	?		?
	partagé	?	x ^{C11}	x ^{C21}	?	x	?	x ^{C14}	?	?	x ^{C17}	x ^{C30}	x ^{C29}	?		?	?	x?	x?
matérialité	réel	x ^{C9}	x ^{C11}	x ^{C21}	x ^{C19}	x ^{C20}	x ^{C33}	x ^{C18}	?	?	?	x ^{C30}	x ^{C29}	x ^{C27}	x ^{C31}	x ^{C26}		x ^{C12}	?
	conceptuel	x ^{C7}						x ^{C14}	?	?	?						x ^{C12}		?
réalisation	concret/explicite				x ^{C19}		?		?	?		x ^{C28}	x ^{C29}	x ^{C27}	x ^{C31}	x ^{C26}		?	?
	abstrait/implicite	x ^{C6}	x ^{C11}	x ^{C21}		x ^{C20}	x ^{C33}	x ^{C18}	?	?	x ^{C17}		x ^{C29}				(x)	?	?
conscience	conscient	?	?	?	?		?	?	?	?	?	?	?	?	?	?	?	?	?
	inconscient	?	?	?	?	x ^{C20}	?	?	?	?	?	?	?	?	?	?	?	?	?
élément musical	modalité de jeu										?	x ^{C28}							
	partie	(x) ^{C4}			x ^{C19}			x ^{C18}	x ^{C13}		?				x				
	pièce	x				x	x ^{C33}				?		x ^{C29}	x ^{C27}		x		x	x
fonction pour les tenant-e-s	groupe de pièces		x ^{C11}	x ^{C21}							?								
	réalisation	x ^{C8}	x ^{C11}	x ^{C21}	x ^{C19}	x ^{C20}	x ^{C33}	x ^{C14}	?	?	x ^{C17}			x ^{C27}	x ^{C31}	non ^{C26}	?		
	transmission								?	?		x ^{C30}	x			x ^{C26}	?	x ^{C12}	
simplifié par l'interprète	identification	(x) ^{C11}							?	?							?		x ^{C12}
					x ^{C19}					?				x ^{C27}			?		x ^{C12}

TAB. 3.1 : Comparaison des concepts de cognition musicale utilisés par différent-e-s ethnomusicologues africanistes français-e-s. Pour chaque paramètre, les différentes couleurs ne servent qu'à mieux visualiser les modalités prises par chaque concept : il n'y a donc pas de point commun entre des couleurs identiques pour des paramètres différents, ni de hiérarchie dans les couleurs. Seules les cellules orange montrent une ambiguïté dans la manière dont le concept considère un certain paramètre. Les notes du type « Cxx » renvoient aux citations de l'annexe B.1 pages 378 à 386. Les informations présentées ici ont été déterminées à partir de ces seules citations. Elles ne visent qu'à donner un aperçu des nuances entre ces différents concepts et demanderaient à être affinées. En particulier, chaque case de ce tableau nécessiterait d'être discutée afin d'exposer plus précisément en quoi telle définition de tel concept se rattache aux modalités proposées.

3.4 Conclusion du chapitre

J'ai présenté dans ce chapitre certains outils fondamentaux de l'ethnomusicologie structuraliste. Je me suis attaché en particulier à exposer les principes de la modélisation des *pièces* musicales et de la *catégorisation musicale*, qui permettent de mettre en évidence la *systematique musicale* des cultures étudiées. J'ai montré pourquoi ces outils étaient à la fois utiles et nécessaires pour une comparaison pertinente de différentes cultures musicales. J'ai aussi discuté de la notion de *modèle* ainsi que certains concepts qui en sont dérivés. Bien que cette discussion s'écarte du cœur du sujet de cette thèse, elle m'a paru intéressante à replacer ici. En effet la notion de *modèle* a un rôle central dans l'ethnomusicologie structuraliste et j'espère que les contradictions que j'ai pu relever dans ces concepts pourront permettre de les affiner.

Enrichi des outils ethnomusicologiques sur lesquels fonder une comparaison interculturelle pertinente d'une part et de méthodes d'*inférence historique* issues de la *systematique musicale* d'autre part, j'aurais pu dès à présent entrer dans le cœur du sujet, à savoir l'analyse des données musicales à ma disposition afin d'en comprendre l'histoire. Ma formation première n'étant pas en ethnomusicologie, il m'a cependant paru maladroit d'attaquer mes analyses sans connaissance de la manière dont cette discipline avait pu, au cours de son existence de plus d'un siècle, aborder la question qui m'intéresse. Je chercherai donc à retracer succinctement dans le chapitre suivant⁴⁹ la manière dont l'ethnomusicologie s'est emparée de la question de l'histoire. Toujours dans cette optique, je m'intéresserai aussi à la dimension comparative en ethnomusicologie et plus particulièrement à la manière dont la comparaison a pu être envisagée comme un outil pour étudier l'histoire.

Chapitre 4

Musique, comparatisme et histoire

4.1	Quelle place pour la comparaison ?	130
	Une musicologie de la comparaison	130
	La comparaison n'est pas la spécificité de la discipline	130
	Critiques et rejets de la comparaison	131
4.2	La comparaison dans la pratique de l'ethnomusicologie	132
4.2.1	Le courant évolutionniste	132
4.2.2	Le courant diffusionniste	133
4.2.3	L'étude des aires musicales	134
	4.2.3.1 Études régionales	134
	4.2.3.2 Le projet <i>Cantometric</i>	134
4.2.4	Diversité des comparaisons	135
	4.2.4.1 La comparaison diachronique	136
	4.2.4.2 Les familles de pièces	137
	4.2.4.3 L'étude de la diversité culturelle	138
	4.2.4.4 La comparaison des comportements musicaux	138
4.3	La comparaison musicale pour l'inférence historique	139
4.3.1	Ethnomusicologie et histoire	139
	Une discipline de la synchronie	139
	Histoire des musiques	139
	La musique porteuse d'histoire	139
4.3.2	L'histoire inscrite au cœur de la musique	140
	L'histoire dans les premières comparaisons ethnomusicologiques	140
	Histoire de pièces	141
	Traces historiques entre cultures proches	141
	Méthodes issues de l'inférence phylogénétique	142
	Musique et migrations	142
4.4	Conclusion du chapitre	143

4.1 Quelle place pour la comparaison ? De la musicologie comparée à l'ethnomusicologie

Une musicologie de la comparaison. Si les premières études des précurseurs de l'*ethnomusicologie*¹ étaient monographiques et peu tournées vers la comparaison [310, p. 150], la discipline naissante place la méthode comparative au cœur de ses préoccupations et de son nom. Dans un article décisif pour l'institutionnalisation de la musicologie en Europe [377, p. 241], Guido Adler [A1]³ définit ainsi la discipline qu'il nomme *vergleichende Musikwissenschaft* (*musicologie comparée*) :

« [...] die vergleichende Musikwissenschaft, die sich zur Aufgabe macht, die Tonproducte, insbesondere die Volksgesänge verschiedener Völker, Länder und Territorien behufs ethnographischer Zwecke zu vergleichen und nach der Verschiedenheit ihrer Beschaffenheit zu gruppieren und sondern. »⁵ Adler (1885) [A1]⁴

Le terme de *musicologie comparée* se répand petit à petit en musicologie, mais sans que les objectifs ou les méthodes de la comparaison ne soient vraiment formalisés [314, p. 73]. Erich von Hornbostel publie en 1905 un article intitulé *Die Probleme der vergleichenden Musikwissenschaft*⁷ [A202]⁸ dans lequel il formalise les buts et les moyens de cette nouvelle discipline. Il insiste sur l'importance épistémologique de la comparaison en dressant un parallèle avec le rôle qu'elle tient en anatomie comparée et en linguistique comparée [377, p. 247]. Il propose aussi une discussion de certains outils de comparaison sans pour autant rentrer vraiment dans la comparaison elle-même [314, pp. 73–74].

La comparaison n'est pas la spécificité de la discipline. Pourtant la plupart des définitions données à la musicologie comparée⁹ par la suite ne mettent plus en avant la méthode comparative, mais une certaine catégorie de musiques étudiée¹⁰ [291, p. 192]. En effet, dès le dé-

1. ↑ Par exemple celles de Joseph-Marie Amiot [A4]², Guillaume André Villoteau [A431]² et Raphael Kiesewetter [A222]².

2. Citées par Nettl [310, p. 150].

3. ↑ Citée par Nettl [314, p. 72].

5. ↑ Alan Merriam [291, p. 199] propose cette traduction en anglais : « [...] comparative musicology has as its task the comparison of the musical works — especially the folksongs — of the various peoples of the earth for ethnographical purposes, and the classification of them according to their various forms. »⁶

6. Voici ma traduction de la traduction proposée par Alan Merriam : « [...] la musicologie comparée a pour objectif la comparaison des productions musicales — en particulier des chansons folkloriques — des différents peuples du monde à des fins ethnographiques, et la classification de celles-ci selon leurs diverses formes. »

7. ↑ « Les problèmes de la musicologie comparée » (ma traduction depuis celle, en anglais, de Richard Campbell : « The Problems of Comparative Musicology » [201, p. 247]).

8. ↑ Citée par Nettl [314, p. 73].

9. ↑ Voir l'entrée de glossaire concernant la *musicologie comparée* pour plus de détails sur la définition de ce terme.

10. ↑ Une exception, parmi les définitions relevées par Alan Merriam, est celle proposée par Helen Roberts [A352, pp. 233–238]¹¹ dans laquelle elle précise que cette discipline s'intéresse aux « [...] exotic musics as compared with one other and with that classical European system [...] »¹².

11. Citée par Merriam [291, pp. 199–200].

12. « [...] les musiques exotiques comparées les unes aux autres et avec le système classique européen [...] »

but des années 1940, l'appellation de *musicologie comparée* est remise en question puisque, comme le constatent les ethnomusicologues, la comparaison n'est pas plus importante dans cette discipline que dans d'autres disciplines scientifiques, y compris musicologiques [291, pp. 192–193]. La comparaison n'est donc pas remise en question, puisqu'elle est considérée comme fondamentale à toute activité scientifique, mais elle perd sa place au sein de l'appellation de la discipline. Jaap Kunst acte ce changement en constatant une fois encore l'inadéquation du nom de la discipline et en proposant l'adoption du terme d'*ethno-musicologie* :

« The name of our science is, in fact, not quite characteristic ; it does not 'compare' any more than any other science. A better name, therefore, is that appearing on the title page of this book : *ethno-musicology*. » ¹⁴ Kunst (1950) [A230, p. 7]¹³

Après une brève période de chevauchement à la fin des années 1950, ce nouveau nom sera rapidement adopté et, à partir de 1961, le terme de *musicologie comparée* ne sera plus utilisé que pour évoquer l'histoire de la discipline et, occasionnellement, pour désigner une sous-discipline de l'ethnomusicologie [291, p. 192].

Critiques et rejets de la comparaison. Les critiques de la démarche comparative elle-même (et non plus seulement du nom de la discipline) ne viennent qu'à partir des années 1960 et sont principalement de deux ordres [291, p. 193]. La première critique, portée notamment par Leonard Meyer [A293, pp. 49–50]¹⁵ et John Blacking [A44, p. 218]¹⁵, vient de ce que les structures musicales n'ont pas les mêmes significations d'une culture à l'autre : les comparer reviendrait à comparer des choses sans rapport entre elles [291, pp. 193–194]. Si ce point de vue porte une quasi-impossibilité de la comparaison, la seconde critique lui reconnaît une légitimité, mais la considère comme prématurée car elle nécessite une bonne connaissance de chaque musique avant de pouvoir les comparer entre elles [291, pp. 193–194] :

« [...] it seems a bit foolish in retrospection that the pioneers of our field became engrossed in the comparison of different musics before any real understanding of the musics being compared has been achieved. » ¹⁷ Hood (1969) [A200, p. 233]¹⁶

La méthode comparative a aussi été écartée du discours des ethnomusicologues de la deuxième moitié du XX^{ème} siècle pour des raisons socio-politiques [314, pp. 82–85]. Tout d'abord, à la sortie de la seconde guerre mondiale, la musicologie comparée est injustement associée à l'idéologie nazie et raciste [314, p. 83]. Par ailleurs la comparaison des musiques populaires a parfois été considérée comme nationaliste, là aussi généralement sans véritable raison [314, p. 84]. Durant

14. ↑ « Le nom de notre science n'est, en fait, pas vraiment spécifique ; celle-ci ne "compare" pas plus que n'importe quelle autre science. Par conséquent un meilleur nom est celui figurant sur la couverture de ce livre : *ethno-musicologie*. »

15. ↑ Citées par Merriam [291, pp. 193–104].

17. ↑ « [...] rétrospectivement il paraît un peu insensé que les pionniers de notre discipline aient pu s'absorber dans la comparaison de différentes musiques avant d'être parvenu à une réelle compréhension de musiques comparées. »

cette période, la comparaison a ainsi perdu la place de prédilection au sein de l’ethnomusicologie que lui avait conférée Guido Adler¹⁸.

4.2 La comparaison dans la pratique de l’ethnomusicologie

Ce revirement théorique n’est pourtant pas si marqué en pratique et la comparaison a toujours eu une place importante dans la pratique de l’ethnomusicologie [314, pp. 85–86]. Cependant, son rôle en tant qu’outil de compréhension est souvent peu explicite et le but comme les moyens de la comparaison sont variables en fonction des études et au cours du temps [310].

4.2.1 Le courant évolutionniste

À la fin du XIX^{ème} siècle et au début du XX^{ème} siècle se développent en Europe des études comparatives à grandes échelles entre des cultures souvent éloignées [310, p. 151], voire sur l’ensemble des musiques du monde. Ces travaux sont largement inspirés par les conceptions évolutionnistes issues de l’anthropologie. Selon ces théories, le développement culturel en général et musical en particulier passe par des stades successifs de développement, identiques pour toutes les cultures. La description des musiques contemporaines, issue de cultures diverses, permettrait d’accéder à différents stades de l’évolution de la musique. Certaines musiques d’aujourd’hui, qualifiées de *primitives*, sont donc censées représenter la musique de l’origine de l’humanité [314, p. 110]. L’ouvrage *Primitive Music*²¹ de Richard Wallaschek [A435]²² propose ainsi une classification des musiques du monde basée sur les « stades de développement ou d’évolution »²³ [377,

18. ↑ On peut cependant noter que Savage et Brown [369] ont récemment proposé de redonner une place centrale à la comparaison en préconisant de retrouver l’usage du terme de *comparative musicology*, avec un sens substantiellement différent de celui qu’il avait au début du XX^{ème} siècle :

« [...] we argue for an inclusive, constructive, and multidisciplinary field that analyzes the world’s musical diversity, from the broadest of generalities to the most culture-specific particulars, with the aim of synthesizing the full range of theoretical perspectives and research methodologies available. »¹⁹
19 20 Savage et Brown [369]

19. [...] nous plaidons en faveur d’une discipline ouverte, constructive et multidisciplinaire qui analyse la diversité musicale du monde entier, des généralités les plus larges aux particularités culturelles les plus spécifiques, ayant pour but la synthèse de tout l’éventail des perspectives théoriques et des méthodologies de recherche disponibles. »

20. Cette citation est issue du chapeau de la version à interligne simple de l’article de Savage et Brown [369], et n’est pas présente dans la version publiée.

21. ↑ *La musique primitive*

22. ↑ Citée par Nettl [310, p. 150] et Schneider [377, p. 247].

23. ↑ Isabelle Mayaud [281] rattache le travail de Wallaschek au courant diffusionniste, mais le passage concernant cet auteur dans l’article de Albrecht Schneider [377, p. 247] me paraît montrer plutôt une tendance évolutionniste que Mayaud caractérise ainsi « [...] les populations primitives qui sont considérées comme les vestiges de la musique occidentale. ». En effet Wallaschek considère que certains groupes ethniques « still preserve the primitive stage which represents mankind at the time of the stone age »²⁴ [A436, p. 66]²⁵.

24. « conservent encore le stade primitif qui représente l’humanité à l’époque de l’âge de pierre »

25. Citée par Schneider [377, p. 247]. Cette citation en anglais est probablement la traduction faite par Albrecht Schneider du texte de 1903 (en allemand) de Wallaschek, car je ne l’ai pas retrouvée dans l’édition anglaise de 1893. La titre de l’édition de 1903 indiqué par Schneider est *Anfänge der Musik*, mais il s’agit probablement d’une erreur, l’ouvrage de Wallaschek ayant pour titre *Anfänge der Tonkunst* [A436].

p. 247]. De même *Die Anfänge der Musik*²⁶ de Carl Stumpf [A402]²⁷ propose la reconstitution d'une supposée musique originelle à partir de la comparaison des structures rythmiques, des formes vocales²⁸ et des instruments de musique issus de l'analyse d'observations de terrain et d'enregistrements phonographiques et photographiques [377, p. 244].

Ce courant évolutionniste s'exprime aussi dans l'étude des musiques dites « folkloriques » européennes. Ce sont alors les musiques des populations paysannes qui sont censées donner accès aux origines de la musique savante occidentale [281, p. 41]. Comme le remarque Isabelle Mayaud [281, p. 41], la comparaison est ici pensée de manière **diachronique** : même si ce sont bien des musiques contemporaines qui sont comparées, les ethnomusicologues considèrent ces musiques comme des témoignages de musiques passées et donc qu'il comparent de l'ancien avec de l'actuel.

4.2.2 Le courant diffusionniste

Les théories diffusionnistes de l'évolution de la culture se développent au début du XX^{ème} siècle en anthropologie. Le concept de *cercles culturels*²⁹, élaboré par Leo Frobenius [A149]³⁰ puis par Wilhelm Schmidt³¹ et Fritz Graeber³¹, considère que les traits culturels apparaissent en des points particuliers³² puis diffusent en cercles à partir de ces points. La comparaison est ici **synchronique** : on cherche à expliquer l'histoire des traits musicaux en comparant les musiques de cultures distantes géographiquement [281, p. 41].

Concernant l'ethnomusicologie, cette théorie sera principalement utilisée pour l'étude de l'origine des instruments de musique. Erich von Hornbostel constate que l'accordage de certaines flûtes de pan d'Océanie et d'Amérique latine, ainsi que celui de certains xylophones d'Asie du sud-est et d'Afrique centrale est identique et correspond à ce qu'il appelle le *cycle des quintes soufflées*³³ [398, p. 28]. Il explique cette ressemblance par l'apparition en Chine de ce type d'accordage puis sa diffusion dans les autres cultures où il est observé [398, p. 28]. Curt Sachs [A364]³⁴ propose une reconstitution de l'origine et de la diffusion de tous les instruments de musique à partir de la comparaison de leur distribution observée aujourd'hui, expliquée à l'aide de 23 cercles de diffusion [290, p. 287] [314, p. 76] [398, pp. 29–30].

La diffusion en cercles culturels servira à expliquer d'autres histoires que celle des instruments. Ainsi Marius Schneider [A378]³⁵ retrace l'histoire mondiale de la diffusion des techniques polyphoniques [314, p. 76] et Werner Danckert [A105]³⁶ s'intéresse aux liens entre pentatonisme et structures sociales à partir du même paradigme [314, p. 76].

26. ↑ « *Les commencements de la musique* » (ma traduction de celle, en anglais, proposée par Albrecht Schneider [377, p. 244] : « *Beginnings of Music* »).

27. ↑ Citée par Nettl [310, p. 150] et Schneider [377, p. 244].

28. ↑ Carl Stumpf introduit pour la première fois le terme d'*hétérophonie* [377, p. 244].

29. ↑ *Kulturkreis*, en allemand.

30. ↑ Citée par Merriam [290, pp. 286–287].

31. ↑ Citées par McLeod [286, p. 101].

32. ↑ Il y a donc *monogenèse* des traits culturels, alors que l'évolutionnisme envisage la *polygenèse*.

33. ↑ *cycle of blown fifths*, en anglais.

34. ↑ Citée par Nettl [314, p. 76].

35. ↑ Citée par Nettl [314, p. 76].

36. ↑ Citée par Schneider [A376], elle-même citée par Nettl [314, p. 76].

4.2.3 L'étude des aires musicales

4.2.3.1 Études régionales

Le concept d'*aire culturelle*³⁷ est créé en 1885³⁸, puis développé par les anthropologues Clark Wissler [A446]³⁹, Alfred Kroeber [A228]³⁹ et Harold Driver [A126]³⁹. Il s'agit de délimiter des zones géographiques, les *aires culturelles*, dont les caractéristiques culturelles sont relativement homogènes et se distinguent des zones voisines.

Concernant l'ethnomusicologie, la circonscription de zones géographiques à partir de traits musicaux permet de définir des *aires musicales*⁴⁰ [311, p. 332]. Dans un premier temps, le concept a surtout été utilisé dans les études concernant les amérindiens d'Amérique du Nord [311, p. 332]. Maurice Herzog évoque en 1928 l'utilité de définir des aires musicales pour appréhender et organiser les nombreuses données disponibles concernant leurs musiques [A191]⁴¹. Le concept sera utilisé de manière plus approfondie à partir des années 1930. La première étude détaillée est sans doute celle d'Helen Roberts sur les musiques amérindiennes [A351]⁴². Celle-ci définit des aires musicales à partir de l'étude des styles vocaux d'une part et à partir de celle des instruments d'autre part (les deux types d'aires ne se superposent pas tout à fait)⁴³. Elle synthétise la distribution des styles d'Amérique du Nord en six⁴⁴ zones géographiques, fondées largement sur des aires culturelles (et pas seulement musicales)⁴⁵. Elle décrit par ailleurs en détail les instruments de musique et leur répartition, ce qui lui permet de définir cinq aires « instrumentales » entièrement fondées sur cette répartition⁴⁶. Cette approche sera poursuivie par Bruno Nettl [A308]⁴⁷ qui divise l'Amérique du Nord en six aires musicales, indépendamment d'aires culturelles préétablies [310, p. 156].

La méthode a par la suite été utilisée pour étudier d'autres zones géoculturelles. Ainsi Alan Merriam se base sur les aires culturelles définies par Melville Herskovits pour décrire sept aires musicales en Afrique [A289]⁴⁸.

4.2.3.2 Le projet *Cantometric*

Alan Lomax met en place le projet *Cantometric*, qui s'étalera sur l'ensemble des années 1960 et aboutira à la publication en 1968 de l'ouvrage *Folk Song Style and Culture* [265] qui en présente les résultats. Cette vaste étude établit une grille d'analyse des musiques enregistrées qui

37. ↑ *Culture-area*, en anglais.

38. ↑ Harris [A183, p. 374], cité par Nettl [311, p. 321].

39. ↑ Citées par Nettl [313, p. 177].

40. ↑ *Musical areas*, en anglais.

41. ↑ Citée par Nettl [311, p. 332].

42. ↑ Citée par Nettl [313, pp. 177–178].

43. ↑ Nettl [311, pp. 332–333].

44. ↑ Au cours de deux résumés du travail de Roberts, Bruno Nettl donne un nombre différent d'aires musicales basées sur les styles vocaux : six [311, pp. 332–333] ou huit [313, pp. 177–178].

45. ↑ Nettl [311, pp. 332–333].

46. ↑ Citée par Nettl [313, p. 177].

47. ↑ Citée par Merriam [290, pp. 295–296] et Nettl [311, p. 333].

48. ↑ Citée par Nettl [311, pp. 332–333].

se veut applicable à n'importe quelle production vocale humaine [265, p. 34] afin de pouvoir les comparer entre elles. Cette grille est constituée de 37 paramètres (appelés *lines*) distincts, dont les valeurs sont ordonnées sur un continuum [265, pp. 22–23, pp. 38–74]. Il définit par ailleurs 56 aires culturelles à partir de celles proposées par George Murdock [A299]⁴⁹, comprenant chacune de deux à onze « sociétés » et regroupées en neuf larges régions du monde [265, pp. 29–33]. Un ensemble d'au moins dix pièces vocales⁵⁰ est choisi pour chaque société. Cette sélection vise à refléter la diversité interne à cette société [265, p. 322]. Un ensemble d'environ 40 pièces est ainsi constitué pour chaque aire culturelle [265, p. 75]. La valeur de chaque paramètre est déterminée pour chacune des 3525 pièces [265, p. 322].

Pour chaque aire et pour chaque paramètre, un graphique est tracé montrant la proportion de la quarantaine de pièce prenant chacune des valeurs de paramètres [265, p. 77]. La superposition des graphiques de deux aires culturelles différentes délimitent des surfaces (entre la courbe et l'axe horizontal des abscisses) qui permettent de définir une mesure de similarité entre deux aires. Pour chaque paramètre, la surface commune sous les deux courbes est divisée par la surface totale. Ainsi plus les courbes sont proches, plus la surface commune est vaste et plus la similarité calculée entre ces deux aires culturelles et pour ce paramètre est grande. La similarité totale entre les deux aires est déterminée en calculant la moyenne de la similarité pour chaque paramètre [265, p. 77]. Le calcul de la similarité entre chaque paire d'aire culturelle forme une matrice de similarité, laquelle est utilisée pour chercher des ensembles d'aires proches géographiquement et dont tous les membres ont deux à deux une grande valeur de similarité.

Lomax et coll. [265, pp. 80–81] définissent ainsi six macrorégions stylistiques correspondant à peu près à 1) l'Amérique du Nord, 2) l'Amérique du Sud, 3) le Pacifique insulaire, 4) l'Afrique, 5) l'Europe et 6) l'Asie. Trois aires culturelles ne sont rattachées à aucune macrorégion, et sont donc considérées comme des isolats : l'Australie, l'Asie arctique et l'Inde tribale [265, pp. 80–81].

4.2.4 Diversité des comparaisons

Malgré le tournant du milieu du XX^{ème} siècle marqué par l'abandon du terme de *musicologie comparée* et le rejet (théorique) de la comparaison, celle-ci reste importante dans la plupart des travaux ethnomusicologiques. Mais la fin du XX^{ème} siècle est aussi marquée par une diversité des méthodes et des objectifs de la comparaison.

L'analyse de cette diversité peut utilement s'inspirer du travail de Oscar Lewis [262] qui étudie 248 travaux d'anthropologie publiés entre 1950 et 1954 et qui ont une dimension comparative. Il s'intéresse aux types d'objets comparés, aux zones géographiques concernées, aux méthodes utilisées, mais aussi aux buts poursuivis. Il distingue ainsi différents objectifs [262, p. 264] : établir des lois générales ou des régularités, documenter la variabilité d'un phénomène ou la distribution d'une caractéristique culturelle, reconstruire l'histoire culturelle ou encore tester des hypothèses.

49. ↑ Citée par Lomax [265, p. 75].

50. ↑ Il s'agit de pièces strictement vocales ou voco-instrumentales.

Nettl [310, pp. 152–153] s’inspire de cette méthodologie pour proposer une analyse succincte de 270 articles publiés dans la revue *Ethnomusicology* de 1958 à 1973. Albrecht Schneider [377] s’intéresse lui aussi aux différents buts de la comparaison en ethnomusicologie et note qu’elle peut servir à [377, pp. 248–250] :

- établir des parallèles ethnographiques servant comme illustration (dimension heuristique) ;
- établir des parallèles ethnographiques ayant statut de preuve ;
- identifier des universaux musicaux ;
- identifier des convergences structurelles et/ou fonctionnelles ;
- établir une typologie au sein d’une culture, ou entre plusieurs cultures ;
- inférer des relations historiques ou génétiques.

Dresser une revue détaillée de ces usages très diversifiés de la comparaison en ethnomusicologie dépasse bien entendu la portée de cette thèse. Je me contenterai de présenter ci-dessous quelques tendances dans l’utilisation de la comparaison en ethnomusicologie depuis la fin du XX^{ème} siècle, sans prétendre rendre compte de manière exhaustive de ces pratiques.

4.2.4.1 La comparaison diachronique

Un petit nombre de travaux s’intéresse à la comparaison d’un même objet musical à différentes époques historiques. Ce type de comparaisons, réellement diachroniques, est qualifié par Oscar Lewis [A261]⁵¹ de *restudies*. Pratiquées en anthropologie dans les années 1950 et 1960, ces approches sont plutôt rares en ethnomusicologie ou alors la comparaison y est implicite et peu formalisée [310, pp. 156–157]. Nettl [A309]⁵² compare par exemple la culture musicale des indiens pieds-noirs qu’il a étudiée dans les années 1960 avec les descriptions fournies (entre autres) par Clark Wissler au début du XX^{ème} siècle [310]. Dans une logique similaire, Ruth Katz [A221]⁵³ compare les chants de deux générations différentes dans une culture juive de Syrie, mais la comparaison n’est pas réellement diachronique puisque les chants sont collectés à une même époque. Le fait d’utiliser la musique des personnes les plus âgées pour avoir une idée de la musique jouée il y a plusieurs décennies comporte des biais évidents, puisque les pratiques musicales peuvent se modifier avec l’âge [310, p. 157].

Le travail d’Émeline Lechaux [249] est particulièrement intéressant ici puisqu’il concerne une étude diachronique de musiques gabonaises. Elle compare entre eux trois corpus d’enregistrements⁵⁴ faits par Pierre Sallée en 1966, 1968 et 1969 ainsi que deux enregistrements faits par Sylvie Le Bomin en 2001 d’un même type de cérémonie de *Bwétè* chez les Tsogho du Gabon.

51. ↑ Citée par Nettl [310, p. 156].

52. ↑ Citée par Nettl [310, p. 157].

53. ↑ Citée par Nettl [310, p. 157].

54. ↑ Dont un corpus audio-vidéo.

Une étude détaillée des différents paramètres musicaux de ces enregistrements, et en particulier l'analyse structurale du répertoire Mópòsè, lui permet d'en dégager les constantes et les variantes sur une période de 50 ans. De tels travaux sont évidemment indispensables si l'on veut parvenir à comprendre quel matériau musical peut être utilisé pour étudier l'histoire et par quels moyens il peut l'être. Le travail de Claire Lacombe [232]⁵⁵ concernant les xylophones Fang s'inscrit dans la même perspective.

4.2.4.2 Les familles de pièces

La comparaison peut aussi être utilisée pour étudier différentes versions de pièces, provenant éventuellement de différentes populations⁵⁶. Je détaille ici une sélection de quelques une de ces études car elles portent souvent des questionnements (explicites ou implicites) d'ordre historique.

Cette approche a tout d'abord eu un rôle classificatoire et se développe au tout début du XX^{ème} siècle et jusque dans les années 1960 [314, p. 77]. Frances Densmore [A118]⁵⁷ [A119]⁵⁷ propose par exemple une classification de pièces amérindiennes en onze groupes sur la base de leurs caractéristiques musicales [310, p. 153]. En 1950, Samuel Bayard [A37]⁵⁸ propose une théorie de la *famille de pièces*⁵⁹ qui doit servir de base à la classification des pièces, à l'étude de l'histoire des pièces et aux études stylistiques [98, p. 495]. En effet, Wilhelm Tappert [A407]⁶⁰ suggère déjà en 1890 que les pièces peuvent circuler d'une culture musicale à une autre, qu'il s'agisse de cultures savantes ou populaires [314, p. 77]. Identifier les ressemblances entre différentes pièces devrait ainsi permettre de retrouver l'histoire des pièces et de leur circulation entre différentes populations. Ce type d'études décline à la fin du XX^{ème} siècle, mais continue par exemple à travers les travaux de James Cowdery [A99]⁶⁰.

Le Bomini et Bikoma [238] ont plus récemment amorcé un travail de ce type concernant les musiques de la côte ouest du Gabon. Dans leur étude sur les musiques Myènè, elles décrivent et comparent les musiques de trois populations de l'ouest du Gabon : les Galoa, les Orungu et les Nkomi. À cette occasion, elles comparent entre elles 3 versions de la pièce *Awiri na wiri*⁶¹, enregistrées dans chacune de ces populations [238, pp. 51–58]. Elles comparent aussi deux versions (Galoa et Nkomi) de la pièce *Ologo ma mè*⁶² ainsi que deux versions (Galoa et Orungu) de la pièce *Ologo mbèla*⁶² [238, pp. 63–73]. Elles comparent enfin deux versions (Galoa et Nkomi) de la pièce *Mbumba osaka*⁶³ [238, pp. 73–75]. Ce travail avant tout descriptif permet d'évaluer la

55. ↑ Ce travail est présenté rapidement dans le paragraphe *Enregistrements d'archives* page 281.

56. ↑ La comparaison est aussi largement utilisée pour l'étude des différentes productions musicales au sein d'une même culture. L'analyse paradigmatique, par exemple, est un outil efficace pour comparer entre elles les réalisations successives des différents cycles d'une pièce. Elle permet ainsi d'identifier la référence mentale présidant à la réalisation de la pièce. Je ne détaille pas plus cet aspect, me concentrant sur les comparaisons de pièces provenant de populations différentes.

57. ↑ Citées par Nettl [310, p. 153].

58. ↑ Citée par Cowdery [98, p. 495].

59. ↑ *Tune family*, en anglais.

60. ↑ Citée par Nettl [314, p. 77].

61. ↑ Répertoire *Imbwiri yi ighono*.

62. ↑ Ces deux pièces appartenant au répertoire de l'*Ologo*.

63. ↑ Répertoire du *Mbumba*.

stabilité des différents paramètres musicaux entre des pièces équivalentes dans différentes populations et constitue la base d'une recherche à visée historique qui sera présentée rapidement dans la partie 4.3.2 page 141.

Pièce	<i>Awiri na wiri</i>		<i>Ologo ma me</i>		<i>Ologo mbela</i>		<i>Mbumba osaka</i>	
	Métricité	Échelle	Métricité	Échelle	Métricité	Échelle	Métricité	Échelle
Galoa	8/2		8/2	heptatonique	8/3	hexatonique	8/2	pentatonique
Orungu	8/2				8/3	pentatonique		
Nkomi	8/2		8/2	pentatonique			8/2	heptatonique

FIG. 4.1 : **Comparaison des types d'échelles et des métricités de quatre pièces présentes dans plusieurs populations.** Chaque pièce a été enregistrée et analysée dans au moins deux populations différentes (Galoa, Orungu et/ou Nkomi). Pour chacune des pièces, la **métricité** et type d'**échelle** (caractérisé ici par le nombre de degrés qu'elle contient) ont été identifiés. Ces données sont issues du livre de Le Bomin et Bikoma [238, pp. 49–92].

4.2.4.3 L'étude de la diversité culturelle

La comparaison est aussi utilisée pour décrire la diversité des musiques existantes dans une région plus ou moins grande. Une des premières études de la diversité musicale à l'échelle globale est celle d'Alan Lomax [265]⁶⁴. Plus récemment Rzeszutek et coll. [363] s'intéressent à la diversité des musiques de l'est de l'Asie à partir de questionnements et d'outils statistiques classiquement utilisés dans des disciplines telles que la biologie, mais peu répandus dans les études ethnomusicologiques. À partir des données concernant seize populations de langues austronésiennes, ils comparent ainsi la diversité musicale existante au sein d'une même population (diversité interne) à la diversité musicale observée entre différentes populations (diversité externe) et constatent que la diversité interne est bien plus importante que la diversité externe.

4.2.4.4 La comparaison des comportements musicaux

Une autre tendance récente est celle de la comparaison non pas des **paramètres musicaux internes**, mais des comportements de production de la musique [87, p. 158]. Un important travail sur ce sujet est actuellement porté par la chaire *Geste et acoustique musicale* (GeAcMus) de l'université Paris-Sorbonne. Par une combinaison des méthodes de l'ethnomusicologie et des récentes avancées techniques en matière d'enregistrement et d'analyse du geste, il s'agit de décrire et de comparer les gestes de la pratique instrumentale dans différentes cultures. Fabrice Marandola, par exemple, compare le mouvement des membres antérieurs de xylophonistes du Cameroun lors de la réalisation instrumentale.

64. ↑ Voir partie 4.2.3.2 page 134.

4.3 La comparaison musicale pour l'inférence historique

4.3.1 Ethnomusicologie et histoire

Une discipline de la synchronie. L'ethnomusicologie s'est longtemps détournée de l'histoire et de la dimension diachronique. Peut-être en raison de la rareté des sources habituelles de l'analyse historique dans les terrains de prédilection de l'ethnomusicologie naissante. Mais aussi sûrement pour des raisons idéologiques et méthodologiques : étudiant des sociétés considérées « sans histoire » et dont l'état observé était censé révéler l'histoire de la musique (supposée suivre un développement similaire dans toutes les cultures), l'ethnomusicologie ne pouvait s'intéresser à l'histoire de ces musiques. Par la suite, inspirée d'une anthropologie fonctionnaliste puis structuraliste, l'ethnomusicologie s'est concentrée sur la dimension synchronique [87, p. 16], s'attachant à comprendre et décrire les musiques contemporaines plutôt qu'à étudier leur histoire. Ce détournement de la diachronie est cependant révolu et l'étude de l'histoire des musiques est devenu un sujet incontournable de l'ethnomusicologie.

Histoire des musiques. Il existe une multitude d'approches pour retrouver l'histoire des pratiques musicales. On peut bien entendu passer par les sources classiques de l'analyse historique que sont les documents écrits. Les enregistrements ethnomusicologiques, parfois vieux de plus d'un siècle [380], sont aussi des documents témoignant des musiques passées [151, pp. 16–17], comme dans les travaux déjà présentés d'Émeline Lechaux [249] et de Claire Lacombe [232]. L'enquête auprès des personnes les plus âgées permet elle aussi d'obtenir des informations sur cette histoire. Par ailleurs l'archéologie peut apporter des informations sur les pratiques passées, en particulier en ce qui concerne les instruments de musiques [87, p. 28].

L'ethnomusicologie peut aussi s'appuyer sur les apports de l'ethnohistoire qui, pour retrouver l'histoire des populations, utilise l'histoire telle que racontée par les populations elle-même. En plus des informations historiques telles que rapportées à la demande de l'ethnologue, cela inclut l'analyse des éléments culturels qui peuvent comporter des données d'ordre historique. L'analyse des épopées, récits racontant des faits passés, et des mythes fondateurs, souvent ancrés dans l'histoire des peuples, peut être d'une grande utilité pour étudier l'histoire des musiques [87, p. 26]. L'origine et la fabrication des instruments de musique sont ainsi souvent mythifiées, fournissant des informations sur la transmission de ces instruments [87, p. 27]

La musique porteuse d'histoire. La valeur de la musique comme source historique⁶⁵ est reconnue depuis longtemps, comme en témoignent les rencontres de l'Unesco de 1967–69 concernant son projet d'histoire générale de l'Afrique, en particulier la publication de Alagoa [A2]⁶⁶ nommée *Le chant comme source d'information historique*. Il est intéressant ici de s'arrêter sur la

65. ↑ Mais de quelle histoire la musique peut-elle être une source ? L'histoire des populations ? De leur culture ? De leur musique ? Je discuterai de cette question dans la partie 7.2 pages 251 à 268.

66. ↑ Citée par Charles-Dominique [87, p. 23].

question des textes associés aux pratiques musicales. Le texte du chant, porteur de signification, peut être utilisé à ce titre pour retracer l'histoire d'une population et de sa musique, comme le suggérait déjà Constantin Brăiloiu [A62]⁶⁶ en 1960. Il est un élément charnière, à la fois source ethnohistorique et élément de la pratique musicale pouvant être utilisé pour inférer l'histoire. Dans le même ordre d'idées, puisque le langage tambouriné permet dans certaines cultures de transmettre des significations, il peut être utilisé pour véhiculer une histoire [87, p. 25], à la manière d'une épopée. D'autres paramètres liés au musical peuvent être porteurs d'informations historiques. Ainsi la *philologie organologique* de Pierre Bec [A38]⁶⁷ s'intéresse à ce que la diversité des désignations des instruments de musiques, dans les cultures actuelles et dans les témoignages passés, peut nous apprendre sur leur histoire [244]. Là encore c'est un élément intermédiaire dans notre problématique, puisqu'il est intimement lié au musical, mais que son étude relève plutôt des méthodes de la philologie et de la linguistique historique. Je m'intéresse maintenant plus particulièrement à ce que l'analyse des **paramètres musicaux internes**⁶⁸ peuvent nous apprendre sur l'histoire, question évidemment liée, mais qui requiert la mobilisation des outils spécifiques de l'ethnomusicologie, en particulier ceux de l'analyse de la **systematique musicale**.

4.3.2 L'histoire inscrite au cœur de la musique

L'histoire dans les premières comparaisons ethnomusicologiques. Comme on l'a vu, la comparaison de musiques de différentes cultures avait déjà, aux premières heures évolutionnistes de l'ethnomusicologie, une dimension historique puisqu'elle était basée sur l'idée que le lointain donnait accès à l'ancien.

« Ich möchte darum heute [...] zu zeigen versuchen, wie gerade die Probleme der vergleichenden Musikwissenschaft auf die allgemeinsten Fragen : nach dem Ursprung und der Entwicklung der Musik und dem Wesen des Musikalisch-Schönen hinführen. »⁶⁹ von Hornbostel (1975 [1905]) [201, p. 249]

L'identification des cercles de diffusion culturelle proposée par le courant diffusionniste était là encore mû par cette recherche de l'origine et de l'histoire des instruments et des pratiques musicales. Mais au-delà de ces recherches souvent peu rigoureuses et aux présupposés idéologiques aujourd'hui largement abandonnés, la recherche de l'histoire des populations et de leurs musiques à travers la comparaison musicale s'est perpétuée en ethnomusicologie, bien que de manière plus diffuse et sans former un courant fort de la discipline.

Il est par ailleurs intéressant de noter que, bien que le concept d'*aire culturelle* ait été développé en anthropologie avec une intention historique, l'étude ethnomusicologique des aires musicales

67. ↑ Citée par Charles-Dominique [87, p. 28].

68. ↑ **Métrique, rythme et système scalaire.**

69. ↑ Voici la traduction en anglais proposée par Richard Campbell : « I would like to try today [...] to show just how the problems of comparative musicology lead directly to the most general questions, namely, the origin and growth of music and the nature of the musically beautiful. »⁷⁰ [201, p. 249].

70. Voici ma traduction faite à partir de celle, en anglais, de Richard Campbell : « Je voudrais essayer aujourd'hui [...] de montrer comment les problèmes de la musicologie comparée conduisent directement aux questions les plus générales, à savoir l'origine et le développement de la musique et la nature de la beauté musicale. ».

s'intéressera plutôt peu à la question historique, se concentrant sur des aspects descriptifs et classificatoires [314, pp. 76–77].

Histoire de pièces. L'une des finalités de l'identification de *familles de pièces* (via la comparaison de différentes pièces entre elles⁷¹) est aussi la recherche historique :

« Bayard sought “to identify specific melodies in as many of their variant forms as possible” (Bayard 1950 :4⁷²), a study which he believed would be “a worthy partner of tune-indexing efforts; a logical precursor and accompanier of melodic-variation studies, analyses of musical style, inquiries into the migrations and history of songs, and studies in the relations of melodies and words” (ibid.). »⁷³

Cowdery (1984) [98, p. 495]

Dans la continuité du travail présenté ci-dessus⁷⁴ Sylvie Le Bomin et moi-même avons par ailleurs initié un travail de collecte dans la province de l'Estuaire (Gabon) afin d'analyser les ressemblances entre différentes pièces jugées équivalentes par différentes populations et de tenter de les relier à l'histoire de celles-ci. Pour 4 pièces différentes, nous avons pu enregistrer les versions jugées équivalentes sur des critères contextuels (nom endogène ou identification par les tenant-e-s de la culture) telles que réalisées par 5 populations différentes. Pour 15 autres pièces, nous avons pu enregistrer les versions de 3 ou 4 populations différentes. L'analyse paradigmatique de ces enregistrements nous permettra d'une part de confronter le discours endogène au matériau sonore et d'autre part d'essayer d'identifier des degrés de ressemblance musicale entre ces différentes réalisations. Il s'agira ensuite de relier ces données à nos connaissances sur l'histoire de ces populations afin d'évaluer dans quelle mesure le matériau musical peut nous renseigner sur l'histoire des populations et de leurs musiques⁷⁵.

Traces historiques entre cultures proches. La comparaison de musiques est aussi fréquemment utilisée pour étudier les relations historiques entre des populations que l'on sait⁷⁶ apparentées. En étudiant en détail différents paramètres musicaux, et souvent d'autres types d'informations (sources écrites, ethnolinguistique, ...), il s'agit de déceler des éléments qui ne peuvent être partagés par hasard entre les différentes populations et ne peuvent être expliqués que par une cause historique (ascendance commune ou contact entre populations).

Ainsi Jean-Jacques Nattiez [304] compare, au cours d'une analyse détaillée, les jeux de gorges pratiqués dans trois cultures boréales : les Inuits (Amérique boréale), les Aïnoues (Sakhaline) et

71. ↑ Voir partie 4.2.4.2 page 137.

72. ↑ L'article Bayard (1950) cité par Cowdery correspond à la référence numéro [A37] de ma bibliographie.

73. ↑ « Bayard a cherché “à identifier des mélodies spécifiques sous toutes leurs formes possibles”, une étude qu'il pensait pouvoir être “une aide louable dans les efforts d'indexation des pièces; un précurseur logique et un compagnon des études sur la variation mélodique, des analyses du style musical, des enquêtes concernant la migration et l'histoire des chants, et des études sur les relations entre mélodies et textes.” . »

74. ↑ Voir la description du travail de Sylvie Le Bomin et Bikoma [238] dans la partie 4.2.4.2 page 137.

75. ↑ Nous n'avons malheureusement pas encore eu le temps d'effectuer ces analyses et cette étude ne sera donc pas plus détaillée dans cette thèse.

76. ↑ Grâce à des études provenant d'autres disciplines, comme la linguistique ou la génétique.

les Tchouktches (Extrême-Orient boréal). Constatant la similarité de ces pratiques, il envisage trois hypothèses pour expliquer cette similarité : une hypothèse universaliste, une hypothèse diffusionniste et une hypothèse phylogénétique. Après une discussion s'appuyant sur son analyse des trois pratiques, il conclut à l'existence d'un lien historique entre ces trois populations.

Un autre exemple est celui du travail de Susanne Fürniss [151] dans lequel elle étudie la circulation du rituel *Bèkà* dans le sud-est du Cameroun. Elle s'intéresse en détail aux pratiques musicales liées à ce rituel dans trois populations (les Baka, les Bangando et les Kwelé) en analysant en particulier leur polyrythmie. Elle constate que la formule polyrythmique des Kwelé Esel est très proche de celle des Bangando, mais très différente de celle des Kweké Djako qui s'apparente plutôt à celle des Baka. Cette observation entre en contradiction avec les discours des différentes populations concernant l'origine de ce répertoire, lesquels ne sont par ailleurs pas tous en accord. Susanne Fürniss se sert de cette analyse musicale, combinée avec d'autres indices anthropologiques et ethnolinguistiques, pour discuter de l'histoire de ces populations. Même si elle ne démêle pas complètement l'écheveau complexe que constituent les migrations des populations du sud-est du Cameroun, elle montre comment les données musicales peuvent être utilement employées comme sources de connaissances historiques.

On pourrait sans doute citer de nombreux autres travaux allant dans ce sens⁷⁷, mais il est difficile⁷⁹ de les recenser, d'autant plus que l'aspect historique est souvent une facette de l'analyse parmi d'autres, sans en constituer la question centrale. Ce type d'études nous apprend combien l'analyse musicale, tout en étant une source potentielle d'informations historiques, ne dévoile pas facilement le passé et nécessite d'être maniée avec prudence compte tenu de l'histoire souvent complexe des populations et de la nature indirecte des preuves apportées.

Méthodes issues de l'inférence phylogénétique. Comme je l'ai déjà exposé⁸⁰, les méthodes issues de l'inférence phylogénétique ont aussi été quelquefois utilisées pour étudier des données musicales. Si ces méthodes n'ont pas toujours été utilisées à des fins d'inférence historique, c'est parfois le cas (éventuellement de manière implicite). Dans cette perspective, le travail le plus consistant, tant du point de vue de la taille de l'échantillon que de la réflexion méthodologique, est sans doute celui de Le Bomin et coll. [240]⁸¹.

Musique et migrations. Victor Grauer [165] propose quant à lui un projet d'inférence historique d'une envergure bien plus importante. Il discute des relations historiques entre les musiques du monde entier qu'il prétend pouvoir retracer jusqu'à la première migration de l'homme moderne hors du continent africain⁸² il y a environ 85 000 ans. Pour cela, Grauer se sert des descriptions

77. ↑ On peut par exemple évoquer l'article de Voisin et Cloarec-Heiss [A432]⁷⁸ ou encore ceux du dossier du *Journal des africanistes* (2014) « Ethnomusicologie et histoire » coordonné par Susanne Fürniss.

78. Citée par Fürniss [151, p. 10].

79. ↑ Et hors de la portée de cette thèse.

80. ↑ Voir partie 2.1.5 page 70.

81. ↑ Voir partie 2.1.5.3 page 73.

82. ↑ Cette théorie de l'origine unique de l'homme moderne issu d'une population venant d'Afrique est parfois surnommée *Out-of-Africa*.

musicales produites par le projet *Cantometric*⁸³ auquel il a participé.

L’observation d’une certaine similarité entre les musiques pygmées des forêts tropicales africaines et les musiques bochimanes du désert du Kalahari⁸⁴ le conduit à suggérer une origine commune de ces populations permettant d’expliquer cette ressemblance, définissant ainsi un style commun qu’il nomme *Pygmée/Bochiman* (P/B). Il constate ensuite la présence dans des parties très éloignées du monde de polyphonies en hoquets (île Salomon, Afrique, Europe, Amérique du Sud, Flores, Indonésie), de flûtes de pan ou d’ensembles de flûtes jouées en hoquets (Russie, îles Salomons, Afrique, Nouvelle-Guinée) et d’ensembles d’idiophones joués en hoquets (tubes frappés et hochets d’Indonésie, gongs des Philippines et de Flores, gamlan de Java et de Bali), autant de pratiques musicales qu’il rattache au style P/B. Il explique finalement ces observations, en s’inspirant des travaux des généticiens sur les populations humaines, par la transmission du style P/B depuis son origine en Afrique en suivant les migrations humaines. La diversité des pratiques, et notamment l’absence dans certaines zones des caractéristiques du style P/B (Australie, Amérique du Nord, Asie centrale) est interprétée comme le résultat de goulets d’étranglements provoqués par des événements écologiques (éruption du mont Toba, ère glaciaire).

On voit comment cette approche se sert de la comparaison entre différentes musiques (issues de *Cantometric*) pour proposer des hypothèses sur l’histoire des musiques et des populations qui les pratiquent. Cependant, de par son envergure, ce projet ne peut se baser sur une analyse détaillée et fine des pratiques musicales. Comme le relève Stock [397, p. 78]⁸⁵, la comparaison détaillée des musiques des Pygmées Aka et des Bochimans Ju’hoansi faite par Fürniss et Olivier [153] tend plutôt à contredire les propositions de Grauer [165]. Cette approche fait ainsi l’impasse sur l’énorme diversité tant intraculturelle que régionale qui existe au sein des musiques étudiées. Les propositions qui en sont issues ne peuvent qu’être spéculatives, comme le reconnaît volontiers l’auteur [163, p. 104], et on peut même douter de la portée heuristique d’un tel travail. L’article de Grauer a d’ailleurs été largement critiqué, en particulier par Stock [397].

4.4 Conclusion du chapitre

J’ai exposé succinctement dans ce chapitre la place de la comparaison dans l’histoire de l’ethnomusicologie, montrant en particulier qu’elle est restée importante dans la pratique de l’analyse quand bien même elle était minimisée dans les discours. J’ai aussi présenté la diversité des formes que pouvait prendre la comparaison en ethnomusicologie ainsi que la variété des buts qu’elle pouvait servir. Après avoir rappelé que l’étude de l’histoire fait partie des champs de l’analyse ethnomusicologique, j’ai présenté comment la comparaison de musiques a été utilisée en ethnomusicologie pour renseigner sur l’histoire.

83. ↑ Ce projet est décrit dans la partie 4.2.3.2 page 134.

84. ↑ Voir à ce sujet l’analyse de Fürniss et Olivier [153], prise comme exemple dans la partie 3.2 pages 114 à 118, qui propose une analyse détaillée des musiques des Pygmées Aka et des Bochimans Ju’hoansi.

85. ↑ Et comme je l’ai expliqué dans la partie 3.2.1.1 page 116.

Conclusion de la première partie

Cette première partie a été l'occasion de donner un aperçu des travaux réalisés dans plusieurs disciplines différentes. La présentation des méthodes d'**inférence historique** par comparaison des structures biologiques et culturelles a été prolongée par une revue de l'application de méthodes issues de la **systematique biologique** pour étudier des données culturelles. La description des outils de l'ethnomusicologie structuraliste a par ailleurs été complétée par une synthèse sur la manière dont l'ethnomusicologie a pu aborder les relations entre comparaison et histoire.

Chacune de ces revues, bien que succincte, m'a paru nécessaire pour bien comprendre les enjeux et les problématiques de cette thèse. On voit à l'issue de cette introduction que la mise en lien de ces disciplines permet de poser les bases d'une problématique cohérente et unifiée. Au vu des avancées obtenues dans l'étude de l'histoire culturelle grâce aux méthodes de comparaison structurelles et compte tenu de la base méthodologique offerte par l'ethnomusicologie permettant la comparaison interculturelle, la comparaison de musiques paraît offrir (même si l'histoire de la discipline laisse apercevoir les difficultés qu'il peut y avoir) une source prometteuse d'informations sur l'histoire des musiques et de leurs populations.

Faire converger les apports de ces différentes disciplines pour en faire un sujet de recherche. Faire coïncider des cadres épistémologiques et des traditions disciplinaires disjointes. Compte tenu du sujet fondamentalement interdisciplinaire qui m'intéresse, cela constitue en soi, me semble-t-il, une petite avancée sur le plan méthodologique et heuristique. C'est parce que cet apport me paraît intéressant que j'ai choisi de le développer si longuement dans cette première partie de ma thèse.

Mais ce cadre de réflexion et les problématiques de recherches soulevées ne sont intéressantes que si elles sont mises à l'épreuve de l'analyse, soumises à la difficile confrontation avec le réel. La deuxième partie de cette thèse s'attachera donc à présenter les résultats des analyses que j'ai menées. Je montrerai que mes premières tentatives, réalisées alors que les réflexions présentées dans cette introduction n'étaient sans doute pas suffisamment abouties, manquaient à la fois d'une bonne connaissance des données que j'avais à ma disposition et d'un cadre théorique et méthodologique plus clair. Ces constatations m'ont amené à décrire plus en détail le jeu de données collecté par Sylvie Le Bomin dans une vingtaine de localités du Gabon et en particulier les **paramètres musicaux internes**. Elles m'auront aussi poussé à m'interroger plus en profondeur sur ce que peuvent nous apprendre l'analyse comparative des musiques du Gabon. J'ai ainsi

cherché à comprendre plus précisément quelle était cette « histoire » que je cherchais à inférer : Celle des populations ? Celle de la culture ? Celle de la musique ?

On verra une nouvelle fois comment, réflexion et analyse, théorie et terrain, réussissent bon an mal an à travers leur permanente opposition à nous emmener vers des champs d'exploration nouveaux et stimulants.



Muséum National d'Histoire Naturelle

École Doctorale Sciences de la Nature et de l'Homme – ED 227

Année 2017

N° attribué par la bibliothèque

THÈSE

Pour obtenir le grade de

DOCTEUR DU MUSÉUM NATIONAL D'HISTOIRE NATURELLE

Spécialité : Évolution culturelle

Présentée et soutenue publiquement par

Jérémy Gardent

le 30 juin 2017

Mesurer les musiques pour parler du passé La comparaison des musiques du Gabon comme source d'informations historiques

Volume II : Résultats — Discussion

Sous la direction de : Évelyne HEYER et Sylvie LE BOMIN

JURY :

Dana Rappoport,	Directrice de Recherche, CNRS, Paris	Rapporteuse
Stephan Shennan,	<i>Professor</i> , University College, Londres	Rapporteur
Olivier Tourny,	Chargé de Recherche, CNRS, Paris	Examineur
Olivier Morin,	<i>Research group leader</i> , Max Planck Institute, Jena	Examineur
Évelyne Heyer,	Professeure, MNHN, Paris	Directrice de Thèse
Sylvie Le Bomin,	Maître de Conférence, MNHN, Paris	Co-directrice de Thèse



Travail sous licence Creative Commons
Attribution – Partage dans les Mêmes Conditions
4.0 International

Table des matières

Volume I : Introduction	7
Remerciements	7
Préambule	9
Introduction	13
1 L'inférence historique par comparaison des structures formelles	15
1.1 De quoi s'agit-il?	17
1.2 Historique disciplinaire	18
1.3 Description de méthodes d'inférence historique	24
1.4 Conclusion du chapitre	47
2 Les méthodes issues de l'inférence phylogénétique appliquées aux données culturelles	49
2.1 État de l'art disciplinaire	53
2.2 Considérations générales sur ces transpositions interdisciplinaires	75
2.3 Tableau comparatif de quelques études	81
2.4 Conclusion du chapitre	103
3 Systématique musicale	105
3.1 Outils théoriques et méthodologiques	107
3.2 Des outils au service de la comparaison interculturelle	114
3.3 Réflexions autour de la notion de <i>modèle</i>	118
3.4 Conclusion du chapitre	128
4 Musique, comparatisme et histoire	129
4.1 Quelle place pour la comparaison?	130
4.2 La comparaison dans la pratique de l'ethnomusicologie	132
4.3 La comparaison musicale pour l'inférence historique	139
4.4 Conclusion du chapitre	143
Conclusion de la première partie	145
Volume II : Résultats — Discussion	153
Résultats	153
5 Essai d'inférence phylogéographique	155
5.1 Contexte et objectifs	157
5.2 Méthodes	158
5.3 Résultats	165
5.4 Discussion	180
5.5 Perspectives	184

6	Un aperçu de la diversité des musiques du Gabon	187
6.1	Introduction	188
6.2	<i>Diversity and distribution of musics in Gabon</i>	191
6.3	Conclusion et perspectives	245
7	Quelles données, quelle(s) histoire(s) ?	247
7.1	Positionnement épistémologique	249
7.2	Croisements d'histoires	251
7.3	Les indicateurs de choix des marqueurs historiques	268
7.4	Conclusion du chapitre	274
	Conclusion de la deuxième partie	275
	Discussion	277
8	Pistes de recherches	279
8.1	Comprendre la transmission et la transformation des musiques	280
8.2	Approfondir l'étude des formules rythmiques	287
9	Réflexions sur la pratique de la recherche	295
9.1	La relation aux données	297
9.2	Le rôle d'analyste	301
9.3	Limites de l'interdisciplinarité	304
9.4	La transposition de méthodes	306
	Conclusion générale	315
	Volume III : Glossaire — Annexes	327
	Glossaire argumenté	327
	Annexes	370
A	Conventions	373
A.1	Citations et traductions	373
A.2	Sources et références bibliographiques	374
B	Suppléments	377
B.1	Chapitre 3 : Citations autour de la notion de <i>modèle</i>	378
B.2	Chapitre 5 : Inférence phylogéographique	387
B.3	Test de Mantel	421
C	Bibliographie	425
C.1	Publications consultées	425
C.2	Autres publications citées	445
D	Listes des figures et des tableaux	457
D.1	Liste des figures	458
D.2	Liste des tableaux	460
E	Table exhaustive des matières	461

Volume II
Résultats — Discussion

Deuxième partie

Résultats

Chapitre 5

Essai d'inférence phylogéographique

5.1	Contexte et objectifs	157
5.2	Méthodes	158
5.2.1	L'arbre et les nœuds étudiés	158
5.2.2	Le modèle d'évolution	159
5.2.3	La méthode de MCMC	162
5.2.3.1	Chaîne d'états	162
5.2.3.2	Marche aléatoire guidée	162
5.2.3.3	Allumage de la chaîne	164
5.2.3.4	Échantillonnage de la chaîne	164
5.3	Résultats	165
5.3.1	Vérifications	165
5.3.1.1	Convergence	165
5.3.1.2	Autocorrélation	165
5.3.2	Localisations inférées des nœuds	168
5.3.3	Cas d'incertitudes	168
5.3.3.1	Distribution dans les itérations échantillonnées	168
5.3.3.2	Association entre longitude et latitude d'un même nœud	171
5.3.3.3	Caractérisation des positions alternatives de chaque nœud	171
5.3.3.4	Associations entre les positions alternatives de différents nœuds	176
5.4	Discussion	180
5.4.1	Méthodologie	180
5.4.1.1	Autocorrélation et échantillonnage	180
5.4.1.2	Surparamétrisation	180
5.4.1.3	Arbre des relations entre patrimoines	180
5.4.2	Limite du modèle utilisé pour l'inférence	181
5.4.2.1	Le modèle arborescent d'évolution des patrimoines	181
5.4.2.2	Le patrimoine musical comme unité évolutive	181
5.4.2.3	Patrimoines musicaux et populations	182
5.4.3	Pertinence de l'analyse	182
5.4.3.1	Tester une méthode	183
5.4.3.2	Produire de la connaissance historique	183
5.5	Perspectives	184

5.5.1	Résumé du chapitre	184
5.5.2	Mieux connaître nos données	184
5.5.3	Affiner les fondements théoriques et méthodologiques	185

5.1 Contexte et objectifs

À l'issue de leur analyse cladistique des patrimoines musicaux de 58 populations du Gabon, Le Bomin et coll. [240] obtiennent un arbre reliant ces patrimoines, présenté dans la figure 5.1^{1,2}. En supposant que cet arbre puisse représenter des relations historiques entre les patrimoines musicaux et en connaissant les localités où ces patrimoines ont été collectés, il est possible de faire des hypothèses sur les régions dont ces patrimoines sont issus³. En effet, si deux patrimoines sont proches géographiquement et regroupés dans l'arbre, l'ancêtre de ces deux patrimoines était vraisemblablement situé dans la même région géographique. Inversement, si deux patrimoines sont rapprochés par l'arbre, mais sont très éloignés géographiquement, on peut supposer qu'il y a eu chez une ou plusieurs des populations pratiquant ces musiques des déplacements de grande ampleur.

Je présente ici une analyse cherchant à inférer les positions de certains nœuds de l'arbre de la figure 5.1, considérés comme des ancêtres hypothétiques des patrimoines musicaux actuels. Cette analyse est fondée sur une méthode d'inférence bayésienne dite de *phyloplastie*. Cette méthode a été proposée et utilisée pour la première fois en biologie par Venditti et coll. [430] pour étudier l'évolution des mammifères et a récemment été appliquée par Currie et coll. [103] pour étudier l'évolution des langues subsahariennes. Après avoir expliqué le principe de cette méthode⁴ je présenterai son application au cas des patrimoines musicaux du Gabon⁵ avant de discuter les limites de cette approche⁶ et de présenter les perspectives qui en découlent, orientant la suite de mon travail⁷.

Cependant, comme expliqué dans la partie 5.4⁸, l'analyse présentée ici a posé de nombreuses questions méthodologiques et m'a conduit à une importante réorientation de mes analyses. J'ai néanmoins choisi de présenter cette partie de mon travail afin de partager ces essais et ce que j'ai pu en tirer. L'analyse n'étant pas complètement aboutie, j'ai cherché à reproduire le cheminement théorique et méthodologique qui m'a amené à expérimenter cette méthode d'abord puis à l'abandonner ensuite. On ne s'étonnera donc pas que certaines hypothèses de départ soient manquantes ou peu discutées dans la première partie, car ce sont justement ces insuffisances qui se sont révélées problématiques par la suite. De même il peut être surprenant que je ne remette pas les résultats obtenus en perspectives avec nos connaissances de l'histoire des populations du Gabon et de leurs musiques. J'ai pourtant choisi de ne discuter de cette partie de mon travail que dans ses aspects méthodologiques. En effet, ne pouvant avoir une confiance suffisante dans ses fondements théoriques et méthodologiques, discuter des apports historiques de mon analyse

1. ↑ Page 160.
 2. ↑ Les buts et résultats de l'article de Le Bomin et coll. [240] dont il est question sont présentés dans la partie 2.1.5.3 page 73.
 3. ↑ Le lien n'est cependant pas si évident, comme j'en discuterai dans la partie 5.4.2.3 page 182.
 4. ↑ Partie 5.2 pages 158 à 164.
 5. ↑ Partie 5.3 pages 165 à 177.
 6. ↑ Partie 5.4 pages 180 à 183.
 7. ↑ Partie 5.5 pages 184 à 185.
 8. ↑ Page 180.

aurait risqué de me conduire au mieux dans une impasse, au pire à des discussions stériles.

5.2 Méthodes

La méthode utilisée ici pour estimer les coordonnées géographiques correspondantes à différents nœuds de l'arbre, appelée phyloplastie⁹ [103, p. 3], est une méthode d'inférence bayésienne¹⁰ dérivée de la méthode des contrastes indépendants¹¹ et implémentée dans le logiciel *BayesTraits*¹² [333]. C'est une méthode probabiliste qui consiste à chercher les positions de ces nœuds qui permettent d'expliquer le mieux possible les positions actuelles des populations en supposant que leurs relations de parentés sont celles décrites par l'arbre de la figure 5.1. Pour cela la longitude et la latitude de chaque nœud sont considérées comme des caractères continus qui sont modifiés aléatoirement au cours du temps. Ceci constitue le modèle d'évolution de ces caractères¹³. On peut alors calculer, pour chaque position supposée des nœuds, la vraisemblance des données et de l'arbre selon ce modèle : quelle probabilité a-t-on d'obtenir les données observées à partir de ce modèle et en supposant que les nœuds étaient situées à ces positions ? La méthode de MCMC permet alors de rechercher les positions des nœuds qui permettent d'obtenir la plus grande vraisemblance des données¹⁴.

5.2.1 L'arbre et les nœuds étudiés

Je suppose ici que les patrimoines musicaux analysés par Le Bomin et coll. [240] évoluent selon un modèle strictement arborescent¹⁵ suivant l'arbre proposé dans leur analyse et reproduit dans la figure 5.1¹⁶. Cet arbre contient une cinquantaine de nœuds internes et il est impossible d'inférer dans un temps raisonnable, avec la méthode utilisée ici, les positions géographiques de la totalité de ces nœuds. J'ai donc choisi de n'étudier que les groupes dont les membres montrent une cohérence soit géographique, soit ethnonymique, soit du point de vue du système de filiation, ainsi que la racine de l'arbre. Ces nœuds sont mis en évidence sur la figure 5.1 par des ronds colorés. Il s'agit des 20 groupes suivants :

- 2 groupes montrant une cohérence du système de filiation :

Pa (nœud 90)¹⁷ : patrilinéaire

9. ↑ *Phyloplasty*, en anglais.

10. ↑ Le principe général des méthodes bayésiennes est rapidement présenté dans la partie 1.3.7.2 page 39.

11. ↑ La méthode est décrite en détail dans l'article de Venditti et coll. [430, supplément, pp. 12–16] : [http : //www.nature.com/nature/journal/v479/n7373/extref/nature10516-s1.pdf](http://www.nature.com/nature/journal/v479/n7373/extref/nature10516-s1.pdf).

12. ↑ Les détails des paramètres utilisés pour l'analyse sont donnés dans l'annexe B.2.1 page 387.

13. ↑ Voir partie 5.2.2 page 159.

14. ↑ Voir partie 5.2.3 page 162.

15. ↑ Ce modèle suppose donc que les patrimoines musicaux se séparent au cours du temps en plusieurs patrimoines (ramification) et qu'il n'y a pas d'échange entre les patrimoines une fois séparés. Je discute cette hypothèse dans les parties 5.4.1.3 page 180 et 5.4.2.1 page 181.

16. ↑ Le modèle d'évolution utilisé est décrit ci-dessous nécessite de spécifier les longueurs relatives des différentes branches de l'arbre. J'ai utilisé les longueurs de branches issues de l'analyse cladistique de Le Bomin et coll. [240], qui correspondent aux nombres de changements d'états de caractères intervenant sur cette branche.

17. ↑ Les numéros de nœuds se réfèrent à ceux indiqués sur la figure 5.1 page 160.

Ma (nœud 107) : matrilinéaire

- 8 groupes montrant une cohérence géographique :

Bo (nœud 72) : région de la ville de Boumango

La (nœud 78) : région de la ville de Lastourville

Oi (nœud 87) : province de l'Ogooué Ivindo

Wn (nœud 83) : province du Woleu Ntem

Ho (nœud 68) : province du Haut-Ogooué

OI (nœud 81) : province de l'Ogooué Lolo

N (nœud 88) : partie nord du Gabon

Ng (nœud 100) : province de la Ngounié

- 9 groupes montrant une cohérence *ethnonyme* ethnonymique :

Ngo (nœud 62) : Ngom

Tek (nœud 64) : Teke

Oba (nœud 66) : Obamba

Bko (nœud 80) : Bongo de la région de Koulamoutou

F (nœud 82) : Fang

Bt (nœud 91) : Bongo voisins des Tsogho

My (nœud 102) : Myènè

Pu (nœud 105) : Punu

Ts (nœud 93) : Tsogho

- La racine de l'arbre (nœud 108).

5.2.2 Le modèle d'évolution

Au cours de cette simulation, la position géographique (longitude et latitude) de chacun des nœuds est modifiée de manière aléatoire (on dit qu'elle *évolue*). Cependant la vitesse d'évolution (c'est-à-dire le nombre de modifications de ces paramètres dans un même intervalle de temps) peut varier au cours de cette évolution. Autrement dit, à certains points de l'arbre le déplacement géographique des populations peut être plus rapide (ou au contraire plus lent). Cette variation de la vitesse d'évolution peut affecter soit une seule branche de l'arbre, soit toute une partie de l'arbre c'est-à-dire toutes les branches issues d'un même nœud [430, supplément pp. 13–14].

Ces hypothèses concernant la manière dont évolue la position géographique au cours du temps sont implémentées par un modèle¹⁸ dans le logiciel *BayesTraits*. Concrètement il s'agit

18. ↑ Voir note 11 page 158.

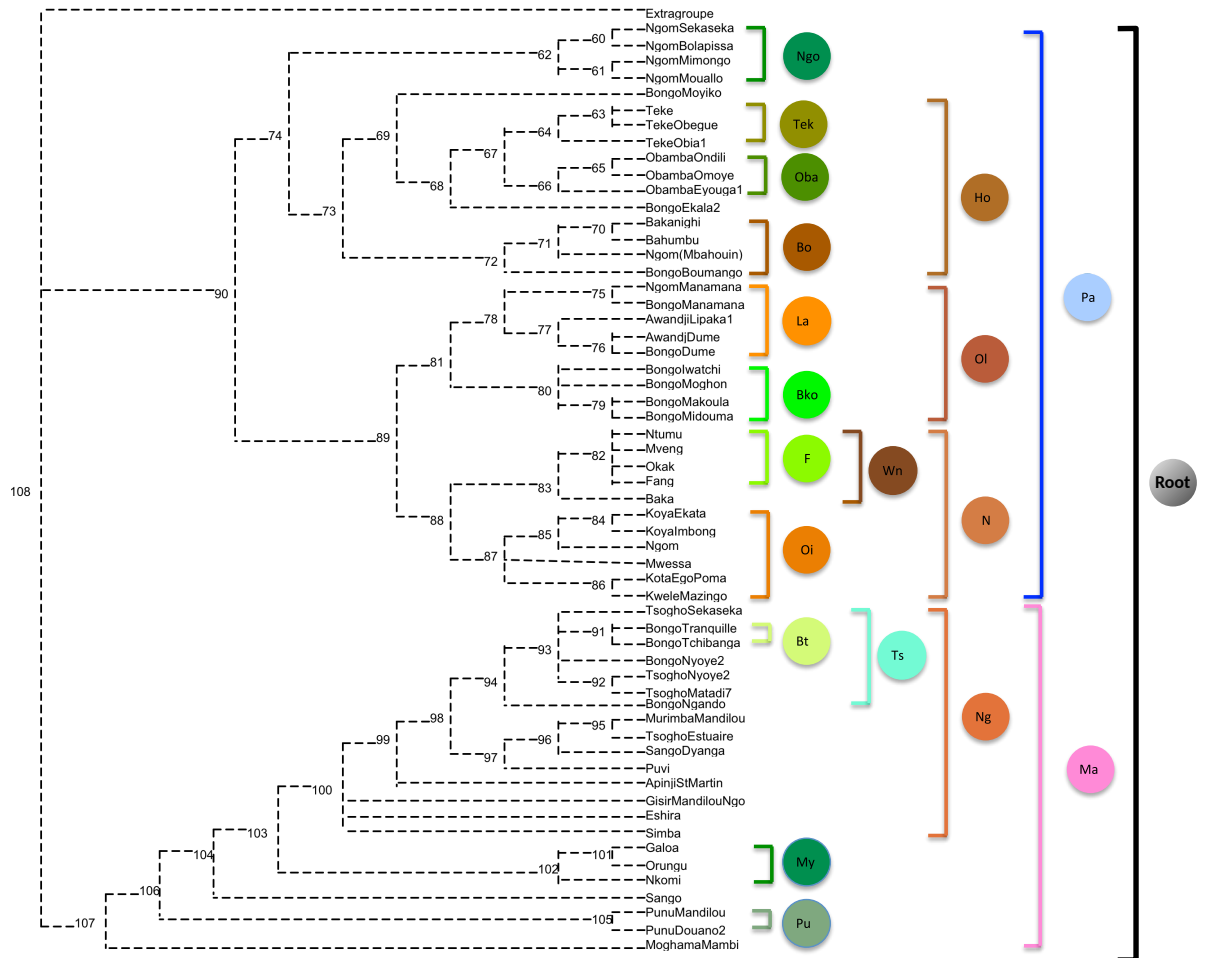


FIG. 5.1 : Arbre des patrimoines musicaux de 58 populations du Gabon. Cet arbre est issu de l'analyse cladistique de données musicales réalisée par Le Bomin et coll. [240]. Les numéros inscrits aux nœuds servent à désigner les groupes correspondant. Les 19 groupes indiqués par des ronds colorés sur la partie droite correspondent aux *clades* ayant une pertinence soit ethnonymique, soit géographique, soit du point de vue du *système de filiation*. Le rond noir correspond à la racine (*Root*, ensemble de tous les patrimoines musicaux étudiés). Ce sont les positions géographiques de ces 20 nœuds qui ont été inférées par l'analyse bayésienne.

d'une formule mathématique permettant de calculer, pour chaque position possible des nœuds internes étudiés¹⁹, une valeur de vraisemblance qui prenne en compte l'emplacement actuel des populations dont le patrimoine musical a été étudié²⁰ ainsi que les relations entre les différents patrimoines musicaux, décrites par l'arbre de la figure 5.1. Autrement dit cette formule répond à cette question :

- En considérant que les relations de filiation entre ces patrimoines sont celles représentées sur l'arbre 5.1
- et que les positions géographiques évoluent au cours du temps selon le modèle décrit ci-dessus
- en supposant que les 20 nœuds internes étudiés¹⁹ sont situés aux positions $(x_1, y_1), (x_2, y_2), \dots, (x_{20}, y_{20})$
- => quelle est la probabilité que les patrimoines musicaux étudiés se retrouvent à la position réellement observée aujourd'hui²⁰ ?

Sans rentrer dans le détail de cette formule mathématique²¹, ce modèle permet donc de calculer une valeur de vraisemblance Lh ²² en fonction de chaque ensemble de coordonnées géographiques des 20 nœuds internes dont on cherche à trouver la position :

$$Lh = f((x_1, y_1), (x_2, y_2), \dots, (x_{20}, y_{20})) \quad (5.1)$$

Dans cette équation, Lh est la valeur de vraisemblance calculée et chaque couple (x_i, y_i) représente la position géographique du nœud interne i , où x_i est la longitude et y_i la latitude. Pour simplifier l'écriture, on peut aussi désigner par X l'ensemble des 20 couples de coordonnées géographiques des 20 nœuds étudiés :

$$X = ((x_1, y_1), (x_2, y_2), \dots, (x_{20}, y_{20})) \quad (5.2)$$

$$Lh = f(X) \quad (5.3)$$

Selon la méthode d'inférence utilisée ici, rechercher les positions ancestrales des 20 nœuds internes étudiés revient à rechercher l'ensemble de coordonnées X pour lequel la valeur de la vraisemblance Lh est maximale. Il faut donc trouver l'ensemble de 40 paramètres (20 paramètres de longitude et 20 paramètres de latitude) qui maximise la valeur de Lh . Comme chacun de ces paramètres peut varier à l'infini et que le nombre de possibilités augmente pour chaque

19. ↑ C'est-à-dire les nœuds décrits dans la partie 5.2.1 page 158 et représentés par des ronds colorés sur la figure 5.1.

20. ↑ Les positions géographiques où ont été collectés les patrimoines musicaux sont donnés dans l'annexe B.2.1.1 page 388.

21. ↑ On peut la trouver dans le supplément de l'article de Venditti et coll. [430], voir note 11 page 158.

22. ↑ Le sigle Lh utilisé pour symboliser la vraisemblance vient de la traduction en anglais de ce terme : *likelihood*.

combinaison de paramètres, il est complètement impossible de tester la totalité des valeurs possibles²³. L'ensemble de coordonnées maximisant Lh est donc estimé de manière heuristique par la méthode **MCMC** décrite ci-dessous.

5.2.3 La méthode de MCMC

Les **méthodes de Monte-Carlo par chaînes de Markov** (MCMC²⁴) permettent de rechercher les valeurs d'un ensemble de paramètres maximisant une certaine fonction. Dans le cas présenté ici, elles peuvent donc être utilisées pour calculer l'ensemble X des positions géographiques des 20 nœuds internes étudiés maximisant la vraisemblance Lh .

5.2.3.1 Chaîne d'états

Cette méthode est basée sur une succession d'*états* formant une *chaîne d'états*. Un ensemble de valeurs X_0 est choisi au départ pour tous les paramètres : cet ensemble de valeurs constitue l'état E_0 , dont on peut calculer la valeur de la vraisemblance Lh_0 . Puis les valeurs X_0 des paramètres vont être modifiées aléatoirement pour prendre les valeurs X_1 , formant ainsi l'état E_1 et la valeur de la vraisemblance Lh_1 est calculée. On continue ainsi de modifier les valeurs des paramètres étudiés, chaque calcul d'un nouvel état et de la vraisemblance associée constituant une *itération*. Cet ensemble d'états E_0, E_1, \dots, E_n constitue la *chaîne d'états*. La figure 5.2 illustre la construction d'une chaîne d'états pour 3 nœuds. Le déroulement d'une telle chaîne peut être comparée à une marche sans but où la direction de chaque pas est imprévisible et choisie au dernier moment. On parle donc de *marche aléatoire*.

Plus le nombre de paramètres à estimer est important, plus l'espace à explorer par la chaîne d'états est vaste : il faut donc effectuer un nombre de pas d'autant plus grand pour avoir de bonnes chances d'avoir exploré l'ensemble des paramètres possibles et d'avoir trouvé la vraisemblance maximale. Ici j'ai choisi d'effectuer 1,1 milliards d'itérations au cours de l'analyse²⁵.

5.2.3.2 Marche aléatoire guidée

La construction d'une chaîne d'états complètement aléatoire ne permet évidemment pas de trouver la vraisemblance maximale. Les méthodes de **MCMC** introduisent des règles permettant d'orienter la direction de la chaîne afin d'identifier les zones où la fonction est maximisée : on parle de marche aléatoire *guidée*. À chaque état n , un état $n+1$ est proposé en modifiant aléatoirement la valeur des paramètres, comme expliqué ci-dessus. Cependant le nouvel état est parfois jugé insatisfaisant : on dit qu'il n'est pas *accepté* ou qu'il est *rejeté*. On choisit d'accepter le nouvel état

23. ↑ Supposons que l'on calcule toutes les valeurs de vraisemblance pour chaque centaine de kilomètres (ce qui est extrêmement grossier) de chaque paramètre, et uniquement dans le territoire du Gabon (entre 8,7 et 14,5 degrés de longitude et -3,9 et 2,3 degrés de latitude). Chaque paramètre pourrait donc prendre environ 5 valeurs, et le nombre de valeurs de la vraisemblance à calculer serait donc de $5^{40} \simeq 9 \times 10^{27}$, soit environ 9 milliards de milliards de milliards.

24. ↑ Le sigle MCMC provient du terme anglais *Markov chain Monte Carlo*.

25. ↑ Le choix de cette valeur est discuté dans la partie 5.4.1.1 page 180.

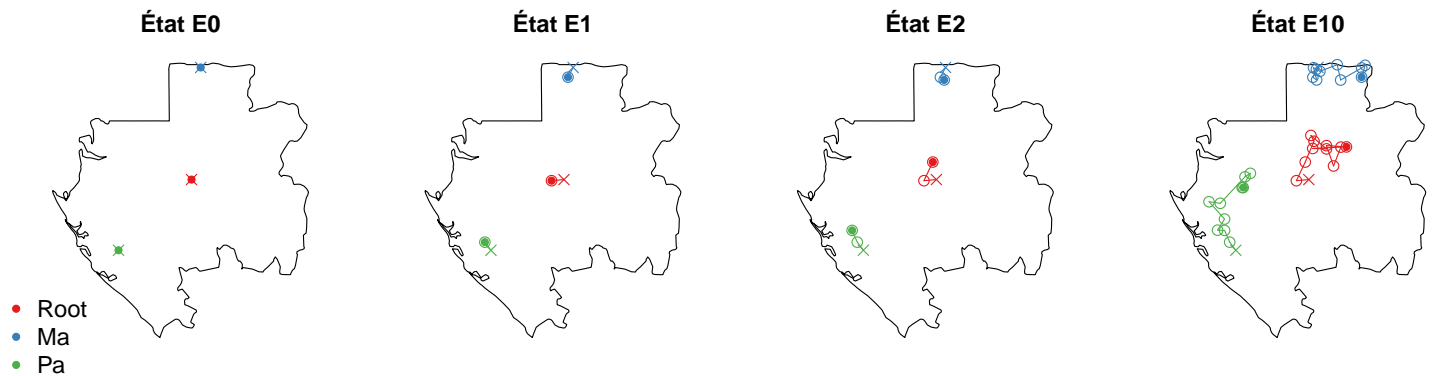


FIG. 5.2 : **Exemple de chaînes d'états.** Pour que la figure soit lisible, seuls trois nœuds (*Root*, *Ma* et *Pa*) sont montrés sur cet exemple, mais le principe est le même pour les 20 nœuds étudiés. À l'état initial E0, les 6 paramètres (3 de longitude et 3 de latitude) donnant les coordonnées des 3 nœuds sont pris au hasard. Cet état initial est représenté, pour chaque nœud, par une croix. À chacun des états E1, E2 et E10, la position de chaque nœud est représentée par un rond plein, les états antérieurs par des cercles vides et l'ordre de succession des états par des lignes reliant les états, dessinant ainsi une chaîne d'états pour chaque nœud.

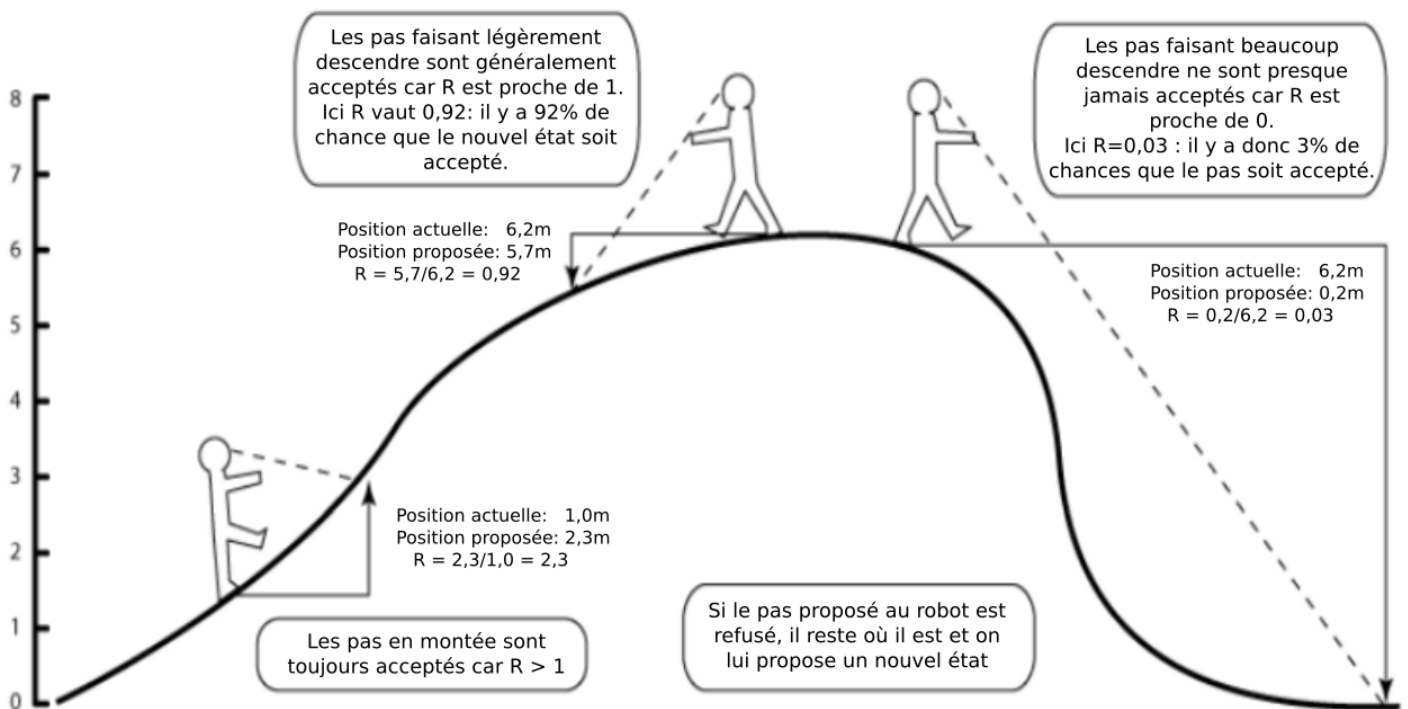


FIG. 5.3 : **Illustration du principe de la marche guidée.** Dans cet exemple on cherche à maximiser la fonction « altitude ». L'axe horizontal représente le paramètre dont on cherche la valeur qui maximise la fonction. Pour chaque valeur de ce paramètre, l'altitude correspondante est représentée par la courbe. Un robot parcourt ce paysage, chaque position correspondant à un état. À chaque itération, un nouvel état proche est proposé. Le choix de se déplacer jusqu'à cet état se fait en fonction des valeurs de la fonction maximisée (l'altitude) à l'état actuel et à l'état proposé. Dans cet exemple, le critère de choix est simple : la probabilité d'accepter le nouvel état est égale au ratio (noté R) entre la nouvelle altitude et l'ancienne. Schéma modifié (traduit) d'après celui de Mecanopsis (CC-by-SA) : https://en.wikipedia.org/wiki/File:Robot_metaphor.png

en fonction de la vraisemblance de l'état actuel et de celle du nouvel état proposé. Ce principe est illustré sur la figure 5.3. Le critère utilisé pour décider ou non d'accepter le nouvel état est celui de Metropolis-Hastings [430, supplément p. 14], dont on pourra trouver une explication détaillée dans l'article de Delsuc et Douzery [117, pp. 77, 79].

5.2.3.3 Allumage de la chaîne

Comme les valeurs initiales des paramètres sont choisies au hasard, la vraisemblance calculée au départ est forcément très faible puisqu'elle n'a pas encore été optimisée. Il faut un certain temps pour que la chaîne explore les paramètres possibles avant d'arriver aux valeurs maximales de vraisemblance. Le nombre d'itérations nécessaires pour atteindre les zones du paysage où la vraisemblance est la plus forte est appelé la *période d'allumage*²⁶. Encore une fois, plus le nombre de paramètres estimés est grand, plus la période d'allumage doit être longue. Les itérations correspondantes à cette période d'allumage sont écartées de l'analyse puisque la vraisemblance n'y est pas encore maximisée. J'ai choisi ici une période d'allumage de 100 millions d'itérations : cela signifie donc que les résultats des 100 premiers millions d'itérations ne sont pas sauvegardées par le logiciel *BayesTraits*.

5.2.3.4 Échantillonnage de la chaîne

Cette marche aléatoire guidée permet en quelque sorte d'explorer le paysage de la fonction de vraisemblance, en particulier les zones où la vraisemblance est maximale. Comme les valeurs successives de la chaîne d'états sont nécessairement relativement proches (puisque la chaîne se déplace par petites variations des paramètres) elles ne sont pas indépendantes. Afin d'avoir une meilleure idée du paysage, on ne choisit donc que des itérations de la chaîne qui sont supposées être suffisamment indépendantes : plus deux itérations sont éloignées sur la chaîne, plus elles sont indépendantes. On dit alors qu'on *échantillonne* les valeurs prises par la chaîne. Deux échantillons successifs doivent être d'autant plus éloignés sur la chaîne que le nombre de paramètres à estimer est grand. Compte tenu du grand nombre de paramètres inférés ici, j'ai choisi d'échantillonner la chaîne toutes les 100 000 itérations et donc de conserver au total les résultats de 10 000²⁷ itérations supposées indépendantes.

26. ↑ *Burn-in*, en anglais.

27. ↑ 1,1 milliard d'itérations sont effectuées, dont les 100 premiers millions ne sont pas enregistrés (allumage) et dont seules une sur 100 000 sont échantillonnées : $\frac{1,1 \times 10^9 - 100 \times 10^6}{100\,000} = 10\,000$

5.3 Résultats

5.3.1 Vérifications

5.3.1.1 Convergence

Comme on l'a vu, la chaîne met un certain temps avant de réunir un ensemble de paramètres permettant d'obtenir une vraisemblance maximale. Il est donc nécessaire de vérifier que la période d'allumage définie est suffisante, c'est-à-dire qu'un nombre suffisant d'itérations initiales ont été écartées. Si c'est le cas, on dit que la chaîne a *convergé*. Pour cela on vérifie que la vraisemblance ne continue plus d'augmenter au cours des itérations échantillonnées. La figure 5.4 montre qu'en effet la vraisemblance reste stable. De même la figure 5.6 montre que les valeurs des longitudes et latitudes prises au cours des différentes itérations restent stables pour les différents nœuds²⁸. La chaîne a donc bien convergé.

5.3.1.2 Autocorrélation

Comme on l'a vu, les différentes itérations échantillonnées doivent être suffisamment espacées sur la chaîne pour pouvoir être considérées comme indépendantes. En effet, une position sera considérée d'autant plus probable qu'elle a été visitée souvent par la chaîne d'état. Cependant, si les échantillons ne sont pas suffisamment espacés, la chaîne d'états sera restée dans une zone relativement proche entre deux itérations successives. Certaines positions de nœuds pourront donc avoir été visitées souvent non pas parce que la vraisemblance correspondante est élevée, mais parce différentes itérations ont été échantillonnées avant que la chaîne ait eu vraiment la possibilité de s'éloigner.

Pour vérifier cela, on peut mesurer l'autocorrélation de la vraisemblance de la chaîne. L'autocorrélation de la chaîne mesure si les valeurs prise par celle-ci sont corrélées à ces même valeurs, mais décalées d'un certain nombre d'itérations. La figure 5.5 montre qu'il existe encore une autocorrélation lorsque le décalage est inférieur à 40 itérations échantillonnées²⁹. Ceci montre que le paramètre d'échantillonnage choisi est insuffisant et aurait dû être au moins 40 fois plus important. Pour avoir le même nombre d'échantillons sans que les itérations échantillonnées ne soient corrélées, il aurait donc fallu effectuer 40 milliards d'itérations échantillonnées toutes les 4 millions d'itérations. Si on rééchantillonne pour ne garder qu'une itération sur 40 déjà échantillonnées, on obtient un échantillon de 250 itérations, ce qui manque de finesse pour explorer les résultats.

Cependant on voit sur la figure 5.5 que l'intensité de cette corrélation décroît rapidement et à partir d'un décalage de 7 itérations, la variance expliquée par l'itération précédente (égale au carré du coefficient de corrélation [368, p. 416]) est inférieure à 1%. On peut ainsi penser qu'il est raisonnable de ne rééchantillonner qu'une fois toutes les 7 itérations, ce qui donne un

28. ↑ Voir aussi les figures correspondantes pour chacun des nœuds dans l'annexe B.2.2 page 393.

29. ↑ Soit 4 millions d'itérations, puisque seule une itération sur 100 000 est échantillonnée.

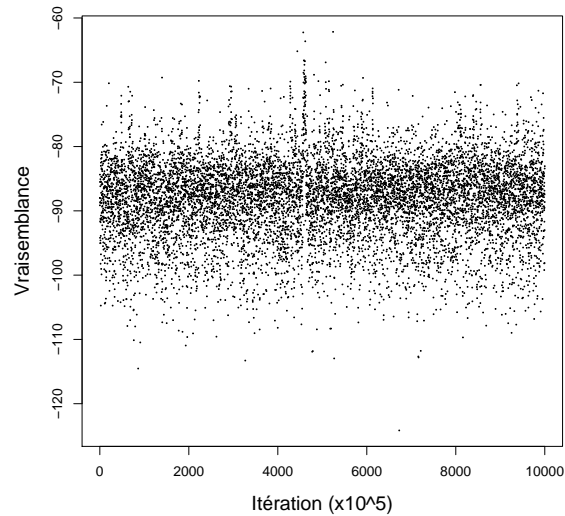


FIG. 5.4 : **Représentation de la vraisemblance de la chaîne pour chaque itération échantillonnée.** Chaque point représente la valeur de la **vraisemblance** de la **chaîne d'états** pour une des **itérations** échantillonnées. L'abscisse indique le numéro de l'itération^a et l'ordonnée la valeur de vraisemblance correspondante. On voit que la plupart des valeurs de vraisemblance sont distribuées entre -81 et -96, la valeur la plus fréquemment étant proche de -88. On voit aussi pour des itérations dont les index sont proches^b les valeurs de vraisemblance sont aussi relativement proches. Ceci révèle une autocorrélation des données, comme expliqué dans la partie 5.3.1.2 page 165 et illustré sur la figure 5.5.

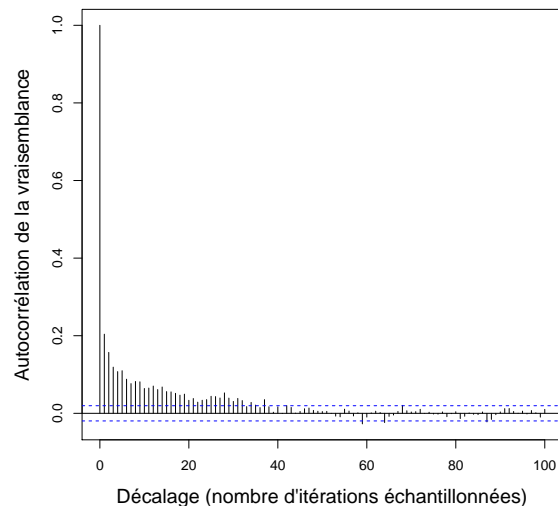


FIG. 5.5 : **Autocorrélation de la vraisemblance de chaque itération échantillonnée.** Pour chaque valeur de décalage (indiquée en abscisse), la barre correspondante indique le coefficient de corrélation de la vraisemblance totale de la chaîne avec ces mêmes valeurs mais décalées d'un certain nombre d'itérations. Les deux lignes en pointillés bleus indiquent les valeurs en dessous desquelles l'autocorrélation (positive ou négative) n'est pas considérée comme significative (intervalle de confiance de 5%). Par exemple on voit que pour un décalage de 1, l'autocorrélation vaut 0,2. Ceci signifie donc que la corrélation entre les valeurs de vraisemblance des itérations échantillonnées numéro 1 à 9999 d'une part et celles des itérations numéro 2 à 10 000 d'autre part est de 20%. Autrement dit, la valeur de la vraisemblance d'une itération échantillonnée peut être prédite à $0,2^2 = 0,04$ (4%) par la valeur de la précédente. Cette figure a été obtenue grâce à la fonction *acf* [A429] du logiciel R [A344].

a. ↑ Comme une itération sur 100 000 est échantillonnée, l'unité d'indice des itérations est multiplié par 100 000. Ainsi le chiffre lui indique directement l'indice de l'itération échantillonnée.

b. ↑ Par exemple entre les 4450^{ème} et 4650^{ème} itérations échantillonnées.

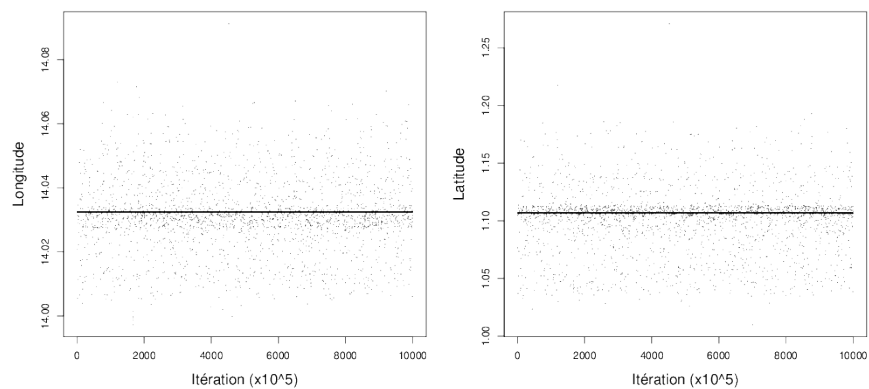
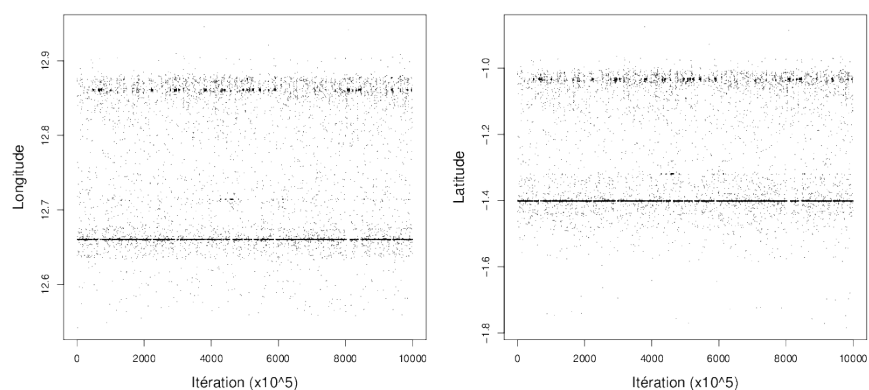
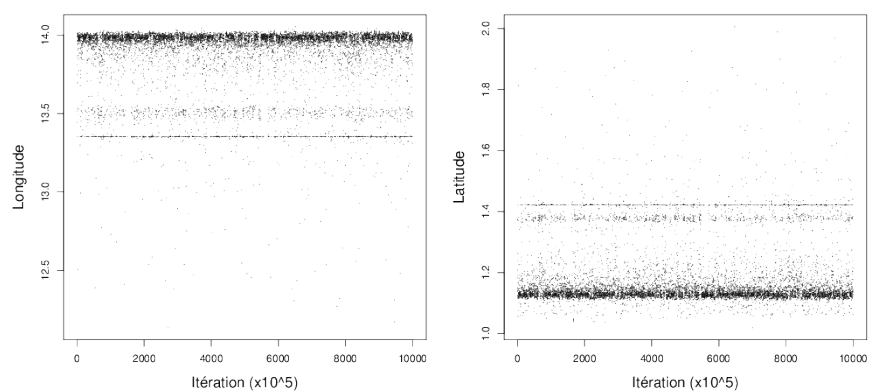
(a) O_i (b) O_l (c) N

FIG. 5.6 : Longitude et latitude de trois nœuds à chaque itération échantillonnée. Chaque graphique montre la valeur de la longitude (gauche) et de la latitude (droite) prises à chacune des itérations échantillonnées pour les nœuds : O_i (a), O_l (b) et N (c). Voir la légende de la figure 5.4 page 166 pour plus de détails. Les représentations équivalentes pour les 17 autres nœuds peuvent être trouvées dans l'annexe B.2.2 pages 393 à 394.

	Longitude	Latitude		Longitude	Latitude
Bko	12,549217	-1,601759	Ngo	11,19271	-1,478339
Bo	13,63009	-2,055776	Oba	13,775131	-1,444938
Bt	11,062316	-1,039871	Oi	14,032493	1,10669
F	11,50188	2,087417	Ol	12,660607	-1,401022
Ho	13,611636	-1,854028	Pa	12,732684	-1,382183
La	12,886063	-0,994725	Pu	11,143998	-2,576079
Ma	11,074904	-2,802026	Root	12,0563195	-1,8227665
My	9,356016	-1,236336	Tek	13,827594	-1,240562
N	13,9636045	1,138823	Ts	11,194935	-1,553732
Ng	11,971143	-1,18442	Wn	11,891512	2,098697

TAB. 5.1 : Valeurs médianes des longitudes et latitudes inférées lors de l'analyse bayésienne.

échantillon de presque 1500 itérations. Étant donné que les résultats de ces échantillons réduits sont semblables à l'échantillon complet, j'ai choisi de présenter les résultats avec l'échantillon complet. Il faut cependant garder à l'esprit que ce sont des résultats provisoires qui demandent à être reproduits avec les paramètres suggérés ci-dessus.

5.3.2 Localisations inférées des nœuds

Pour chacun des 20 nœuds dont la position ancestrale a été inférée par l'analyse bayésienne présentée ci-dessus, la médiane des valeurs de longitude et de latitude des 10 000 itérations échantillonnées ont été calculées. Ces valeurs médianes sont données dans le tableau 5.1. Ces coordonnées définissent l'emplacement inféré de chaque nœud, représenté sur la figure 5.8 page 170.

5.3.3 Cas d'incertitudes

5.3.3.1 Distribution dans les itérations échantillonnées

La figure 5.7 page 169 montre l'histogramme des valeurs prises par la longitude et la latitude dans l'ensemble des échantillons pour trois nœuds (*Oi*, *Ol* et *N*). On voit que pour certains nœuds, comme *Oi* (figure 5.7a) les valeurs échantillonnées sont toutes regroupées en un seul pic très net, écartant toute ambiguïté possible concernant la localisation inférée. Dans d'autres cas, comme *Ol* (figure 5.7b) ou plus nettement encore *N* (figure 5.7c), les valeurs échantillonnées forment plusieurs pics révélant une incertitude quant à la position inférée : une valeur est dominante mais d'autres valeurs ont aussi une probabilité non négligeable. Ainsi, dans le cas du nœud *N*, la longitude est le plus fréquemment échantillonnée autour de 13,97 mais des longitudes proches de 13,35 sont rencontrées relativement souvent ainsi que, dans une moindre mesure, des longitudes autour de 13,51. La même observation peut-être faite pour la latitude avec des pics d'intensité décroissante aux valeurs de 1,13, 1,42 et 1,38. Cette indétermination est notable pour les nœuds

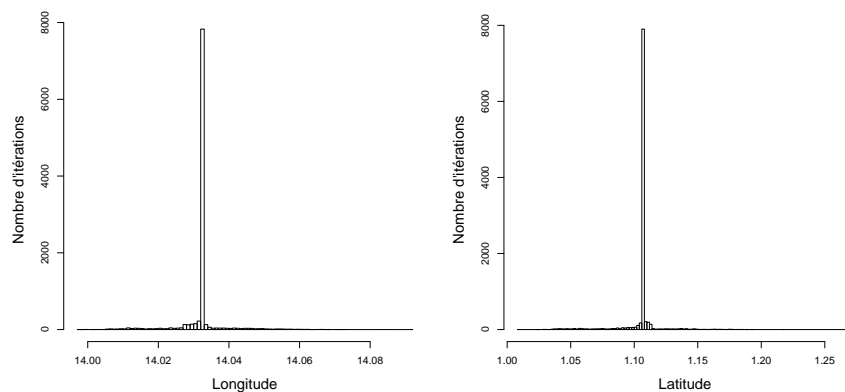
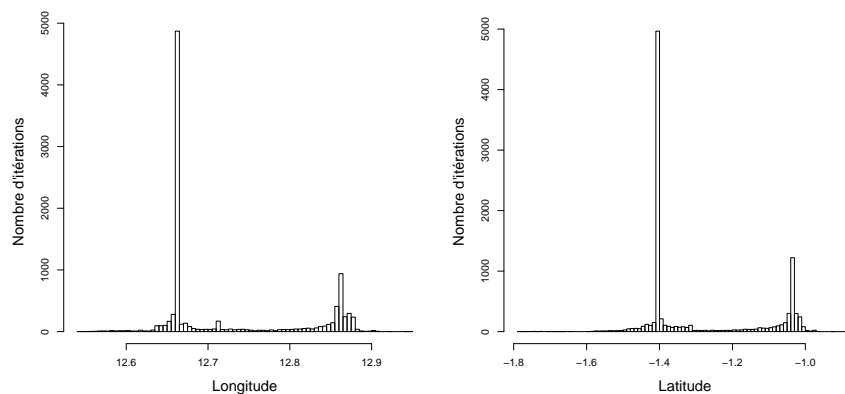
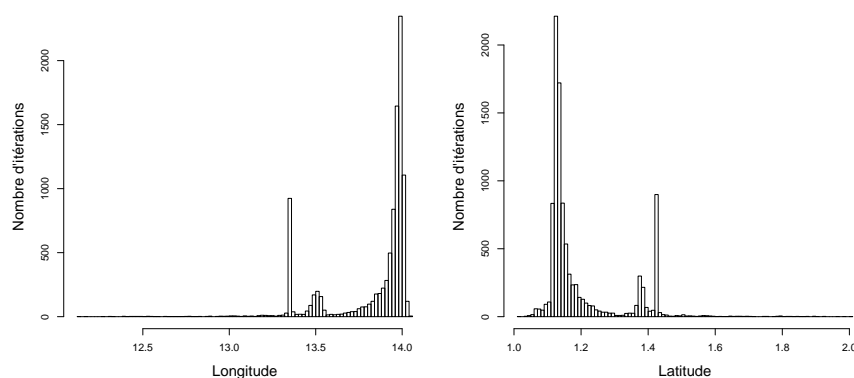
(a) O_i (b) O_l (c) N

FIG. 5.7 : **Histogramme des valeurs de longitude et de latitude de trois nœuds.** Chaque graphique représente l'histogramme des valeurs de longitude (gauche) et de latitude (droite) dans l'ensemble des itérations échantillonnées pour trois nœuds : O_i (a), O_l (b) et N (c). Les représentations équivalentes pour les 17 autres nœuds peuvent être trouvées dans l'annexe B.2.2 pages 393 à 394.

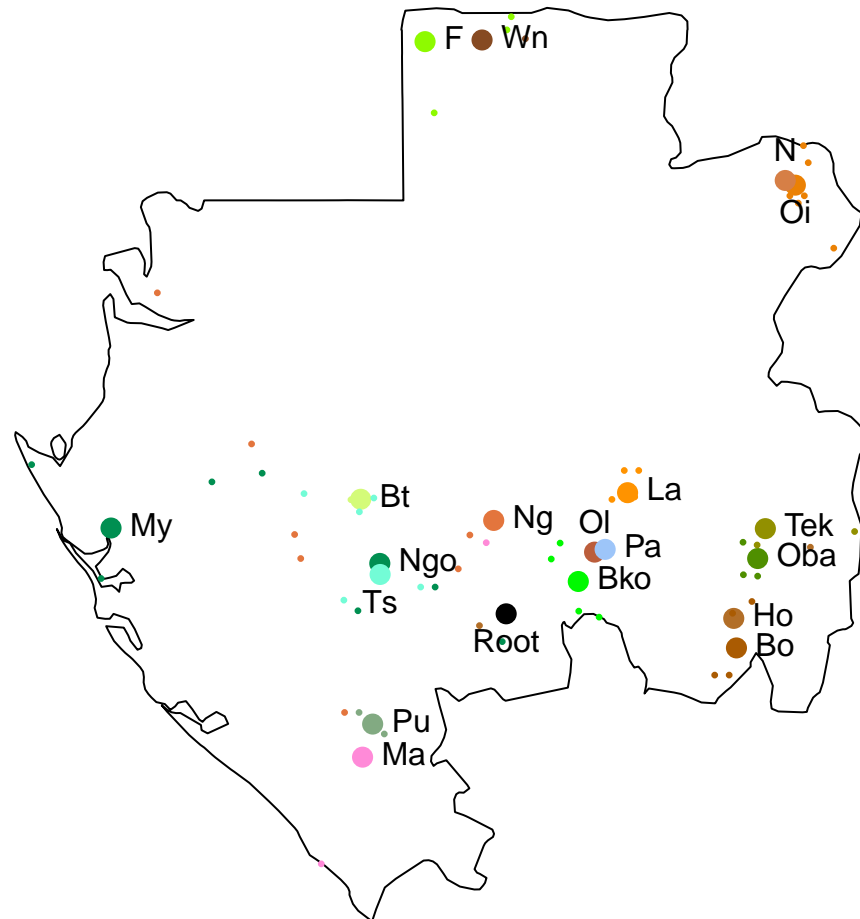


FIG. 5.8 : **Localisation des positions inférées des 20 nœuds étudiés.** Les cercles de grand diamètre montrent la position inférée de chacun des 20 nœuds. Les coordonnées de chaque nœud correspondent à la médiane des valeurs de longitude et de latitude des 10 000 itérations échantillonnées. Les couleurs utilisées pour chaque nœud sont reprises de la figure 5.1 page 160, modifiée d'après la figure 2 de l'article de Le Bomin et coll. [240]. Les cercles de petit diamètre représentent chacune des 58 populations étudiées dans ce même article. La couleur utilisée pour ces populations correspond à celle du groupe du plus bas niveau auquel appartient la population, visible sur la figure 5.1. Par exemple les Téké de Obégué sont de la couleur du groupe *Tek*, et non pas de celle du groupe *Ho* ou du groupe *Pa*, auxquels cette population appartient bien mais qui sont des groupes de plus haut niveau.

Ma^{30} , N^{31} , Ol^{32} , Pa^{33} , $Root^{34}$ et Wn^{35} . On peut la visualiser sur les figures correspondantes à chaque nœud dans l'annexe B.2.2 pages 393 et suivantes.

5.3.3.2 Association entre longitude et latitude d'un même nœud

Pour chacun de ces nœuds, les valeurs échantillonnées pour la latitude et pour la longitude ne sont pas indépendantes. On voit par exemple sur la figure 5.7b page 169 que la longitude inférée du nœud Ol est proche soit de 12,66, soit de 12,86 et sa latitude inférée est proche soit de $-1,40$, soit de $-1,04$. Si l'on représente les valeurs échantillonnées en montrant de manière conjointe la latitude et la longitude du nœud Ol , on observe cependant que les valeurs prises par la longitude ne sont pas indépendantes de celles prises par la latitude. La figure 5.9a page 172 montre ainsi que si, pour une itération donnée, la longitude du nœud Ol est autour de 12,66, sa latitude sera presque toujours proche de $-1,40$. Inversement, si la longitude est autour de 12,86, sa latitude sera presque toujours proche de $-1,04$. Le résultat de l'analyse bayésienne propose donc deux positions possibles pour le nœud Ol : une de coordonnées $(12,86 ; -1,04)$ ($Ol1$, correspondant à 32% des itérations échantillonnées) et une de coordonnées $(12,66 ; -1,40)$ ($Ol2$, 68% des itérations échantillonnées).

Cette analyse a été répétée pour chacun des 6 nœuds dont la position inférée n'est pas complètement déterminée. Les résultats de l'inférence bayésienne suggèrent donc une indétermination entre 3 positions pour le nœud Ma^{36} , 3 positions pour le nœud N^{37} , 2 positions pour le nœud Ol^{38} , 4 positions pour le nœud Pa^{39} , 4 positions pour le nœud $Root^{40}$ et 2 positions pour le nœud Wn^{41} .

5.3.3.3 Caractérisation des positions alternatives de chaque nœud

On voit que, pour chacun de ces 6 nœuds, la position inférée est incertaine et l'on peut visualiser (par exemple sur la figure 5.9a page 172) un nombre limité de pics regroupant la plupart des itérations. Chacun de ces pics représente donc une position possible du nœud, d'autant plus vraisemblable que les coordonnées correspondantes sont trouvées dans un grand nombre d'itérations échantillonnées.

Afin de caractériser chacun de ces pics, j'ai procédé de la manière suivante. J'ai cherché la ou les lignes droites séparant le plus nettement les différents pics observés. Par exemple, pour le nœud

30. ↑ Voir figure B.7 page 400.

31. ↑ Voir figure B.9 page 402.

32. ↑ Voir figure B.14 page 407.

33. ↑ Voir figure B.15 page 408.

34. ↑ Voir figure B.17 page 410.

35. ↑ Voir figure B.20 page 413.

36. ↑ Voir figure B.21 page 415.

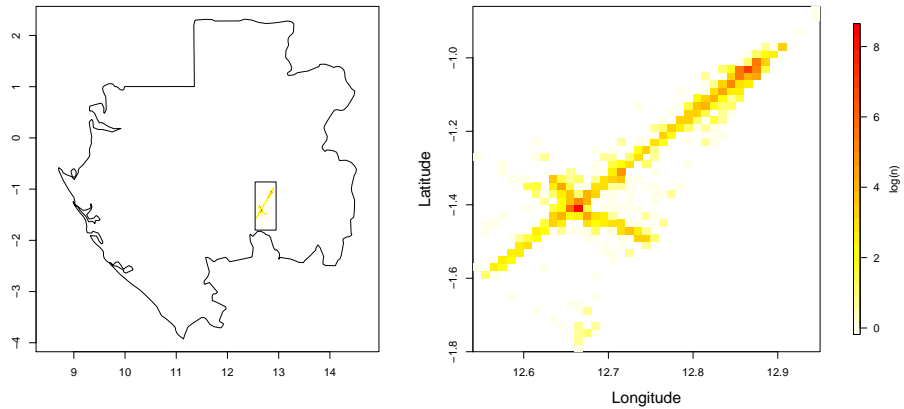
37. ↑ Voir figure B.22 page 416.

38. ↑ Voir figure B.23 page 417.

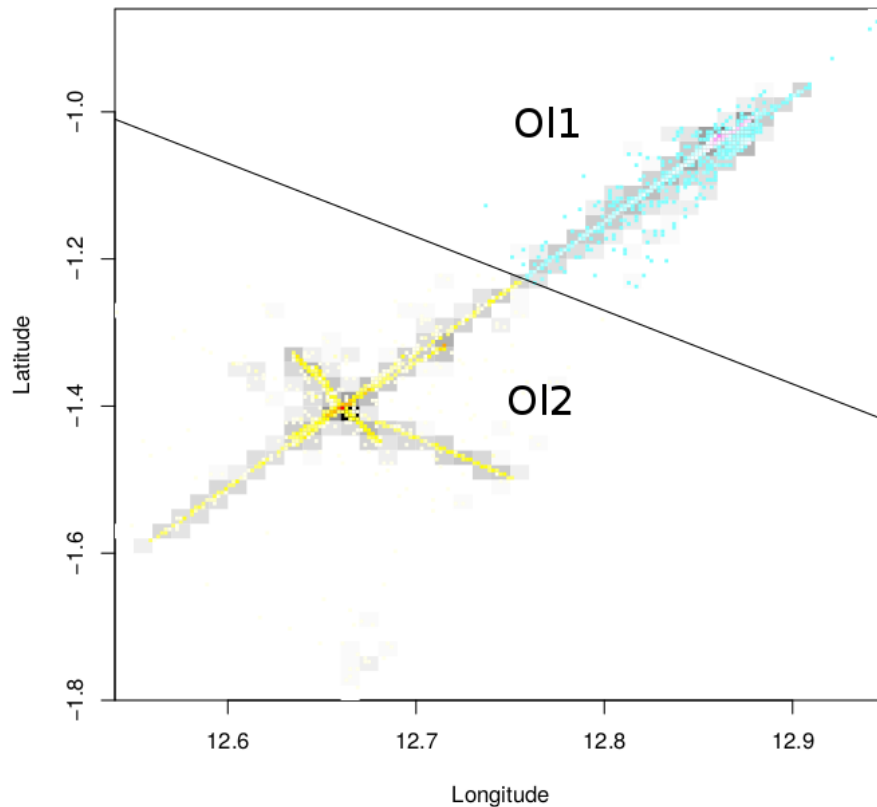
39. ↑ Voir figure B.24 page 418.

40. ↑ Voir figure B.25 page 419.

41. ↑ Voir figure B.26 page 420.



(a) Localisation les plus fréquemment échantillonnées



(b) Représentation des deux zones définies

FIG. 5.9 : **Localisations les plus fréquemment échantillonnées pour le nœud O_l .** Pour chaque zone de $\frac{1}{100}$ ème de degrés de côté, le nombre d'itérations échantillonnées dont la longitude et la latitude du nœud O_l correspondent à cette zone est calculée. Plus ce nombre est élevé, plus la couleur représentant cette zone est proche du rouge. Afin de mieux visualiser la répartition des valeurs, c'est le logarithme du nombre d'itérations échantillonnées qui est représenté. La partie gauche de la figure (a) représente la fréquence des localisations échantillonnées sur la carte du Gabon. La partie droite de la figure (a) offre une visualisation détaillée de la zone encadrée dans la partie gauche. La couleur est d'autant plus proche du rouge qu'un grand nombre d'itérations ont été échantillonnées à la position correspondante, et varie en fonction du logarithme du nombre d'itérations ($\log(n)$). La figure (b) montre le découpage de la figure (a) par une ligne oblique définissant deux zones et donc déterminant deux positions possibles : O_{l1} (couleurs froides) et O_{l2} (couleurs chaudes).

Ol , les deux pics principaux peuvent être séparés par la droite d'équation $y = 11,53 - 1,00 * x$ ⁴², tracée sur la figure 5.9b page 172. Cette droite définit deux zones, une correspondant au pic $Ol1$ (au-dessus de la droite) et l'autre au pic $Ol2$ (au-dessous). Chaque itération est donc assignée à l'une des deux zones, selon les valeurs prises par la longitude et la latitude du nœud Ol dans cette itération. La zone correspondant à $Ol1$ regroupe ainsi 32% de l'ensemble des itérations échantillonnées et la zone correspondant à $Ol2$ en regroupe 68%.

J'ai ensuite calculé, pour chaque zone ainsi définie, la valeur médiane de la longitude et de la latitude de l'ensemble des itérations appartenant à cette zone. Ces deux valeurs déterminent les coordonnées de la position alternative correspondant à cette zone. On obtient par exemple les coordonnées (12,86 ; 1,04) pour la position $Ol1$ et les coordonnées (12,66 ; 1,40) pour la position $Ol2$. Les positions alternatives de chaque nœud sont représentées sur la figure 5.10 page 174. Leurs coordonnées⁴³ ainsi que les figures permettant de visualiser les zones correspondantes⁴⁴ sont listées dans le tableau 5.2 page 175⁴⁵.

42. ↑ Où x et y sont respectivement la longitude et la latitude.

43. ↑ Colonne *Coordonnées*.

44. ↑ Colonnes *Distribution* et *Zones*.

45. ↑ Comme expliqué dans la partie 5.4 page 180, j'ai choisi de ne pas discuter des résultats eux-mêmes de l'analyse, sinon d'un point de vue méthodologique. Il est cependant intéressant de faire quelques remarques sur les six nœuds pour lesquels plusieurs positions probables ont été inférées. Parmi ceux-ci, on trouve les trois nœuds de plus haut niveau de l'arbre utilisé (*Root*, *Pa* et *Ma*). Les trois autres (Ol , N et Wn) sont tous les trois des groupes de haut niveau caractérisés *a priori* comme ayant une pertinence géographique. En revanche une seule position vraisemblable a été inférée pour tous les groupes montrant, selon Le Bomin et coll. [240], une pertinence ethnonymique. On aurait pourtant pu s'attendre à ce que l'analyse permette de déterminer plus certainement une position ancestrale pour les groupes montrant une cohérence géographique.

46.

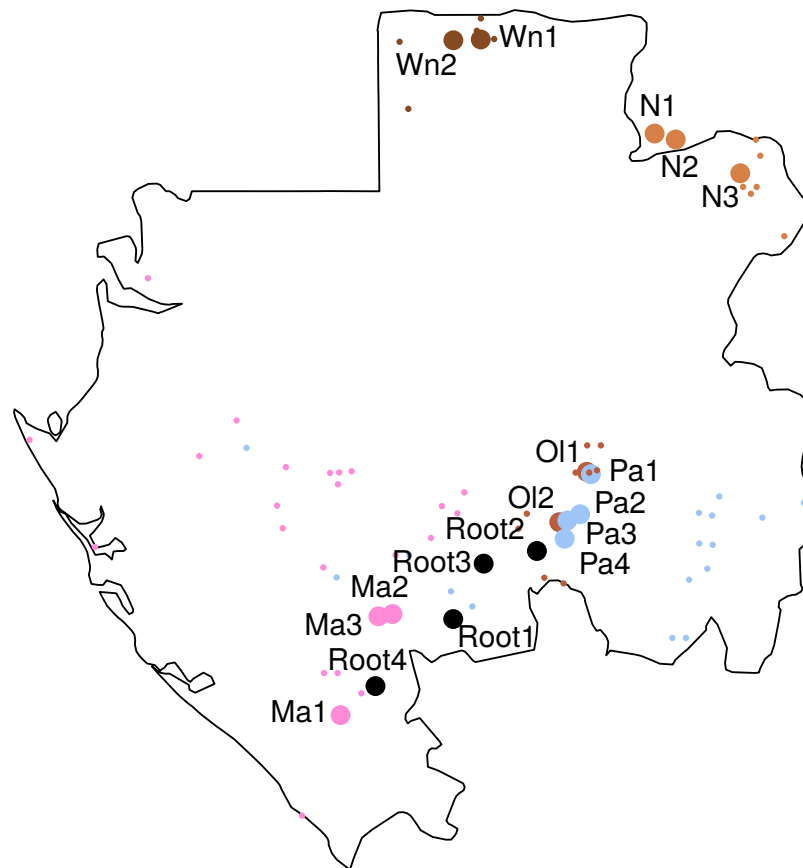


FIG. 5.10 : Localisation des différentes positions vraisemblables pour les 6 nœuds dont la position est incertaine. Pour chaque nœud pour lequel plusieurs localisations sont vraisemblables, des zones ont été définies et sont représentées dans l'annexe B.2.3 pages 415 et suivantes. Chaque cercle correspond ici à une zone d'un de ces nœuds et ses coordonnées sont les médianes des valeurs de longitude et de latitude des itérations correspondant à cette zone. Les couleurs suivent la même logique que celles de la figure 5.8.

Nœud	Position	Coordonnées	Fréquence	Distribution	Zones
Ma	Ma1	(11,07; 2,80)	61%	fig. B.7 page 400	fig. B.21 page 415
	Ma2	(11,45; 2,07)	8%		
	Ma3	(11,35; 2,08)	31%		
N	N1	(13,35; 1,42)	12%	fig. B.9 page 402	fig. B.22 page 416
	N2	(13,51; 1,38)	7%		
	N3	(13,98; 1,13)	82%		
Ol	Ol1	(12,86; 1,04)	32%	fig. B.14 page 407	fig. B.23 page 417
	Ol2	(12,66; 1,40)	68%		
Pa	Pa1	(12,89; -1,05)	7%	fig. B.15 page 408	fig. B.24 page 418
	Pa2	(12,81; -1,34)	27%		
	Pa3	(12,72; -1,39)	51%		
	Pa4	(12,70; -1,52)	12%		
Root	Root1	(11,89; -2,10)	42%	fig. B.17 page 410	fig. B.25 page 419
	Root2	(12,50; -1,61)	18%		
	Root3	(12,11; -1,70)	37%		
	Root4	(11,33; -2,59)	2%		
Wn	Wn1	(12,09; 2,10)	13%	fig. B.20 page 413	fig. B.26 page 420
	Wn2	(11,89; 2,10)	87%		

TAB. 5.2 : **Localisations possibles des nœuds pour lesquels il existe une indétermination.** Pour les 6 nœuds listés dans la colonne *Nœud*, plusieurs positions possibles ont été inférées. Les références des figures montrant la distribution des valeurs de longitude et latitude de ces nœuds dans l'ensemble des 10 000 itérations sont données dans la colonne *Distribution*. Ces positions alternatives sont nommées dans la colonne *Position*. Pour chacune de ces positions alternatives, des zones sont définies⁴⁶, comme expliqué dans la partie 5.3.3.3 page 171. La référence de la figure représentant les différentes zones pour chaque nœud est donnée dans la colonne *Zones*. Le pourcentage d'itérations (sur le total des 10 000 itérations) correspondant à chaque zone est indiqué dans la colonne *Fréquence* et les valeurs médianes des longitudes et latitudes du nœud correspondant dans ce sous-échantillon des itérations est indiqué dans la colonne *Coordonnées*.

5.3.3.4 Associations entre les positions alternatives de différents nœuds

De la même manière que les valeurs de longitude et de latitude de certains nœuds ne sont pas indépendantes dans les itérations échantillonnées, les positions alternatives des différents nœuds décrites ci-dessus ne sont pas nécessairement indépendantes. J'ai donc cherché une liaison éventuelle entre les positions alternatives de différents nœuds. Pour cela j'ai sélectionné pour chaque position caractérisée l'ensemble des itérations correspondant à cette position, comme expliqué ci-dessus⁴⁷. J'ai ensuite représenté la distribution des coordonnées géographiques de chacun des autres nœuds pour lesquels plusieurs positions sont possibles, dans ces itérations sélectionnées. La figure 5.11 page 178 montre une telle représentation pour le nœud *Ol*. J'appelle ce type de figure un *diagramme de correspondance*.

Cette figure permet de visualiser rapidement les associations entre certaines positions vraisemblables de différents nœuds. La première ligne de la figure 5.11 permet de visualiser les deux zones, mais aussi de bien comprendre ce que montre cette représentation. Dans les deux figures de cette ligne (nœud *Ol*), le nombre d'itérations correspondant à chaque zone géographique est représenté en nuances de gris, exactement de la même manière que sur la partie droite de la figure 5.9a page 172, si ce n'est que c'est l'intensité du gris qui marque les hautes fréquences. Cette partie grise constitue le « fond » de la figure et permet de visualiser en arrière plan les données correspondant à l'ensemble des itérations. Par dessus ce fond, je rajoute une couche représentant seulement une partie des itérations, en nuances de couleurs chaudes. Dans le cas de la partie gauche (colonne *Ol1*), seules les itérations appartenant à la zone *Ol1* sont représentées en couleurs chaudes. Dans la partie droite (colonne *Ol2*), ce sont les itérations appartenant à la zone *Ol2* qui sont montrées en couleurs chaudes. On visualise donc sur la première ligne (*Ol*) la zone *Ol1* (colorée sur la figure de gauche) et la zone *Ol2* (colorée sur la figure de droite).

La deuxième ligne montre de la même manière les positions du nœud *N* dans tout ou partie des itérations. Sur la partie gauche, seules les itérations de la zone *Ol1* sont conservées et uniquement celles de la zone *Ol2* dans la partie droite. On voit que l'ensemble de la zone est coloré dans les deux cas. Le fait qu'une itération positionne le nœud *Ol* à l'emplacement *Ol1* ne semble donc pas déterminer si le nœud *N* sera préférentiellement à l'une des positions définies pour ce nœud⁴⁸. La même constatation peut-être faite pour les nœuds *Wn* (ligne 3), *Root* (ligne 4) et *Ma* (ligne 6).

On observe un motif différent sur la ligne 5, correspondant au nœud *Pa*. Lorsqu'on ne sélectionne que les itérations appartenant à la zone *Ol1* (partie gauche), seule une partie de l'ensemble des valeurs observées pour ce nœud est représentée, correspondant aux positions les plus au nord-est. Inversement, dans les itérations appartenant à la zone *Ol2* (partie droite), le *Pa* est toujours situé dans la zone la plus au sud-est. On peut comparer cette répartition avec les zones définies pour le nœud *Pa* et représentées sur la figure B.24 page 418. On constate ainsi une association entre la position *Ol1* du nœud *Ol* et les positions *Pa1* et *Pa2* du nœud *Pa* d'une part, et entre

47. ↑ Voir paragraphe *Caractérisation des positions alternatives de chaque nœud* page 171.

48. ↑ Voir le tableau 5.2 page 175 et la figure B.22 page 416.

la position *Ol2* du nœud *Ol* et les positions *Pa3* et *Pa4* du nœud *Pa* d'autre part.

J'ai ensuite cherché à vérifier quantitativement ces observations. Pour cela, pour une position possible d'un nœud (par exemple *Ol1*), j'ai calculé la proportion des itérations qui appartiennent aux différentes zones d'un autre nœud. Ces données sont présentées dans le tableau 5.3 page 179. Je cherche par exemple à vérifier s'il existe bien une association entre certaines positions du nœud *Ol* et certaines du nœud *Pa*. J'ai donc sélectionné seulement les itérations correspondant à la zone *Ol1*. J'ai ensuite compté dans combien de ces itérations le nœud *Pa* est proche soit de l'emplacement *Pa1*, soit de *Pa2*, soit de *Pa3* soit de *Pa4*. Le tableau 5.3 page 179 indique ainsi que les itérations de la zone *Ol1* sont réparties à 22% dans la zone *Pa1*, à 71% dans la zone *Pa2*, à 6% dans la zone *Pa3* et à 0% dans la zone *Pa4* (lignes *Pa1* à *Pa4*, colonne *Ol1*). On peut comparer avec la répartition de l'ensemble des itérations, indiquée dans la colonne « Tout » du tableau 5.3 page 179 : au total, le nœud *Pa* est placé dans 7% des cas dans la zone *Pa1*, dans 27% des cas dans la zone *Pa2*, dans 51% des cas dans la zone *Pa3* et dans 12% des cas dans la zone *Pa4*. On voit donc que la répartition des différentes positions du nœud *Pa* n'est pas du tout la même si l'on considère toutes les itérations ou seulement celles où le nœud *Ol* est placé à la position *Ol1* : dans ce dernier cas, le nœud *Pa* est presque systématiquement dans la zone *Pa1* et jamais dans la zone *Pa2*. Inversement, si l'on ne garde que les itérations où le nœud *Ol* est dans la zone *Ol2*, le nœud *Pa* est quasiment toujours dans les zones *Pa3* et *Pa4*⁴⁹ et presque jamais dans les zones *Pa1* et *Pa2*⁵⁰. Cela confirme donc bien l'association entre la position *Ol1* du nœud *Ol* et les positions *Pa1* et *Pa2* du nœud *Pa* d'une part, et entre la position *Ol2* du nœud *Ol* et les positions *Pa3* et *Pa4* du nœud *Pa* d'autre part.

On peut aussi constater qu'il n'y a pas d'association avec certaines positions du nœud *Ol* et celles d'autres nœuds. Par exemple les itérations de la zone *Ol1* se répartissent à 56%, 9% et 34% entre les zones *Ma1*, *Ma2* et *Ma3*, des valeurs sensiblement similaires à la répartition dans l'ensemble des itérations (respectivement 61%, 8% et 31%).

De la même manière, l'étude détaillée des diagrammes de correspondance pour les nœuds *Ma*⁵¹, *N*⁵², *Ol*⁵³, *Pa*⁵⁴, *Root*⁵⁵ et *Wn*⁵⁶, ainsi que du tableau 5.3 page 179, nous permet d'identifier d'autres associations entre certaines positions inférées de ces nœuds. Je ne détaillerai pas de manière aussi détaillée que ci-dessus l'analyse des figures permettant d'arriver à ces conclusions, me contentant de décrire les principales associations observées, à savoir :

- la position *Ma1* du nœud *Ma* et les positions *Root1* et *Root2* du nœud *Root*;
- et la position *Wn1* du nœud *Wn* et la position *N2* du nœud *N* (de manière moins prononcée).

49. ↑ Respectivement dans 72% et 17% des cas.

50. ↑ Respectivement dans 0% et 7% des cas.

51. ↑ Figure B.21 page 415.

52. ↑ Figure B.22 page 416.

53. ↑ Figure B.23 page 417.

54. ↑ Figure B.24 page 418.

55. ↑ Figure B.25 page 419.

56. ↑ Figure B.26 page 420.

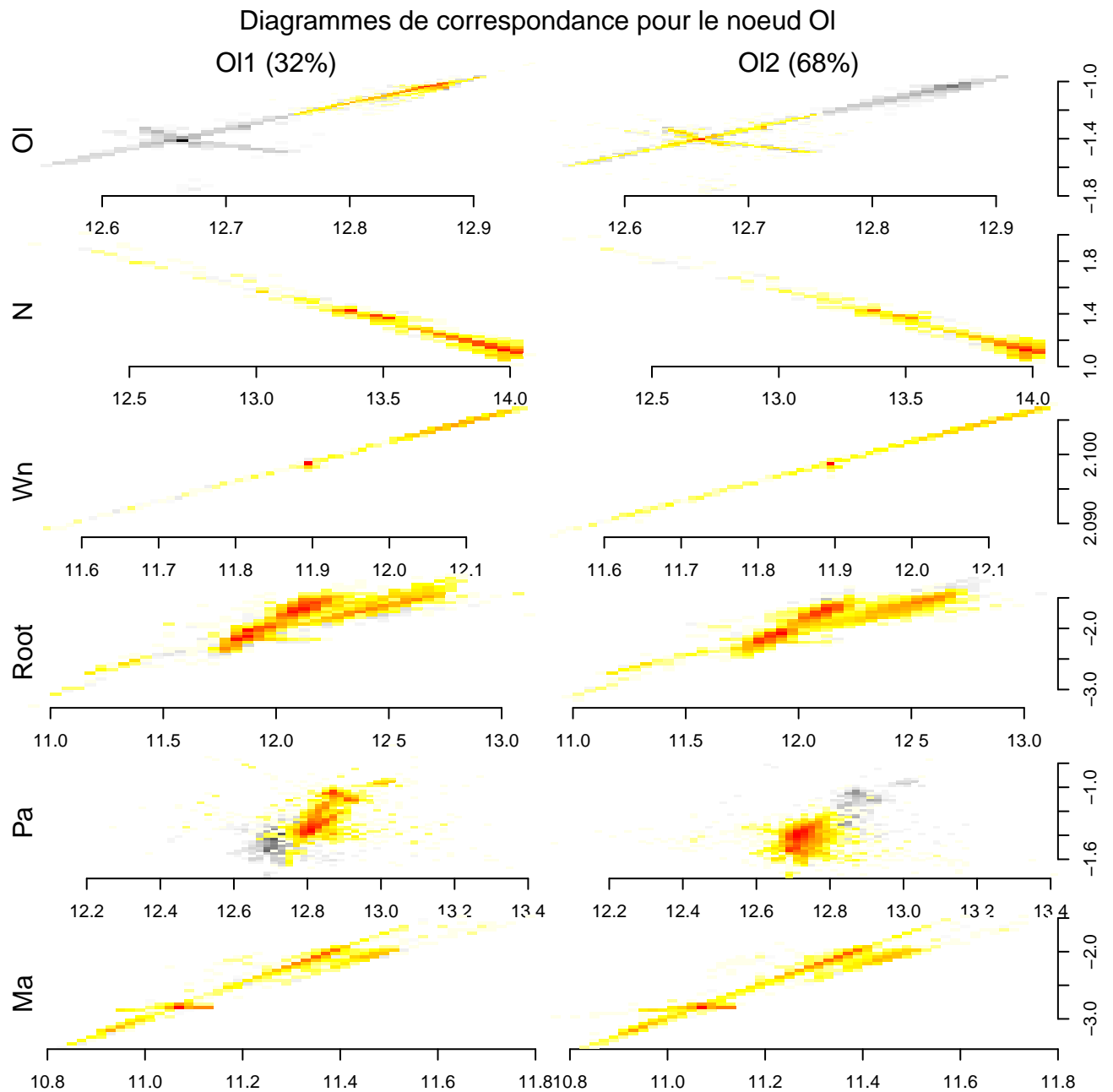


FIG. 5.11 : **Diagramme de correspondance pour le noeud OI.** Chaque ligne correspond à un noeud dont l'analyse bayésienne donne plusieurs positions vraisemblables. Chaque figure montre la fréquence à laquelle le noeud correspondant est trouvé dans les différentes itérations (voir la figure 5.9a page 172 et les explications données dans le paragraphe *Association entre longitude et latitude d'un même noeud* page 171 pour une bonne compréhension de cette représentation). Dans les deux colonnes, la fréquence dans l'ensemble des 10 000 itérations est représentée en nuances de gris : plus la couleur est foncée, plus cette position est échantillonnée souvent. Cette représentation en nuances de gris constitue le fond de la figure. Chaque colonne correspond à une des deux positions possibles pour le noeud *OI* et représente en nuance de couleurs chaudes la fréquence dans les seules itérations correspondant à ce noeud : plus la couleur est proche du rouge, plus cette position est fréquente. Ainsi, dans la colonne de gauche (*OI1*), seules les itérations correspondant à la zone du pic *OI1* sont représentées en couleurs chaudes. De même, dans la colonne de droite (*OI2*), seules les itérations correspondant à la zone du pic *OI2* sont représentées en couleurs chaudes.

	Tout	Ma1	Ma2	Ma3	N1	N2	N3	O11	O12	Pa1	Pa2	Pa3	Pa4	Root1	Root2	Root3	Root4	Wn1	Wn2
Ma1	61				65	50	61	56	63	61	57	61	64	99	85	4	95	55	61
Ma2	8				5	11	8	9	8	7	10	8	6	0	2	21	1	10	8
Ma3	31				30	39	31	34	30	32	34	31	29	1	12	75	4	35	31
N1	12	12	6	11				21	7	33	15	4	22	16	4	11	9	3	13
N2	7	5	9	8				13	3	15	11	1	13	5	6	8	8	49	0
N3	82	82	85	80				66	89	52	74	94	65	79	89	81	84	48	87
O11	32	29	37	35	58	64	25			99	83	4	0	27	33	36	37	48	29
O12	68	71	63	65	42	36	75			1	17	96	100	73	67	64	63	52	71
Pa1	7	7	6	7	20	16	4	22	0					8	2	8	10	11	6
Pa2	27	25	32	29	35	46	24	71	7					22	36	29	30	36	26
Pa3	51	52	49	50	20	11	59	6	72					54	45	51	43	31	54
Pa4	12	12	9	11	23	23	9	0	17					14	11	10	13	18	11
Root1	42	69	1	2	58	33	41	36	45	47	34	45	49					37	43
Root2	18	26	5	7	7	18	20	19	18	6	25	16	18					18	18
Root3	37	2	94	91	34	47	37	42	35	44	40	38	32					43	37
Root4	2	3	0	0	1	2	2	2	2	2	2	1	2					2	2
Wn1	13	12	16	14	3	95	7	19	10	20	17	8	19	11	13	14	14		
Wn2	87	88	84	86	97	5	93	81	90	80	83	92	81	89	87	86	86		

TAB. 5.3 : **Répartition croisée des itérations échantillonnées entre les positions alternatives de différents nœuds.** La valeur donnée pour la ligne correspondant à la position X (par exemple *Ma1*) et la colonne correspondant à la position Y (par exemple *N1*) correspond au pourcentage d'itérations appartenant à la zone Y qui sont aussi dans la zone X. On voit par exemple que, si l'on ne garde que les itérations correspondant à la zone *N1*, dans 65% de celles-ci les coordonnées du nœud *Ma* sont proches de (11,07; 2,80) (zone *Ma1*), dans 5% elles sont proches de (11,45; 2,07) (zone *Ma2*) et dans 30% des cas elles sont proches de (11,35; 2,08) (zone *Ma3*). La colonne *Tout* indique la répartition en considérant toutes les itérations. Elle correspond donc à la colonne *Fréquence* du tableau 5.2 page 175. Ainsi 61% des itérations appartiennent à la zone *Ma1*, 8% à la zone *Ma2* et 31% à la zone *Ma3*. Les valeurs exprimant la répartition croisée des positions alternatives d'un même nœud ne sont pas données car elles sont triviales (100% s'il s'agit de la même position, 0% s'il s'agit de deux positions alternatives différentes d'un même nœud). Toutes les valeurs sont exprimées en %.

5.4 Discussion

Je présente dans cette partie quelques réflexions concernant l'analyse présentée ci-dessus. Je discuterai rapidement de quelques améliorations possibles de la méthode⁵⁷. Je ne m'y étendrai cependant pas outre mesure car, comme je le montre ensuite, cette analyse souffre de nombreuses lacunes dans ses fondements théoriques et méthodologiques⁵⁸. Pour cette même raison, je ne discuterai pas des résultats produits par l'analyse, mais seulement de questions méthodologiques. La constatation de ces insuffisances a constitué un tournant dans l'orientation de ma thèse. Cela m'a conduit à axer mes recherches vers des questions plus élémentaires, mais fondamentales, afin de progresser dans l'inférence historique à partir de données musicales⁵⁹.

5.4.1 Méthodologie

5.4.1.1 Autocorrélation et échantillonnage

Comme on l'a vu⁶⁰, il existe une autocorrélation dans la vraisemblance des itérations échantillonnées, importante jusqu'à 7 échantillons successifs. Cette observation affaiblit la confiance que l'on peut avoir dans les résultats. Il serait nécessaire, pour obtenir un jeu de données contenant le même nombre d'échantillons tout en s'affranchissant de cette autocorrélation, de refaire l'analyse sur 7 milliards d'itérations.

5.4.1.2 Surparamétrisation

J'ai choisi ici d'inférer les positions de 20 nœuds ayant une pertinence géographique, ethnonymique ou du point de vue du système de filiation. Comme chaque position est décrite par deux paramètres (longitude et latitude), l'analyse bayésienne décrite ici cherche simultanément les valeurs de 40 paramètres. On peut ainsi penser que l'analyse est *surparamétrisée* : le nombre de paramètres à estimer est trop important compte tenu des données disponibles. Il aurait sans doute été plus raisonnable de se limiter aux nœuds dont les indices de *robustesse* (en particulier les valeur de *bootstrap*) sont les plus élevés.

5.4.1.3 Arbre des relations entre patrimoines

J'ai supposé ici que les patrimoines musicaux se diversifient selon un modèle strictement *arborescent*. J'ai utilisé pour cela un arbre publié par Le Bomin et coll. [240] issu de l'analyse *cladistique* de 58 patrimoines musicaux. Cet arbre est un *arbre de consensus*⁶¹ de plus de 270 000 arbres également *parcimonieux*. Cependant le logiciel *BayesTraits* permet d'utiliser lors de l'analyse bayésienne un ensemble d'arbres, au lieu d'un seul arbre. Ceci permet de prendre en compte l'incertitude quant à la *topologie* de l'arbre modélisant les relations entre les groupes. L'analyse pourrait donc être améliorée en incluant cet ensemble de plus de 270 000 arbres plutôt que le seul arbre de consensus, afin de moins contraindre l'analyse et autoriser une souplesse concernant les

57. ↑ Voir partie 5.4.1 page 180.

58. ↑ Voir partie 5.4.2 page 181.

59. ↑ Voir partie 5.5 pages 184 à 185, chapitre 6 pages 187 à 245 et chapitre 7 pages 247 à 274.

60. ↑ Voir partie 5.3.1.2 page 165.

61. ↑ Cet arbre est construit en suivant la règle de la majorité (*majority rule*).

relations entre patrimoines qui reflète mieux notre niveau de confiance dans les données utilisées. L'utilisation de cet arbre soulève par ailleurs d'autres questions, discutées dans la partie 5.4.2.1.

5.4.2 Limite du modèle utilisé pour l'inférence

5.4.2.1 Le modèle arborescent d'évolution des patrimoines

L'inférence proposée ici suppose que les **patrimoines musicaux** évoluent par ramification le long d'un arbre, un patrimoine ancestral se divisant en deux patrimoines descendants. Cependant l'arbre proposé par Le Bomin et coll. [240] n'est pas présenté comme un arbre modélisant l'évolution des patrimoines musicaux. En effet, cet article cherche à montrer la présence d'un signal de transmission verticale dans les patrimoines musicaux du Gabon, pouvant renseigner sur la manière dont les musiques se transmettent. L'analyse cladistique effectuée, en particulier l'analyse des *CIs* et des *RIs*, leur permet de discuter de la verticalité de la transmission. Le Bomin et coll. [240] considèrent aussi que l'obtention de **clades** pertinents selon des critères extramusicaux renforce l'hypothèse d'un signal vertical. L'arbre obtenu n'est donc pas présenté comme un modèle des relations historiques de filiation entre les patrimoines musicaux, mais bien comme un outil permettant de discuter de la question de la verticalité de la transmission.

Il n'apparaît donc pas très judicieux d'utiliser un tel arbre pour inférer des localisations ancestrales comme j'ai pu le faire ici. En effet, dans la méthode utilisée l'arbre de départ doit justement constituer un modèle de la filiation des groupes considérés. Même si cet arbre n'a pas été conçu dans ce but, on pourrait éventuellement l'utiliser comme modèle de l'histoire de ces patrimoines musicaux, mais cependant à la condition *sine qua non* que l'on ait évalué auparavant à quel point cette modélisation est justifiée. Comme l'objet de l'article de Le Bomin et coll. [240] n'est pas d'estimer à quel point cet arbre peut servir de modèle à l'histoire des patrimoines musicaux du Gabon, il m'aurait fallu d'abord, avant de poursuivre mon analyse, me pencher sur cette question.

5.4.2.2 Le patrimoine musical comme unité évolutive

Utiliser la notion de **patrimoine musical** comme unité d'analyse pour l'analyse historique présuppose que le patrimoine musical constitue une unité pertinente dans la dimension temporelle. Or un patrimoine musical est constitué par définition d'un ensemble de caractéristiques musicales partagées par une population [240, p. 2] et regroupe donc des données de natures hétérogènes⁶². On peut donc se demander si un patrimoine musical est une unité de transmission suffisamment cohérente⁶³ pour que l'on puisse considérer qu'un patrimoine donné à une époque donnée soit le descendant d'un patrimoine existant à une période antérieure. Cette hypothèse n'a, à ma connaissance, jamais été testée empiriquement. Cependant certains indices peuvent

62. ↑ Par exemple des **paramètres musicaux internes** tels que la **métricité**, les **formules rythmiques** ou les **échelles** musicales, des **paramètres performatifs** tels que les instruments de musique utilisés, ou encore **paramètres contextuel** tels que les cérémonies dans lesquelles interviennent ces musiques.

63. ↑ Voir à ce sujet la discussion dans la partie 7.2.3.2 pages 264 à 266.

faire douter de sa validité, comme je le détaillerai plus loin⁶³. En ce qui concerne les musiques du Gabon en particulier, l'article de Le Bomin et coll. [240] relève des valeurs de *CI*s et *RI*s assez différentes entre différents types de paramètres musicaux (intrinsèques, performatifs et contextuels)⁶⁴, suggérant des transmissions différentes pour chacun d'entre eux. Par ailleurs nos récentes observations sur le terrain suggèrent une grande perméabilité des répertoires entre populations mais aussi des pièces entre répertoires différents. Au vu des quelques éléments dont nous disposons, il me paraît donc plus raisonnable de considérer un patrimoine musical comme un ensemble hétérogène du point de vue de la transmission et donc de la filiation. Cette supposition nécessite de toute façon d'être testée empiriquement, ouvrant le chemin pour une identification plus fine d'unités pertinentes du point de vue historique.

5.4.2.3 Patrimoines musicaux et populations

L'analyse présentée ici suppose que les patrimoines musicaux se transmettent et évoluent comme des unités cohérentes, qui peuvent se ramifier en plusieurs patrimoines musicaux. Inférer la localisation ancestrale d'un nœud de l'arbre reviendrait donc à chercher la zone géographique où un tel patrimoine ancêtre aurait été pratiqué. On voit bien que le « mouvement » d'un patrimoine musical ne peut être complètement décorrélé des populations qui le pratiquent. Si un ensemble de pratiques musicales identiques est retrouvé à un endroit différent, on peut l'expliquer soit par le déplacement des personnes porteuses de ces pratiques⁶⁵, soit par une transmission de ces pratiques aux personnes vivant à cet endroit (et probablement par une combinaison de ces deux hypothèses). Il apparaît que les fondements de l'analyse présentée ici ne sont pas très bien définis. Que cherche-t-on réellement à inférer ? S'il s'agit de la localisation d'un patrimoine musical ancestral, cette notion est très floue : on ne sait comment elle s'articule avec la localisation des populations et on ne sait pas vraiment quelle pertinence a cette notion de « patrimoine ancestral ». S'il s'agit des déplacements des populations ou des échanges musicaux entre populations, il est difficile de savoir à quel point nos données nous permettent d'y avoir accès.

5.4.3 Pertinence de l'analyse

On voit donc que l'analyse présentée ci-dessus a été mal conçue au départ et je propose maintenant quelques réflexions épistémologiques visant à clarifier les questionnements initiaux. On peut à mon avis envisager l'inférence historique à partir de données musicales selon deux perspectives différentes et radicalement distinctes : évaluer une méthode nouvelle ou contribuer à améliorer notre connaissance de l'histoire.

64. ↑ Voir partie 7.2.3.2 page 264.

65. ↑ Éventuellement en remplaçant les personnes qui y vivaient auparavant, on parle alors de *diffusion démique* (*demic diffusion* en anglais).

5.4.3.1 Tester une méthode

Si on ne peut avoir une confiance suffisante dans l'analyse effectuée, que ce soit au niveau de la méthode ou des données utilisées, elle ne peut permettre de proposer des hypothèses fiables sur le passé. L'intérêt d'une telle analyse réside donc dans son caractère innovant et instable, qui nous permet d'améliorer les méthodes utilisées. Pour arriver à cette fin, il faut donc se donner les moyens d'évaluer *a posteriori* les résultats obtenus et par ce biais la qualité de l'analyse.

Comme je l'ai montré ci-dessus, les bases de l'analyse bayésienne effectuée dans ce chapitre ne sont que très peu solides. Si l'on considère que cette analyse a pour but d'expérimenter et d'évaluer une nouvelle méthode pour étudier des données musicales, il est nécessaire de pouvoir tester les résultats obtenus afin de pouvoir apprécier l'intérêt de la méthode. En l'occurrence, il faudrait donc disposer de données indépendantes qui devraient converger avec mes résultats. Je pourrais par exemple utiliser les résultats de Van der Veen et coll. [426] qui proposent un scénario de l'expansion des langues bantoues au Gabon. Malheureusement la nature des localisations inférées restent mal définies⁶⁶ et les bases théoriques de mon analyse ne me permettent pas de supposer que les localisations inférées devraient être cohérentes avec le scénario de Van der Veen et coll. [426]. En effet, je n'ai pas assez d'éléments pour faire l'hypothèse que l'évolution temporelle et géographique des patrimoines musicaux converge avec celle des langues.

5.4.3.2 Produire de la connaissance historique

Si en revanche je dispose de bonnes raisons d'avoir confiance dans la méthode et les données utilisées, l'analyse peut servir à améliorer notre connaissance de l'histoire. Les résultats obtenus peuvent alors être confrontés à des résultats d'autres sources historiques indépendantes et permettre de mieux comprendre à la fois l'histoire telle qu'elle s'est produite et les processus mis en œuvre. Je pourrais par exemple les comparer à ceux de Van der Veen et coll. [426] afin de discuter de l'histoire des populations du Gabon ou de tester s'il existe une congruence entre histoires musicale et linguistique. En effet, comme le soutient Tucker [424], une connaissance historique est d'autant plus fiable qu'elle est supportée par des lignes indépendantes de preuves suffisamment solides. Une « ligne de preuve » ne peut donc avoir un intérêt que dans la mesure où l'on peut avoir confiance en elle. Si la confiance que l'on en a provient de la convergence avec des données indépendantes, la preuve n'est pas indépendante de ces données et ne constitue pas une ligne indépendante. Ainsi, si je choisis de me servir des données de Van der Veen et coll. [426] pour estimer la pertinence de mes résultats, ces derniers ne peuvent être réutilisés conjointement avec ceux de Van der Veen et coll. [426] pour discuter de l'histoire des musiques, langues et populations du Gabon : le raisonnement serait circulaire et les données seraient doublement utilisées (pour tester les résultats et pour produire de la connaissance historique).

66. ↑ Voir paragraphe *Patrimoines musicaux et populations* page 182.

5.5 Perspectives

5.5.1 Résumé du chapitre

J'ai présenté dans ce chapitre une analyse d'inférence bayésienne effectuée avec le logiciel *BayesTraits*. À partir d'un arbre reliant des patrimoines musicaux (issu des travaux de Le Bomin et coll. [240]) et des coordonnées géographiques du lieu de collecte de ces patrimoines, j'ai cherché à inférer l'emplacement géographique de certains **nœuds internes** de l'arbre. L'analyse effectuée pourrait être améliorée sur certains points, par exemple en augmentant le nombre d'itérations et en augmentant l'intervalle d'échantillonnage, ou encore en utilisant l'ensemble des arbres également parcimonieux produits par l'analyse de Le Bomin et coll. [240] (et non pas le seul **arbre de consensus**).

Cependant la pertinence elle-même de l'analyse est contestable. Pour que cette analyse soit intéressante, il faut qu'elle soit fondée sur des méthodes et des données solides, ou bien que l'on dispose de données indépendantes permettant de tester ces résultats. Or je n'ai que peu de données disponibles permettant de tester les résultats de mon analyse et celles dont je dispose ne peuvent être reliées à mes travaux, en partie en raison d'un manque de fondements empiriques et théoriques. Par ailleurs les bases de mon analyse sont très instables : le modèle **arborescent** d'évolution des patrimoines est discutable, l'arbre produit par Le Bomin et coll. [240] n'est pas un modèle de l'histoire de ces patrimoines et l'utilisation des **patrimoines musicaux** comme unité d'analyse me paraît inadéquate (il faudrait *a minima* évaluer la pertinence de ce choix). Plus globalement, les présupposés théoriques sur lesquels se fonde mon analyse sont mal définis et les résultats produits ne sont pas clairs : on ne sait pas vraiment ce que représentent les positions ancestrales des nœuds de l'arbre qui sont inférées ici.

Ainsi il m'est apparu que l'analyse effectuée dans ce chapitre est prématurée compte tenu de nos connaissances des musiques du Gabon et de leur évolution. Il m'a donc semblé nécessaire de réorienter mon travail afin de pouvoir donner des bases plus solides à de futurs travaux d'inférence historique à partir de données musicales. Pour cela, j'ai choisi d'aborder la question sous un angle empirique d'une part, afin de mieux connaître et comprendre les données utilisées et sous un angle théorique d'autre part afin d'affiner les fondements méthodologiques de l'inférence historique.

5.5.2 Mieux connaître nos données

L'hypothèse du patrimoine musical comme unité de transmission peut être contestée, comme je l'ai exposé ci-dessus et comme je le détaille plus loin⁶⁷. Il m'a donc paru important de mieux décrire et de mieux comprendre les différents types de données musicales disponibles et les relations qu'ils pouvaient avoir entre eux. En effet la diversité des différents paramètres musicaux utilisés par Le Bomin et coll. [240] n'a jamais été décrite. De plus, cette dernière étude suggère une transmission différente entre les paramètres musicaux intrinsèques, performatifs et

67. ↑ Voir [partie 7.2.3.2](#) page 264.

contextuels. Les paramètres musicaux intrinsèques en particulier montrent des valeurs de *CI* particulièrement élevées.

J'ai donc choisi de décrire plus finement la distribution géographique de ces paramètres intrinsèque (organisations métrique, rythmique et scalaire), ainsi que le lien aux populations et répertoires dans lesquels on les rencontre. Ces analyses sont l'objet du prochain chapitre⁶⁸.

5.5.3 Affiner les fondements théoriques et méthodologiques

On l'a vu, les fondements théoriques de l'étude de l'histoire à partir de la comparaison de données musicales restent rudimentaires. Quel type de données musicales est-il judicieux d'utiliser ? Sur quelle-s histoire-s ces données nous renseignent-elles : celle-s des musiques, des populations ? Faut-il utiliser toutes les données musicales conjointement ou séparément ? Faut-il les utiliser en même temps que des données non musicales ? Comment définir des unités pertinentes et efficaces du point de vue de l'inférence historique ? Comment évaluer la vraisemblance des résultats obtenus, quelle confiance peut-on avoir en eux ?

Toutes ces questions, largement discutées par les autres disciplines historiques, se doivent d'être posées dans le cadre particulier de l'ethnomusicologie. Si les recherches et propositions issues des autres disciplines sont éclairantes et inspirantes à de nombreux égards, l'ethnomusicologie ne peut se satisfaire de réponses toutes faites et doit les adapter à sa pratique. Le chapitre 7 de cette thèse⁶⁹ sera donc consacré à défricher quelques uns de ces questionnements en s'appuyant sur les apports d'autres disciplines historiques, notamment la systématique biologique, la linguistique historique, la critique textuelle et l'archéologie. Je m'intéresserai en particulier aux moyens pour identifier des paramètres porteurs de traces historiques, à la manière de traiter différents types de traces historiques et à la recherche d'unités homogènes et pertinentes du point de vue de l'inférence historique.

68. ↑ Voir chapitre 6 pages 187 à 245.

69. ↑ Pages 247–274.

Chapitre 6

Un aperçu de la diversité des musiques du Gabon

6.1	Introduction	188
6.1.1	Une thèse avec article	188
6.1.1.1	Objectif de cette introduction	188
6.1.1.2	Un format adapté aux contraintes	188
6.1.1.3	Limites	189
6.1.2	Contexte et objectifs	189
6.1.2.1	Contexte	189
6.1.2.2	Objectifs	189
6.2	<i>Diversity and distribution of musics in Gabon</i>	191
	<i>Introduction</i>	192
	<i>Material and methods</i>	196
	<i>Results</i>	201
	<i>Discussion</i>	213
	<i>Conclusion</i>	219
	<i>Supporting information</i>	221
	<i>References</i>	241
6.3	Conclusion et perspectives	245
	Principaux résultats	245
	Perspectives	245

6.1 Introduction

6.1.1 Une thèse avec article

6.1.1.1 Objectif de cette introduction

Le chapitre qui vient reproduit un article scientifique rédigé par Sylvie Le Bomin et moi-même, tel qu'il a été soumis à la revue *Plos One*. Le choix d'insérer ainsi ce travail est un intermédiaire entre le format de la thèse classique, complètement cohérente d'un bout à l'autre, et la thèse sur articles, composée d'une présentation argumentée d'articles scientifiques reliés entre eux par de brèves introductions permettant une contextualisation générale. Cette partie vise à introduire cet article, le contexte dans lequel il s'inscrit, ses objectifs et à préciser certains points peu ou pas traités dans le corps de l'article lui-même. L'article proprement dit et ses annexes seront suivis d'une conclusion¹ qui reprend, en français, les principaux résultats de l'article et les resitue dans la globalité de cette thèse.

6.1.1.2 Un format adapté aux contraintes

Ce format de thèse avec article est un compromis permettant de répondre à plusieurs contraintes auxquelles a été soumise ma thèse. Il s'agit d'une part de celles issues de l'organisation sociale de la science en général et du doctorat en particulier, relayée principalement par l'école doctorale, d'autre part de celles liées à mon sujet de recherche à la fois interdisciplinaire et exploratoire.

Les canons de la science de ces dernières décennies, particulièrement ceux des sciences « naturelles »² exigent que la thèse se fasse dans une durée courte ne dépassant pas trois ans. Cette durée est issue de la volonté politique d'harmoniser l'ensemble des thèses de doctorat, quelle que soit la discipline, en un contrat doctoral de trois ans. Par ailleurs l'évaluation des scientifiques (en particulier par les institutions de recherche) se fait en fonction de leurs publications. En outre le développement récent de la scientométrie consiste à faire intervenir de manière déterminante le nombre de ces publications (indépendamment, bien souvent, de leurs qualités) dans l'évaluation institutionnelle. Cette tendance se reflète dans la contrainte, imposée par l'école doctorale du MNHN, de publier au moins un article pour pouvoir soutenir sa thèse.

Ces contraintes sociales viennent dans mon cas³ se confronter à d'autres, propres à mon sujet. Mon travail se situe au croisement de deux disciplines aux traditions différentes : la biologie dont le format standard de la thèse est de plus en plus celui de la thèse sur articles et l'ethnomusicologie où le temps long du terrain rend difficile, voire impossible, la soutenance en trois ans. Ce travail interdisciplinaire m'a aussi astreint à un important travail de bibliographie. Il était pourtant essentiel d'intégrer cette partie peu novatrice et difficilement publiable dans cette thèse afin de donner les moyens au jury d'évaluer mon travail.

1. ↑ Voir page 245.

2. ↑ Mais les canons de ces sciences, tendent, de par les faveurs économiques et politiques dont elles jouissent, à s'imposer aux sciences humaines et sociales.

3. ↑ Et dans bien d'autres...

Enfin mon sujet, très exploratoire et fondamentalement interdisciplinaire, rend difficile la publication. Il nécessite d'ajuster ensemble des cadres épistémologiques distincts afin de trouver une stabilité et une pertinence qui assure la reconnaissance à la fois par soi-même et par ses pairs. La publication ne peut être rapide qu'une fois ce cadre établi⁴. Ce sont ces contraintes (la nécessité de publier au moins un article et les difficultés qu'engendrent une telle publication) qui sont à l'origine de ce format : ayant à la fois trop peu publié pour rédiger une thèse sur articles et les contraintes temporelles rendant difficile l'écriture simultanée d'une thèse classique parfaitement cohérente et d'un article, j'ai choisi d'intégrer cet article au corps de ma thèse.

6.1.1.3 Limites

Ce choix de thèse avec article pose néanmoins quelques problèmes. Ainsi la présentation de cette analyse étant soumise aux contraintes de la publication, certaines méthodes ou enjeux pourront paraître trop vite traités. De plus l'article, autonome, est mal intégré au reste de la thèse (bien que cette introduction tente d'y remédier). Je m'en excuse et j'espère qu'elles ne nuiront pas outre mesure à la bonne compréhension de mon travail.

6.1.2 Contexte et objectifs

6.1.2.1 Contexte

Comme on l'a vu dans le chapitre précédent⁵, la possibilité de l'inférence historique est fortement limitée par le peu de connaissances que l'on a concernant la transmission et la transformation des musiques du Gabon. Les données utilisées par Le Bomin et coll. [240] pour évaluer l'importance de la transmission horizontale n'avaient jamais été systématiquement décrites et explorées. Il nous a donc paru nécessaire de revenir à une analyse plus descriptive indispensable aussi bien aux autres scientifiques pour comprendre tous les enjeux de notre travail que pour nous-même afin de comprendre plus finement nos données et ce que l'on pouvait (ou pas) en tirer. L'objectif premier de cet article est donc de décrire les données patiemment amassées par Sylvie Le Bomin sur une grande partie du territoire du Gabon. Jamais présentées en détail jusqu'à ce jour, elles offrent un panorama des musiques gabonaises et affinent nos connaissances sur les musiques du Gabon et de l'Afrique centrale.

6.1.2.2 Objectifs

Le jeu de données collectées par Sylvie Le Bomin étant trop conséquent pour être décrit intégralement, nous nous sommes concentré.e.s sur trois paramètres musicaux internes : l'organisation *métrique*, l'organisation *rythmique* et l'organisation *scalaire*. Nous décrivons donc la diversité et la distribution des différentes valeurs de ces paramètres à l'échelle du Gabon.



4. ↑ Alix Levain, communication personnelle.

5. ↑ Voir chapitre 5 pages 155 à 185.

La musique a un rôle identitaire dans les étudiées et c'est un vecteur de socialisation et de contact entre populations. Nous nous sommes donc aussi intéressé-e-s aux liens qui pouvaient exister entre les valeurs de ces paramètres musicaux et l'*ethnonyme* auquel se réfèrent les populations d'une part et la région géographique dont elles proviennent d'autre part. On sait par ailleurs qu'en Afrique centrale la musique est intimement liée à son contexte d'exécution par le biais du *répertoire*. Nous avons donc cherché à voir la stabilité des différents paramètres musicaux dans un répertoire, aussi bien au sein d'une population donnée qu'entre des populations différentes.

On comprendra mieux la place de cette analyse descriptive au sein de ma thèse à travers le dernier objectif de cet article : à partir des connaissances dont nous disposons sur ces musiques, il s'agit d'évaluer quels sont les paramètres les plus à même de nous informer sur l'histoire des musiques du Gabon et peut être, ainsi, sur celle des populations de ce pays.

Diversity and distribution of musics in Gabon

Jérémy Gardent¹^{*}, Sylvie Le Bomin¹

1 Éco-Anthropologie et Ethnobiologie, UMR 7206 CNRS–MNHN, Université Paris—Diderot, Sorbonne Universités, Muséum national d’histoire naturelle, Musée de l’Homme, Paris, France

 These authors contributed equally to this work.

* jeremy.gardent@mnhn.fr (JG)

Abstract

Almost a century of ethnomusicological studies in Gabon suggests that music of this country is singular among other Central Africa countries. Moreover, previous work suggests that musical patrimonies in Gabon could carry vertical transmission signal, implying that music could be used to infer history. Here we want to describe the diversity and distribution of Gabonese musics. We use musical systematics, quantitative descriptive tools and map plotting to analyse the geographical distribution of three internal musical parameters (metric, rhythmic, and scalar organization). We also look for global tendencies in parameter values among performance context (repertoire), ethnonymic group and geographical areas. This allows us to highlight the specificities of music diversity in Gabon, and to discuss the way music may be used to infer history. Some parameters show clear geographical organization, in particular beat subdivision which is mainly binary in the west side and ternary in the east side of the country. In contrast with what was shown for the rest of Central Africa, repertoires are rarely characterized by a single parameter value. Often shared between geographically close populations, rhythmic figures appear as the best candidate among internal parameters for use as a historical marker. It opens a new range of interesting questions for investigating the use of music in historical inference.

Introduction

Cultural diversity of Central Africa.

Central Africa is an area with important mixing and meeting of populations with varied life styles, languages, social systems and musical practices. For this reason, it is a region of special interest in studying cultural diversity and change. Central Africa is populated with many different human groups, most of them speaking Bantu or Ubanguian languages and living from farming or hunting-gathering. Cultural diversity of these populations has been studied mainly in terms of linguistics, using language diversity as a proxy of cultural diversity. These works increase our knowledge on the history and evolution of these populations.

Musics of Central Africa.

Musical patrimonies of Central Africa have been intensively studied for over fifty years (e.g. Arom [1]). Thanks to the development of new tools for musical systematics and categorization we now better understand the rules of musical systematics in different cultures. This gives a way to better understand what is specific and what is shared by various populations and hence to interrogate the potential for music to reflect cultural identity and diversity.

The great majority of musics performed in Central Africa share several important features. Musical systematics highlighted a series of common rules characteristic to the region. Musical pieces are voco-instrumental and serve specific functions in particular socio-cultural contexts. The great majority of musical pieces are measured, with strict periodicity and cycles of 6, 8, 12, 16 or 24 beats. In the northern part of Central Africa, scales are pentatonic anhemitonic, whereas in the southern part different types of scale co-exist, with an important diversity of degrees and semi-tone [2–4].

Numerous studies have been dedicated to identify endogenous musical categorization processes. It consists in revealing distinctive musical feature(s) of a set of pieces in a given population. Such a set of pieces is called *repertoire* and is linked to a specific performing context (e.g. hunting, initiation ceremony, ...) [5]. Most of the time, repertoires have distinctive endogenous names. The musical features involved are usually a musical instrument, a rhythmic figure or a metricity. Most of repertoires are

characterized by a specific rhythmic pattern performed by idiophones and/or membranophones. In most of Central Africa, repertoires tend to be specific to a population.

Ethnomusicological studies in Gabon.

The first ethnomusicological data have been collected in Gabon by Gilbert Rouget, in the Ogooué River region in 1946 [6, p.198]. These prospective data were limited to a part of the territory then known as French Equatorial Africa, and were not analysed individually. More extensive research campaigns were led by Herbert Pepper [7] and Pierre Sallée [2,8,9] between 1954 and 1966. These campaigns aimed to make a systematic inventory of musics across a wide diversity of geographic areas, cultures and languages. These studies showed both a musical continuity with other Central African countries, and a great diversity of musics across regions and ethnic groups in Gabon.

In recent years, ethnomusicological monographic studies have been carried out among specific ethnic groups of Gabon [10–12]. Furthermore, broader studies studying historically and geographically linked ethnic groups have been also published [3,4], detailing the similarity and diversity of musics across relatively broad geographic areas. Recently, a country-wide comparative study of musical patrimonies in Gabon has been published, being the first of this scope in the region [13] (see below).

Specificity of Gabonese musics.

These recent works have challenged the general consensus over the principles of musical organization and categorization in Central Africa. Endogenous musical categorization among Gabonese populations seems more complex than what was observed in neighbouring countries, e.g. in the Central African Republic [3,4,11,12,14]. In Gabon, a set of musical pieces is rarely characterized by a single unique distinctive feature. Moreover, various repertoires can have the same rhythmic figure, and different pieces of a given repertoire can have various rhythmic figures [4,14].

Comparative approach.

Comparative studies have a long-standing link to ethnomusicology, a discipline known as “comparative musicology” until the middle of the XXth century. First

ethnomusicologists pointed comparison as essential, if not the ultimate goal, of the discipline [15, pp. 72-74]. However, inter-cultural comparative approaches were discredited from the 1950s because of their symbolic association, at the time, with racist, ethnocentrist and nationalist ideologies [15, pp. 82-84]. Championed by some, denigrated by others, comparative studies have played a fairly consistent role in ethnomusicology throughout the XXth century [15, p. 86].

Aside from the ideological debates, large scale inter-cultural comparison studies have been restricted by methodological problems. First, inter-cultural comparison requires good knowledge of the musical cultures across the entire area studied. This ensures that the parameters chosen to carry the comparison are relevant to all cultural contexts. It also permits to check that the same parameters can be considered as equivalent in the cultures, ensuring that the method is sound and the comparison meaningful. Lack of detailed knowledge of the individual cultures has weakened the results of the few attempts to carry large scale comparison (e.g. [16]). Second, we have to make sure that the data to be compared have been obtained in a consistent way. Often, musical data collected over a large area suffer from inter-individual variability resulting from a diversity of collectors with distinct goals in mind when the material was collected. This limits strongly the possibility for meaningful comparison. Third, the comparison needs a robust descriptive and analytical framework applicable to all the cultures included in the comparison.

The musics included in the present study cover a broad cultural area showing a great musical consistency. Moreover almost 17 years of fieldwork throughout Gabon ensured us of the cultural relevance, in all localities visited, of our collecting method and of the parameters used [13]. Finally, longstanding studies in central Africa have permitted the emergence of a musical systematics and categorization of these musics [1, 5], which can be used as a robust framework for our analysis.

Music as a marker of history.

The potential of music parameters as robust markers for historical inference has been pointed in ethnomusicology as early as the beginning of the XXth century, e.g. by Hornbostel [17, p. 249]. Although relatively few empirical studies have been published with this goal in mind, the recent years have shown a renewed interest in these

questions (e.g. [18,19]), although studies sometimes remained mostly speculative (e.g. [20]). Concerning Central Africa, the pioneering work of Fürniss [21] in Cameroon about the origin of the circumcision ritual among Baka pygmies paved the way for studies on cultural borrowing processes among populations from different languages and traditions. Musical comparison studies among several Gabonese populations have revealed exchanges of musical repertoires between pygmies and non-pygmies populations [22]. Recently, the work of Le Bomin et al. [13] has hinted cultural processes of music transmission and evolution in Gabon populations. This study found a relatively high level of hierarchical clustering of musical data, suggesting a predominantly vertical transmission of musical characters. Furthermore, it showed that intrinsic musical features have the highest consistency levels. This led us to believe that these features may be the most promising candidate markers for inferring the history of human populations.

Goals

This study has a double objective of 1) better assessing the generality of musical systematics and categorization in Gabon; and 2) exploring how music material could be used to infer the history of population migrations and contacts. In particular, we wished 1) to show how the tools of musical systematics and categorization can allow intercultural comparison of musical data; 2) to describe the diversity and distribution of three intrinsic musical parameters (metric, rhythmic, and scalar organization) in Gabon; 3) to explore the consistency of these musical parameters inside a repertoire; 4) to compare these results with what was previously known or inferred about traditional musics in Gabon and in Central Africa; and finally, 5) to propose new directions for studying how music may change over time and how musical data could be used to infer musical history.

Material and methods

116

Sampling

117

Populations

118

The elementary units of the analysis are populations, defined as groups of people characterised by an ethnonymic (cultural) and geographic (spatial) unity. Fieldwork was carried out by SLB in Gabon between 2005 and 2009. Data were collected in 20 different populations (S2 Table, S2 Fig) representing seven of the nine provinces of Gabon (no data from the Woleu-Ntem and the Estuaire provinces were used). Below, each population name is followed by a ‘#’ and its corresponding number in S2 Table and S2 Fig. For example, the population corresponding to the Kwele ethnonym and the Mazingo village will be written ‘Kwele - Mazingo #10’. These 20 populations represent 10 distinct ethnonymic groups. In some populations, certain types of musical data could not be collected.

119

120

121

122

123

124

125

126

127

128

Characteristics of pieces studied

129

Musical pieces included in this study are cyclical musics based on repetition and variation. The music is performed by groups of people of variable size. People or groups of people make musical production of different melodico-rhythmic components. Each of these different melodico-rhythmic components is called a *part*. The rhythmic component of a part is called the *rhythmic figure*. All pieces are polyphonic and polyrhythmic. *Polyrhythm* results from the antagonistic layout of the respective distinctive marks of various rhythmic figures [23]. Instrumental parts are performed by drums or sometimes by idiophonic instruments (e.g. a stricken beam or a wood board). Pieces generally have two or three different vocal parts. Pieces studied are cyclical musics.

130

131

132

133

134

135

136

137

138

Recording method

139

In order to allow easy and standardized analysis of musical material, each piece was recorded using *continuous recording of individual parts* (see S1 Protocol and S1 Fig).

140

141

Corpus

About 250 pieces were recorded, 187 of which were transcribed and analysed here. For each population, at least three and up to 20 pieces were analysed, with a mean of 9.35. Metric organization were identified for 177 pieces, rhythmic organization for 137 pieces, and scalar organization for 168 pieces (see detailed dataset in S4 Table).

Musical parameters

Pieces were transcribed and analysed by SLB using to the tools of musical systematics [1]. All musical parameters were identified by aural analysis of recorded material. The musical parameters identified for each piece are described in S4 Table.

Metric organization

Arom defines metric as the segmentation of time in quantities of equal length [23, p. 934]. The pieces considered here have three levels of segmentation: the period, beat and minimal operational values (M.O.V.). Being cyclical and isoperiodic, the larger temporal unit is the *period* which is cyclically repeated all along the performance. Each period can be subdivided in isochronous elements, the *beats*, eventually materialized by culture upholders (e.g. by dance steps or hand clapping). Finally every single beat can be divided in isochronous elements, the *M.O.V.* Pieces are thus characterized by the number of beats per period ('B', the *periodicity*) and by the number of M.O.V. per beat ('S', the *beat subdivision type*). The beat subdivision type can be either *binary* if there are two or four M.O.V. per beat, or *ternary* if there are three M.O.V. per beat. The *metricity* of a piece is the combination of its beat subdivision type and its periodicity. We use the following format to describe the metricity of a piece: 'B/S' (e.g. metricity of a piece with a periodicity of eight beats and a binary beat subdivision is marked '8/2').

Rhythmic organization

Rhythmic organization is the grouping scheme of successive auditive events contrasting by their duration, their accents or their timbre [1, p. 202]. The *rhythmic figure* of a part is the reference rhythmic pattern underlying the realization of each period for this part. We divided each rhythmic figure in elementary units of equal duration, the *rhythmic*

cells. These have a length of two beats when beat has a binary subdivision and one beat when beat has a ternary subdivision. The list of elementary rhythmic cells brought to evidence in our corpus is given in S3 Table. Among the instrumental parts, one or two are relatively stripped-down and more repetitive. These parts are frequently identical in several pieces (e.g. those of a same repertoire). These parts are very often played by drums, and sometimes by idiophonic instruments, like a stricken beam (*obaka*) or a wood board. They have a role of rhythmic landmark for the other parts. Here, we refer to these parts as the *rhythmic reference*. We characterized the rhythmic organization of a piece by listing the rhythmic cells of each part of the rhythmic reference.

As the studied music is cyclic, there is no fixed starting point and music can start on any point of the cycle. In order to avoid irrelevant differences when comparing rhythmic figures, we use a canonical form to describe them. Thus, we always start the rhythmic figure on the rhythmic cell whose identifier is the smallest. The same applies for next rhythmic cells if several options remain. Note that we do not change the sequence order, but only change the starting point of the cyclical sequence. Thus a rhythmic figure first described by the following sequence of rhythmic cells: ‘23,4,4,10’ will be permuted to the canonical form ‘4,4,10,23’.

Scalar organization

A musical scale is a closed set of pitches lying within an octave, each pitch constituting a degree of the scale [1, p. 218]. The scale of a piece is an overall inventory of degrees of all melodico-rhythmic parts of the piece.

A scale can be represented by the set of the melodic intervals separating the successive degrees of the scale: the interval sequence. In the cultures considered, absolute pitch is not relevant. Moreover there is a tolerance margin in the realization of degrees, defining the dispersal field of a degree [24, p. 372] [25]. The realization of intervals also has a tolerance margin. Here, we represented intervals by ‘interval codes’, which are numerals defined as follow: ‘1’ represents an interval of roughly half-a-tone, ‘2’ an interval of roughly one tone, and ‘3’ an interval of roughly one tone and a half. Each scale is then represented as a sequence of such intervals codes.

In the pieces considered, there was generally no obvious hierarchy between degrees, and therefore no reason to choose a degree over another as the starting point of the

scale. A scale can be regarded as a cycle of intervals. In order to allow comparison between these cycles of intervals, we used the following convention: the first interval in the sequence was chosen so that it had the lowest interval code (e.g. '1' rather than '2'), and so on for the following intervals until there was no more ambiguity. For example, the scale II of S6 Fig can be represented by any of the following interval code sequences: 221223, 212232, 122322, 223221, 232212, 322122. According to our convention, when comparing different scales we used the following interval code sequence: 122322 (since '1' is the lowest interval code found here). We called this conventional sequence the *canonical intervals codes sequence* (CICS).

In order to allow a broader comparison, we also defined the *scales class*. A scale class has the following format: '(D-H)', where 'D' is the number of degrees in the scale and 'H' is the number of '1' it contains in its intervals codes sequence. A scale class comprises several CICS.

Here we extracted the scales from the sung parts of the pieces. The vocal parts were also transcribed. The list of occurring degrees was then extracted and sorted in increasing order. Finally, the number of degrees, number of '1' intervals, scale class and CICS of the scale were computed using the 'music21' python module [26] (see below). The full dataset in S4 Table describes the scale of each piece with both solfegic notation and CICS.

Descriptive vocabulary

In order to describe the distribution of parameters values inside the groups of interest (e.g. populations or repertoires), and following the work of Arom et al. [5] we defined three terms:

- specific: A parameter value is *specific* of a group if no piece out of this group have this value. For example, in Bongo from Lengoye #7, the CICS '12522' is specific of *Mavasa* repertoire: it is find in no other repertoire.
- invariant: A parameter value is an *invariant* of a group if all pieces of this group have this value. For example, in Bakoya from Ekata #12, the '16/3' metricity is an invariant of *Ibwema* repertoire: all pieces of this repertoire have this metricity.
- distinctive: A parameter value is *distinctive* of a group if it is both specific and an

invariant of this group, *i.e.* if all pieces of this group and only them have this parameter value. For example, in Ngom from Moualo #5, the ‘8/2’ metricity is distinctive of *Makouma* repertoire.

Formatting and visualization

Barplot and map plotting

For each musical parameter studied (metric, rhythmic and scalar organization), the values taken by the parameter were represented using two complementary approaches. First, the occurrence of the various values of the parameters were represented on barplots. This gives an overview of the diversity of each parameter, and permits to assess their relative abundance. We also represented percentages of all pieces sharing the corresponding value. Because the number of pieces analysed in each population varied, ranging from three (Barimbe from Mandilou #9) to 20 (Ngom from Mimongo #2), we also plotted the percentage of all populations having this value in any of the pieces sampled from it, to avoid bias linked to variation in sampling effort.

Second, for each population, the relative abundance of the values in this population were plotted on an individual barplot. All population barplots were then plotted on a map of Gabon. This permitted to assess the diversity of the values in the different populations. It also gave an overview of the geographical repartition of the values of the parameters.

Population-repertoire analysis

Each parameter was also analysed by taking into account the repertoires of the pieces in the following manner. We defined a *population-repertoire pair* as the set of all pieces in the same repertoire in a given population. For example the repertoire *Bubeyu* in the Awandji from Doumé #13 is a population-repertoire pair from which we analyzed two different pieces. Our dataset contained 104 population-repertoire pairs. We analyzed up to four pieces in a population-repertoire pair, but most of the time only one (38%) or two (47%) pieces. For each of these pairs, we checked if all values of the parameters were identical or not, that is if the parameter is *invariant* in this repertoire (see above). When we analysed a given parameter in a single piece of a population-repertoire pair,

we can not state on invariance of this parameter for this pair. We also look at the number of other repertoires in which a given parameter value is found. If a parameter value is found in only one repertoire of this population, this value is said *specific* of this repertoire (see above). When we analysed only one repertoire in a given population, we can not state on the specificity of this parameter for this population-repertoire pair.

Software and geographic data

Data were formatted and analysed with the R software (version 3.3.3) [27] and standard packages. Populations locations were recorded by SLB with GPS device. Other geographical data such as country boundaries, cities positions and routes come from the national cartography institute of Gabon, except for the course of the Ogooué river which was downloaded from OpenStreetMap using the ‘osmar’ R package [28]. Scales were analysed thanks to the ‘music21’ python module (version 2.2.1) [26] and the ‘rPython’ R package [29].

Results

Metric organization

We retrieved metric information on 177 pieces from 20 populations. All were measured with a regular beat. Metric organization were available in at least two pieces for 61 population-repertoire pairs. In 27 of those (44%), all pieces had the same periodicity, and in 57 of those (93%) all pieces had the same beat subdivision type.

Periodicity.

Periods of the pieces analysed contained between four and 24 beats. The number of beats per period was diverse among the pieces studied. We noticed an overall predominance of 8-, 12- and 16-beats cycles. This was true in all aspects, considering the number of populations, the number of repertoires or the number of pieces where they were found. Shorter periodicities (4 and 6 beats) were far less common, and found in 16 pieces from nine distinct repertoires. Likewise, longer periodicities (18, 20, or 24 beats) were found in seven pieces from six different repertoires and four different populations.

We also observed a geographical organization of the number of beats per period. There was an overabundant amount of cycles with high number of beats in the north-east region, where cycles of 18, 20 and 24 beats were found. In contrast, there was an overabundance of short cycles in the center and the south-east of the country. Most of 4- and 6-beats cycles were found near the cities of Mbigou (Bongo from Lengoye #7 and Ngom from Bolapessa #8), Pana (Bongo from Moghonbofoala #18 and Bongo from Iwatsi #19) and Franceville (Ngom from Ndjokaye #16 and Bakanighi from Mopia #17).

Pulsation subdivision.

Overall, binary and ternary beat subdivisions were found roughly in the same proportions. However, this repartition was not homogeneous among populations and most of them used mostly one type of subdivision (Fig 1). In 10 out of the 20 populations, all pieces analysed had the same beat division type (binary for four populations and ternary for six). In the 10 remaining populations, two showed a clear predominance of ternary beat subdivision (Kwele from Mazingo #10 and Bakoya from Imbong #11). Binary/ternary proportion was roughly equivalent in only eight populations (Galoa from Nombedouma #3, Nkomi from Amendje #6, Bongo from Lengoye #7, Barimbe from Mandilou #9, Awandji from Doumé #13, Awandji from Lipaka #14, Bongo from Moghonbofoala #18 and Bongo from Iwatsi #19).

We also observed a clear regional organization of beat subdivision (Fig 1). Most of populations with predominance of binary beat subdivision were in the west side of the country, whereas most of populations with predominance of ternary beat subdivision were in the east side. The Ogooué River has been proposed as a geographic barrier delimiting areas of different musical features [13]. All populations with a predominance of ternary beat subdivision were situated north of the Ogooué. Almost all populations with a predominance of binary beat subdivision or with roughly equal binary and ternary proportions were situated south of the Ogooué. The only exception is the Orungu from Port-Gentil #4, located south of the Ogooué (close to the mouth of the river) with ternary beat subdivision. The Ogooué River could indeed play a barrier role. This would need to be further checked since most of the samples came from the south of the Ogooué.

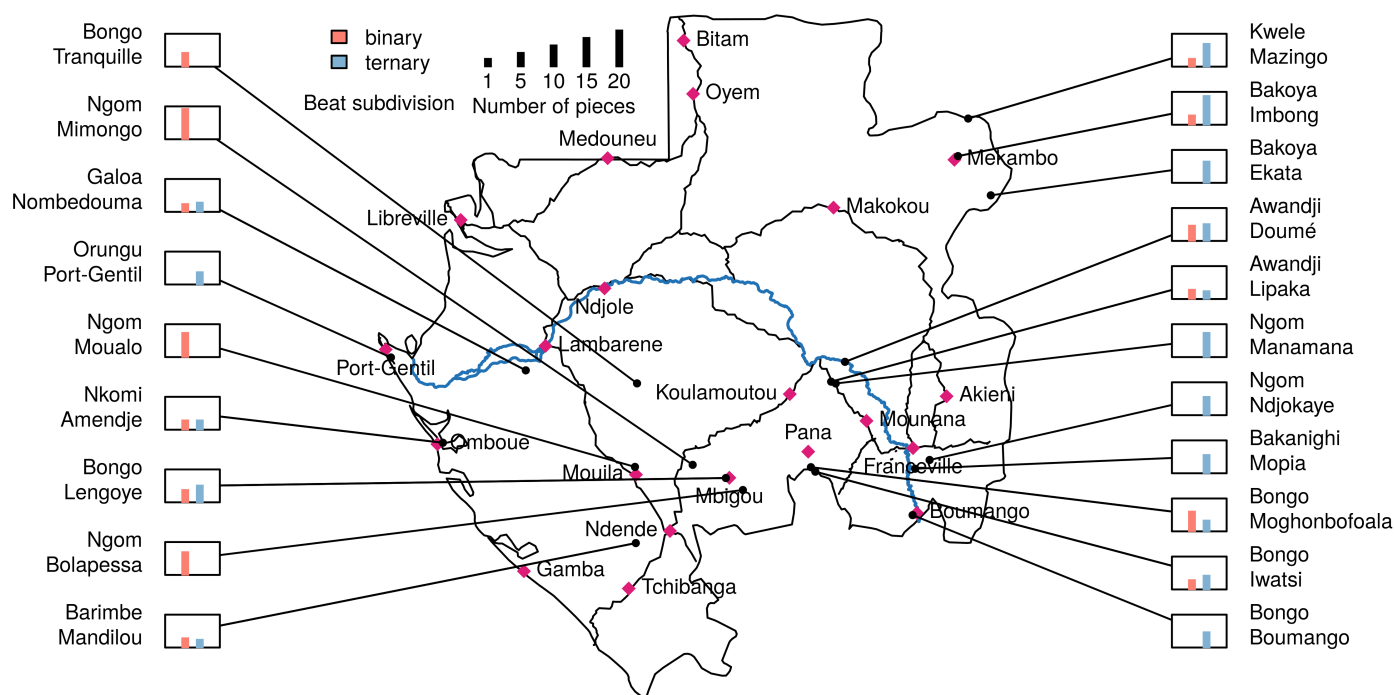


Fig 1. Geographical distribution of beat subdivision. For each population, the number of pieces with binary and ternary subdivision is indicated by colored bars, respectively in red and blue. Reading example: in Kwele from Mazingo #10, beat subdivision is binary in one piece and ternary in 11 pieces.

Distinctiveness

318

There was no distinctive feature in periodicity, beat subdivision or whole metricity across populations. Most metricities were widespread throughout the populations and repertoires. Some periodicities (6, 20 and 24 beats) were much less common and found in only one piece, and therefore could be considered as particular cases rather than population specificities.

319

320

321

322

323

Metricity was generally not distinctive of a repertoire in a given population, with only six exceptions over 60 population-repertoire pairs (Table 1). This low distinctiveness is mostly due to low diversity and consequently low specificity of values.

324

325

326

Among the 14 different metricity values we found, 10 were shared by various populations, with the two most widespread ('8/3' and '8/2') being shared by 15 and 12 different populations, respectively. The '4/2' metricity was found only in three populations in the south of the country, all located within a relatively small geographic area: the Bongo from Lengoye #7, the Bongo from Moghonbofoala #18 and the Ngom

327

328

329

330

331

Table 1. Distinctive parameters values in a repertoire of a population. Fields were left empty when no distinctive value was found. Fields were set to ‘NA’ (‘not applicable’) when there was not enough data to identify distinctive values. It was the case if there was less than two pieces in a population-repertoire pair (as one cannot state on invariability), or when there was data from a single repertoire in this population (as one cannot state on specificity). ‘RF’ stands for ‘Rhythmic Figure’. Population-repertoire pairs for which no distinctive parameter value was found are not shown.

Population	Repertoire	Metricity	CICS	RF (drum 1)	RF (drum 2)	Rhythmic cells
Ngom - Bolapessa #8	<i>Lisimbu</i>			30,31,46	NA	
Ngom - Manamana #15	<i>Bubeyu</i>		NA			7,8
Ngom - Manamana #15	<i>Lisimbu</i>			25,25,25,25		
Ngom - Mimongo #2	<i>Kunde</i>			30,46	NA	
Ngom - Mimongo #2	<i>Lisimbu</i>			30,31	NA	
Ngom - Mimongo #2	<i>Makouma</i>	16/2		NA	NA	NA
Ngom - Moualo #5	<i>Makouma</i>	8/2	12522	NA	NA	NA
Bongo - Boumango #20	<i>Lisimbu</i>	8/3		2,12,4,3	2,2,27,3	12,27
Bongo - Moghonbofoala #18	<i>Laka</i>			41,41	41,42	41,42
Bongo - Moghonbofoala #18	<i>Ngwata</i>			34,36		31,34,36
Bongo - Moghonbofoala #18	<i>Nzobi</i>			3,19,20	3,3,3	19,20
Bakanighi - Mopia #17	<i>Mopfougou</i>				NA	25
Bakanighi - Mopia #17	<i>Nzobi</i>			3,3,3	6,19,26	26
Bakoya - Imbong #11	<i>Boumba</i>		1123212	3,4	NA	
Bakoya - Imbong #11	<i>Bwiti koji</i>					21
Bakoya - Ekata #12	<i>Ngodja</i>	8/3			NA	
Bakoya - Ekata #12	<i>Nzokou</i>			2,10,3,4		
Galoa - Nombedouma #3	<i>Mimbwiri</i>	NA	1221222	NA	NA	NA
Ngom - Ndjokaye #16	<i>Lisimbu</i>					4
Nkomi - Amendje #6	<i>Mabandji</i>	8/3		3,10	NA	3,10
Nkomi - Amendje #6	<i>Mimbwiri</i>	8/2	122322		NA	
Orungu - Port-Gentil #4	<i>Mabandji</i>			3,10	NA	3,10

from Bolapessa #8. For all other metricities, the corresponding populations were
widespread over the country. Moreover, no metricity was specific of a particular
ethnonym.

Rhythmic organization

Overview

We transcribed rhythmic data from 137 pieces collected across 19 populations. For most
pieces (70%), two rhythmic parts were transcribed and analysed. Only one rhythmic
part was transcribed for 40 pieces, and three rhythmic parts were transcribed for one
piece. In total, 160 different rhythmic figures were found, most of them (76%) being
found in only one piece. Rhythmic figures were composed of one one (9%), two (30%),
three (39%) or four (21%) different rhythmic cells, very rarely above four (1%).

We observed a great diversity in rhythmic cell occurrences. Thus 24 (40%) of the 60

different rhythmic cells were found in only one population (from which 17 were found in a single piece), whereas only 23% were found in more than 5 populations.

One rhythmic cell had a special importance: the cell ‘3’, constituted by three quavers, used to describe pieces with ternary beat subdivision. When we took only pieces with ternary beat subdivision, it was found in 80% of pieces, 86% of repertoires (19 over 22) and all populations (15). In pieces where beat subdivision was binary, the most frequent rhythmic cell was the ‘31’ (constituted by 2 crotchets), found in 62% of the pieces, 71% of the repertoires (17 over 24), and 82% of the populations (nine over 11).

Distinctiveness

In total, 39 rhythmic figures were found in at least two different pieces, from which 19 rhythmic figures were found in at least two populations (Fig 2 and Table 2). We looked at the mean distances between all populations sharing a given rhythmic figure. Among the 19 rhythmic figures found in at least two populations, 13 (68%) were found in populations whose average distance was lower than 150km. Cases where this is not true (e.g. ‘3’ or ‘3,10’) often correspond to rhythmic figures composed of few widespread rhythmic cells. Moreover, these populations often had the same ethnonym. This was also true for first (60%) and second drums (75%) taken separately.

We observed a geographical proximity of the populations where they were found for roughly half of the 29 rhythmic cells found in at least three populations. For example:

- all populations having the ‘22’ rhythmic cell are found in south-east part of the country;
- almost all populations having the ‘31’ rhythmic cell are in the west part of the country;
- ‘47’ rhythmic cell is specific of south of the country.

Rhythm in a repertoire of a given population

Rhythmic figures were available in at least two pieces for 42 population-repertoire pairs for the first drum. In 14 of those (33%) rhythmic figures were exactly the same in all pieces. Concerning the second drum, rhythmic figures were identical in all pieces in six (21%) of the 29 population-repertoire pairs available. When the same rhythmic figure

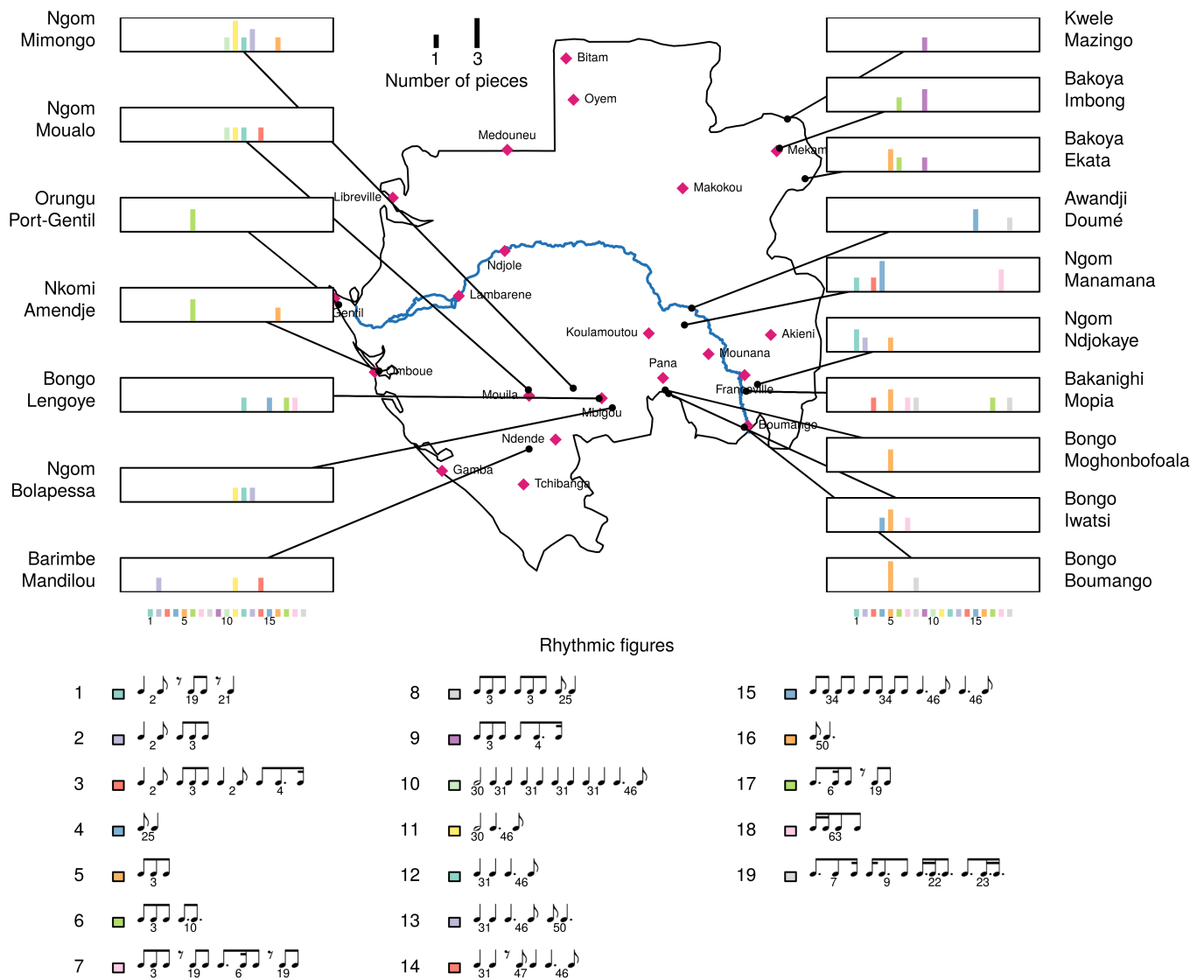


Fig 2. Geographical distribution of rhythmic figures. For each population, the number of pieces in which a given rhythmic figure is used is indicated by a color bar. Only rhythmic figures found in at least two populations are shown. Reading example: in Barimbe from Mandilou #9, rhythmic figure ‘2,3,2,3’ is found in one piece and ‘30,46,30,46’ in one piece.

was found in different pieces from a given population, the pieces came from the same repertoire in roughly two out of three cases. For rhythmic cells found in different pieces from a given population, the repertoire was identical in only 23% of cases. A population frequently used the same rhythmic cell, and sometimes the same rhythmic figure, in different repertoires.

In total, 13 population-repertoire pairs showed a distinctive rhythmic figure for the

Table 2. Rhythmic figures shared by various populations. ‘Distances’ column gives the distances as the crow flies (in kilometers) between all the populations sharing a same rhythmic figure. Reading example: for the rhythmic figure ‘3,4’: Bakoya from Imbong #11 are 52 km from Bakoya from Ekata #12; Kwele from Mazingo #10 are 81 km from Bakoya from Ekata #12 and 39 km from Bakoya from Imbong #11.

Rhythmic figure	Population	Repertoire	Distances
2,19,21	Ngom - Ndjokaye #16	<i>Mungala, Nzobi</i>	123
	Ngom - Manamana #15	<i>Nzobi</i>	
2,3	Ngom - Ndjokaye #16	<i>Lisimbu</i>	310
	Barimbe - Mandilou #9	<i>Mimbwiri</i>	
2,3,2,4	Bakanighi - Mopia #17	<i>Mungala</i>	118
	Ngom - Manamana #15	<i>Mungala</i>	
25	Bongo - Iwatsi #19	<i>Nzobi</i>	91
	Ngom - Manamana #15	<i>Lisimbu</i>	
3	Bongo - Moghonbofoala #18	<i>Nzobi</i>	6 121, 117 106, 102, 17 114, 108, 58, 47 330, 331, 274, 286, 332
	Bongo - Iwatsi #19	<i>Mungala, Nzobi</i>	
	Ngom - Ndjokaye #16	<i>Nzobi</i>	
	Bakanighi - Mopia #17	<i>Nzobi</i>	
	Bongo - Boumango #20	<i>Mungala, Wuya</i>	
	Bakoya - Ekata #12	<i>Mungala</i>	
3,10	Bakoya - Ekata #12	<i>Lisimbu</i>	52 611, 598 632, 612, 101
	Bakoya - Imbong #11	<i>Ibwema</i>	
	Nkomi - Amendje #6	<i>Mabandji</i>	
	Orungu - Port-Gentil #4	<i>Mabandji</i>	
3,19,6,19	Bongo - Iwatsi #19	<i>Lisimbu</i>	102
	Bakanighi - Mopia #17	<i>Mungala</i>	
3,3,25	Bakanighi - Mopia #17	<i>Mopfougou</i>	47
	Bongo - Boumango #20	<i>Wuya</i>	
3,4	Bakoya - Ekata #12	<i>Ibwema</i>	52 81, 39
	Bakoya - Imbong #11	<i>Boumba</i>	
	Kwele - Mazingo #10	<i>Magnele</i>	
30,31,31,31,31,46	Ngom - Moualo #5	<i>Lisimbu</i>	59
	Ngom - Mimongo #2	<i>Mundukwe</i>	
30,46	Barimbe - Mandilou #9	<i>Ologo</i>	121 77, 112 98, 57, 59
	Ngom - Bolapessa #8	<i>Makouma</i>	
	Ngom - Moualo #5	<i>Makouma</i>	
	Ngom - Mimongo #2	<i>Kunde, Makouma</i>	
31,46	Bongo - Lengoye #7	<i>Mujaja</i>	21 93, 112 36, 57, 59
	Ngom - Bolapessa #8	<i>Makouma</i>	
	Ngom - Moualo #5	<i>Lisimbu</i>	
	Ngom - Mimongo #2	<i>Kunde</i>	
31,46,50	Ngom - Bolapessa #8	<i>Lisimbu</i>	57
	Ngom - Mimongo #2	<i>Minonze, Mundukwe</i>	
31,47,46	Barimbe - Mandilou #9	<i>Mbumba</i>	77
	Ngom - Moualo #5	<i>Minonze</i>	
34,34,46,46	Awandji - Doumé #13	<i>Doka, Mbadi</i>	169
	Bongo - Lengoye #7	<i>Mujaja</i>	
50	Ngom - Mimongo #2	<i>Mungala</i>	255
	Nkomi - Amendje #6	<i>Mimbwiri</i>	
6,19	Bakanighi - Mopia #17	<i>Lisimbu</i>	193
	Bongo - Lengoye #7	<i>Mavasa</i>	
63	Ngom - Manamana #15	<i>Nzobi</i>	147
	Bongo - Lengoye #7	<i>Mwiri</i>	
7,9,22,23	Bakanighi - Mopia #17	<i>Lisimbu</i>	129
	Awandji - Doumé #13	<i>Bubeyu</i>	

first drum (Table 1). Among these, four also showed a distinctive rhythmic figure for the second drum. In seven of these 13 pairs, one or several rhythmic cells were distinctive, whereas in the remaining six pairs no individual cell was distinctive. In four pairs, no rhythmic figure was distinctive but one or two rhythmic cell(s) were.

Overall, 12 rhythmic figures (31%) were found in various pieces of the same repertoire and the same population, without being present in other repertoires or populations. All in all, this suggests that repertoires in a given population are better discriminated by rhythmic figures than by rhythmic cells.

Scalar organization

Overview

We determined the musical scales from 168 pieces collected across 19 populations. Scales had at least three, and could have up to nine, but the majority (86%) had five, six or seven degrees. Over the 168 scales extracted, we found 51 different cyclical interval codes sequences (CICS) belonging to 15 different scale classes.

In spite of a wide overall diversity, there was a huge difference in abundance between scales (Fig 3). Most of the CICS (94%) were found four times or less, whereas three sequences were found more than 10 times, representing 54% of the scales we found. These CICS were ‘1221222’ (heptatonic hemitonic), ‘122322’ and ‘123222’ (both hexatonic hemitonic). They are shown in S6 Fig. These CICS were found in respectively 17 (89%), 12 (63%) and 11 (58%) populations. They were also the most frequent scales in respectively 12 (63%), six (32%) and three (16%) populations. Most of the CICS were found only in a few pieces, whereas a few CICS were found several times. This pattern was also found at the population level, with most of CICS found in three or less populations, whereas the same three CICS were found in most of populations. Moreover, the most frequent CICS of a population was always one of these three CICS.

The diatonic scale (corresponding to the ‘1221222’ CICS) was of special importance and was found in 32% of all pieces and 58% of repertoires. It was found in all populations but two: the Bongo from Tranquille #1, and the Orungu from Port-Gentil #4.

There was also a great geographical diversity in scale distribution among

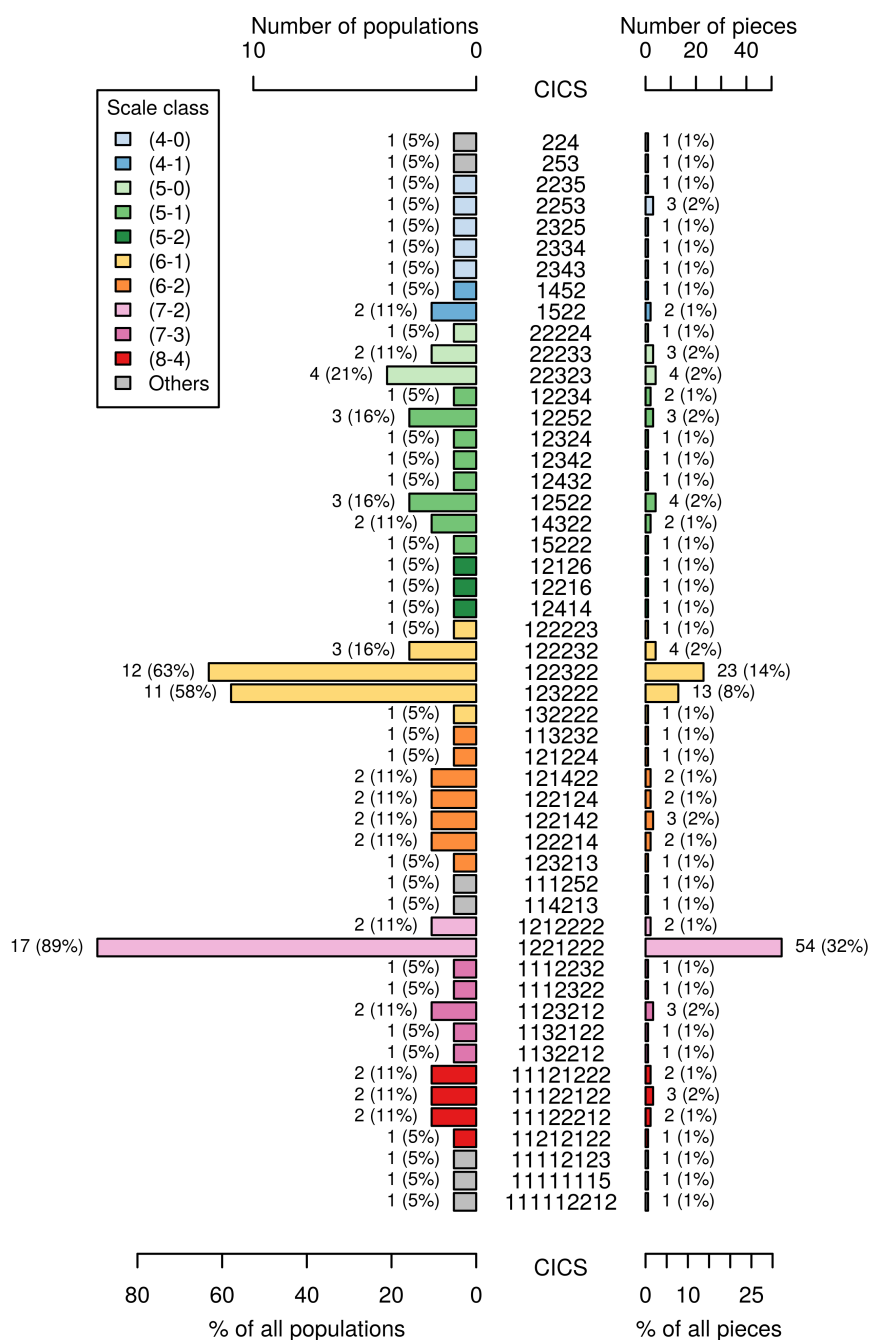


Fig 3. Repartition of the pieces analyzed according to their scalar organization. Scales found less than three times are gathered in the ‘others’ category. Reading example: the CICS ‘122322’, belonging to the scale class ‘(6-1)’, was found in 23 pieces (14% of all pieces) from 12 populations (63% of all populations).

populations. Some populations seemed to have a relative homogeneity in their scales, as
in the case of the Awandji from Doumé #13, where only three CICS, belonging to two
scale classes were found in the 11 pieces transcribed. Conversely, we found a lot of
different scales in some populations, as in the case of the Bakoya from Ekata #12,
where seven different CICS were found (belonging to six different scale classes) in the
nine pieces transcribed.

Distinctiveness.

In total, 32 out of 51 CICS (63%) were found in only one population. From these, all
were found in only one piece and could be considered as special cases, except the ‘2253’
scale found in three pieces of three different repertoires among the Ngom from
Mimongo #2. Among the 19 other scales which were shared by at least two populations,
three were widespread (found in more than 10 populations). Among the 16 CICS which
were found in two, three or four populations, five were found in geographically close
populations (less than 150 km apart). The remaining 11 CICS were found in
populations who shared neither ethnonym nor geographical proximity.

Scales were available in at least two pieces for 61 population-repertoire pairs. In 22
of those (36%), scales had the same number of half-tones in all pieces, and in 24 of those
(39%), scales had the same number of degrees in all pieces. In 13 pairs (21%), all pieces
had the same CICS, but four pairs showed a distinctive CICS value (Table 1).

Diversity of musical parameters in repertoires shared by various populations

Overview

In total, 50% of the 48 different repertoires found were collected in more than one
population, out of which 14 were collected in only two populations. three repertoires
were found in relatively wide area: the *Lisimbu*, the *Mungala* and the *Nzobi*, which were
found in 12, 10 and six different populations, respectively.

The number of different values for a parameter found in a given repertoire tended to
be lower than expected given our full sample. This was true even considering that
pieces of the same repertoire in a given population tended to have similar parameter

values. Low diversity of parameters values inside repertoires was significant for CICS, metricity (periodicity, beat subdivision, and both together) and rhythmic figure of the first drum. As discussed below, most of this low diversity is due to data from *Lisimbu* repertoire. If we exclude data from the *Lisimbu* repertoire, low intra-repertoire diversity remains significant only for metricity parameters.

No parameter value was distinctive of a repertoire. Pulsation subdivision was invariable in 13 repertoires over the 22 repertoires found in at least 2 populations (59%). This was significant in 3 repertoires (*Ibwema*, *Makouma* and *Nzobi*) found respectively in 3, 3 and 6 populations (see S5 Protocol and S5 Table for details concerning statistical significance testing). In some repertoires, like the *Lisimbu*, *Mungala* or *Mimbwiri*, beat subdivision showed a clear geographical organization, with binary subdivision on the west part of the country and ternary subdivision on the east part, in agreement with findings at the level of the entire musical patrimonies. Whole metricity (periodicity and beat subdivision) showed great variability in a given repertoire across populations.

CICS were not significantly stables excepted in *Lisimbu* and *Bubeyu* repertoires (S5 Table). Although rhythmic figures were usually conserved in a given repertoire of a population (33% of population-repertoire pairs had the same rhythmic figure for the first drum in all pieces), few rhythmic figures were shared by various populations. Over the 17 repertoires shared by at least two and up to 11 populations (for which we have rhythmic data), only five rhythmic figure were found in the same repertoires in several populations (Table 3).

The special case of *Lisimbu* repertoire.

The *Lisimbu* repertoire was the most shared repertoire in our sample, since it was found in 12 populations. It showed significant low diversity among populations for scalar (CICS, scale class and number of degrees) and metricity (though not for beat subdivision) parameters, but not for rhythmic parameters (S5 Table). Over the 21 pieces for which we have scalar information, 13 had the same CICS ('1221222'). Over the 23 pieces for which we have metric information, 12 had the same periodicity and beat subdivision type (8/3). Metricity of pieces inside *Lisimbu* repertoire showed a clear geographical organization:

Table 3. Rhythmic figures found in the same repertoire from various musical patrimonies.

Rhythmic figure	Repertoire	Populations
3	<i>Mungala</i>	Bongo - Boumango #20 Bongo - Iwatsi #19 Bakoya - Ekata #12
3	<i>Nzobi</i>	Bongo - Iwatsi #19 Bongo - Moghonbofoala #18 Bakanighi - Mopia #17 Ngom - Ndjokaye #16
2,19,21	<i>Nzobi</i>	Ngom - Manamana #15 Ngom - Ndjokaye #16
2,3,2,4	<i>Mungala</i>	Ngom - Manamana #15 Bakanighi - Mopia #17
30,46	<i>Makouma</i>	Ngom - Bolapessa #8 Ngom - Mimongo #2 Ngom - Moualo #5
3,10	<i>Mabandji</i>	Nkomi - Amendje #6 Orungu - Port-Gentil #4

- binary subdivision was distinctive of the west part of the country; 468
- ternary subdivision was distinctive of the east part; 469
- 12/3 metricity was distinctive of three close Ngom population from the center of Gabon; 470 471
- 8/3 metricity tended to be found only in the north-east part of Gabon, though other metricities could also be found for *Lisimbu* pieces in this area; 472 473

It is noteworthy that no rhythmic figure of the *Lisimbu* repertoire was shared by several populations. 474 475

Other repertoires 476

Other repertoires showed less variability in internal parameters. In the *Mungala* 477 repertoire, the CICS was '1221222' in seven over 16 pieces and the beat subdivision was 478 mostly ternary. However the number of different values for these parameters was not 479 significantly low (S5 Table). Concerning the *Nzobi* repertoire, the beat subdivision was 480 significantly invariant (ternary) and the rhythmic figures were significantly stable with 481 two rhythmic figures found in respectively two and four populations (cf. supra). In 482 other repertoires, no parameter value other than beat subdivision show significant low 483

diversity, except for CICS of the *Bubeyu* which is '1221222' in the two populations where this repertoire was found.

Discussion

Diversity and distribution of musics in Gabon

Are musical parameters distinctive of populations or geographical areas ?

We previously thought that almost all internal musical parameters values could be found in any population, and that only the presence of a given value in a given repertoire was distinctive of populations [13, p. 4]. Here, we showed that few values were found in most of populations, like the CICS '1221222' (17 populations) or the height-beats periodicity (20 populations). Most of values were found in only a few populations, the median of the number of populations in which a given value of parameter was found being lower or equal to five populations (for each parameter except beat subdivision).

Moreover, no musical parameter value was distinctive for a given population. This is due to great intra-population diversity, with almost no parameter value being invariant in a given population. The only exception was the beat subdivision type which was invariant in 10 populations, but it was never specific to a population. In many cases, some values were specific to a population. However, most of the time it corresponded to special cases. Thus, 32 CICS were specific to one population, but only two of those was found in more than one piece. Identically, no metricity specific to a population was found in more than one piece.

Some values of metric or rhythmic organization were specific of geographical areas. For example, long periods tended to be found only in north-east populations (with one exception in the south-east: Ngom from Ndjokaye #16). Similarly, beat subdivision showed a clear east-west (ternary-binary) organization. A great number of rhythmic cells showed geographical consistency in their distribution. Likewise, populations sharing rhythmic figures were almost always geographically close. In contrast, scales distribution did not show any geographic pattern.

Are musical scales invariants in populations or geographical areas ?

511

We found that the ‘1221222’ heptatonic scale was largely dominant in Gabon. This is noteworthy, since for the majority of Central African musics scales are pentatonic [30, p. 21] [31].

512

513

514

Moreover, our results contrast with knowledge from other regions of central Africa, where musical scale is often invariable in a given population but also across wide geocultural areas with different linguistic groups [4, p. 41]. In Gabon, we observed a wide diversity of musical scales that applied to all levels of analysis: within repertoires, within populations, within ethnonymic groups, and within geographic areas. Le Bomin and Bikoma [4, p. 41] made a similar observation among the Myènè populations (here represented by Galoa, Nkomi and Orungu people). Our study allows extending this observation to the entire country of Gabon. High intra-population CICS diversity seems to apply to all populations. Only the Awandji from Doumé #13 showed lower diversity, with only three different musical scales in 11 pieces collected.

515

516

517

518

519

520

521

522

523

524

Are musical scales mostly hexatonic ?

525

Early study suggested that most Gabonese musics use hexatonic scales [2, p. 198] [4, p. 41]. Further studies suggested that scales were overwhelmingly hemitonic hexatonic in the center and north of Gabon [3, p. 23]. Our results contrast these previous claims, and show that only 56 pieces over the 168 analyzed (33%) have a hexatonic scale, which is even lower than the percentage of heptatonic scales (38%). Moreover, the eight populations which show a majority of hemitonic hexatonic scales are not concentrated in the center and northern regions. In fact, many populations of these areas don not have a majority of hemitonic hexatonic scales (e.g. the Kwele from Mazingo #10, the Ngom from Ndjokaye #16), whereas some populations in other areas do (e.g. the Bongo from Tranquille #1, the Nkomi from Amendje #6)

526

527

528

529

530

531

532

533

534

535

How distinctive of repertoire and population are rhythmic figures ?

536

Previous studies in central Africa show the central role of rhythmic figures in discriminating both different populations and different repertoires inside the musical patrimony of a given population. According to these studies, a scheme is largely

537

538

539

dominant: in a given population, all pieces of a same repertoire have the same rhythmic figure, which allows identifying the population [4, p. 44]. Here, we show that the invariance of rhythmic figures among a repertoire is less important than previously thought as far as the musics of Gabon are concerned. The rhythmic figure of the first drum was invariant in only 33% of population-repertoire pairs, and this dropped to 21% for the second drum. This could be explained by previous observations by Le Bomin and Bikoma [4, p. 44]: performers' testimonies suggest that the rhythmic figures of various repertoires could proceed from a same basic rhythmic figure. This is consistent with the flexibility of rhythmic figures in Gabonese musical patrimonies, as in the case of Myènè and Teke musics [3, 4] and in populations from the center of the country [10]. This contrasts with what was observed in Central African Republic [32–34] and Cameroon [35], where rhythmic figures are strong markers of cultural identity, both at population and repertoire levels. In spite of this low invariance level, rhythmic figure was the most distinctive parameter of a repertoire in a population.

It has also been said that a rhythmic figure was very uncommonly shared between several populations [4, p. 44]. Our data suggest that it may be more frequent than previously thought. Among 39 rhythmic figures found in at least two pieces, 19 were found in at least two populations. Interestingly, most of rhythmic figure shared by several populations (13/19) were found in populations relatively close geographically (less than 150 kilometers apart; Table 2). This could be compared to the median value of the distance between two populations for all 19 populations for which we have rhythmic data, which was 252 kilometers. In only five cases did shared rhythmic figures belong to the same repertoire (Table 3).

Diversity of parameters inside repertoires of a given population

We have sufficient data to search for distinctive parameters values in a given repertoire of a population for 61 population-repertoire pairs. In 62% of pairs, no parameter value was both specific of this repertoire and invariant in all pieces of this repertoire. Repertoires were scarcely distinct from each other based on a single parameter value. Rhythmic organization was the parameter which was the most distinctive of a repertoire, especially rhythmic figure of the first drum. In contrast, metric and scalar

organization had low distinctive potential. This observation has different explanations 570
for metric and for scalar organization. Indeed, metricity tends to be invariant inside a 571
repertoire but it is poorly specific of a given repertoire. In contrast, musical scales tend 572
to be specific of a repertoire but are rarely invariant within a given repertoire. Both 573
cases result in low distinctive potential. 574

These results contrast with previous studies concerning musical categorization in 575
some central Africa countries which show that a parameter value is usually distinctive of 576
the musical patrimony of a given population (e.g. pluriarc is found only in Teke 577
populations) or of some of its repertoires. However, they also confirm the predominant 578
role of rhythmic organization compared to other musical parameters in identifying a 579
repertoire in a population, as was already observed in Gabon [4, p. 44]. 580

In most cases, we only analyzed two musical pieces of each repertoire in each 581
population. Thus, the distinctive features we found here may need to be checked by 582
further fieldwork in the corresponding populations. Our observation of unusually low 583
discriminative potential of intrinsic parameters does not seem to vary much across 584
different regions or different populations in Gabon. 585

Have more widespread repertoires less diversity of musical 586 parameters ? 587

Contrary to what has been observed in other parts of central Africa, repertoires can be 588
shared by various populations with different ethnonyms, speaking different languages, 589
and separated by large geographic distances. The diversity of intrinsic musical 590
parameters values of the pieces of these repertoires tends to be lower than what would 591
be expected by chance only. 592

Interestingly, the most widespread repertoire (*Lisimbu*) is also the most conservative 593
regarding internal musical parameters, particularly scalar and metric organization. 594
Nzobi repertoire, which is also widespread, showed low variability in rhythmic figures 595
and beat subdivision. Conversely, repertoires shared by less populations seemed to show 596
more variability of internal musical parameters. 597

Similar observations have been made in linguistics by Pagel et al. [36], who proposed 598
that the more often a word is used, the more slowly it evolves and the lower variability 599

it shows. Concerning our musical data, we thought of two explanations for this observation. First, it could be due to the necessity of maintaining identity and recognition of the repertoire over various populations. Secondly, we know that in Gabon different populations could perform together shared repertoires. This could constrain musical evolution so that populations could continue to perform together and consequently decrease variability of a repertoire among populations.

This result is mitigated by some facts, however. First the *Mungala* repertoire, which is the second most widespread repertoire in Gabon, showed relatively high internal musical parameters diversity compared to *Lisimbu* and *Nzobi*. Furthermore, we did not find significantly low levels of variation in repertoires found in few populations (S5 Table). This could be due to limited statistical power, i.e. not enough pieces in these repertoires to reach statistical significance. Our findings nevertheless open interesting perspectives for future research aiming to better understand the dynamics and mechanisms of musical change.

Which musical parameters are linked to other markers of ancestry ?

Geographical proximity and common ethnonym as clues of historical relatedness

Arom [37] and Furniss [35] have shown that metric organization could be distinctive of a geographical area; for example most of musical pieces of Aka pygmies from Central African Republic have a distinctive eight beats with a ternary subdivision. As shown by Le Bomin [33], the metric organization could also be distinctive of repertoires or social categories within a population. Here we show that some repertoires widespread across Gabon share musical characteristics. The amount of shared parameters values increases with geographical proximity and when population have the same ethnonym. This suggests that these repertoires not only have the same name but also share related musical material. This similarity could not be explained by chance alone, and must be due to some amount of historical continuity between concerned repertoires. Two types of mechanisms (and likely both of them) can explain these similarities: 1) community of descent of the populations, which transmit the repertoires in a sufficiently conservative

way; 2) cultural exchange between populations, one of the population borrowing the repertoire practice from another, and retaining at least some musical characteristics. In both cases, this could be interpreted in terms of community of descent from the point of view of the repertoire itself.

This confirm findings by Le Bomin et al. [13] that showed a strong vertical transmission signal among musical data: both results suggest that musical parameters can be used as a marker of history. They also show that some parameters have a strong signal indicating vertical transmission, especially internal parameters. Here we want to discuss the relative ability of different parameters to reveal history. In congruence-based historical inference methods (like phylogenetic inference), two criteria must be met for a set of parameters to be relevant: 1) each individual parameter must not be prone to convergence (historical signal); 2) most of markers must share (and reflect) the same history (congruence). Concerning the second point, the shared history needs not to be that of populations, since cultural packages can have a different history than the populations [38].

Both horizontal and vertical transmission are more likely to occur between geographically close populations, or between populations revealing signal of common ancestry (like shared ethnonym). Moreover, if a parameter can take a greater number of values, the observation of a shared value between populations is less likely to be due to chance alone. Thus the more variable a parameter is, the more likely a shared parameter values would reveal common history (and not convergence). Thus, we use geographical proximity and common ethnonym to assess the likelihood that parameters values might be shared by common ancestry rather than convergence.

Rhythmic figures may be less prone to convergence than other internal musical parameters

Scales seem prone to convergence since they more likely gather population without ethnonymic link or geographical proximity. Scale is hardly distinctive of a given repertoire in a population, thus it seems that the use of the same scale in the same repertoire in different population could frequently be attributed to chance. Most periodicities values are widespread in Gabon, but some are found predominantly in restricted areas, such as long cycles in the North-East and short cycles in the South.

Pulsation subdivision marks a clear East-West division and extreme periodicities (short and long cycles) are preferentially found in specific regions. Metricity seems to be prone to convergence as well. Rhythmic figures are the most distinctive parameters of all studied. They are also more shared between populations than previously thought. Furthermore, most of the time the corresponding populations are also linked by geographical proximity. This is true for rhythmic cells as well, but less pronounced. It implies that rhythmic figures may be more historically informative than rhythmic cells, though yielding fewer comparison values. All in all, this suggest that rhythmic figure could be less prone to convergence than other internal parameters, making it a good candidate as a marker for historical inference.

We previously thought that shared internal parameter values could better reflect history when considered in a particular repertoire [13]. However, our findings suggest that it may not be the case since in a given population the same parameter value is often found in pieces from various repertoires. In particular, rhythmic figures are less specific of a given repertoire in a population than what was previously thought. Rhythmic figures shared by various populations usually belong to different repertoires. This suggests that musical pieces could be transmitted independently of repertoires. This is consistent with our recent fieldwork observations (unpublished) and those of de Ruyter [14], which revealed that a same piece, identified by its name and recognized by culture upholders, may be found in various population in different repertoires.

All these considerations could not be taken as established tendencies considering music transmission in Gabon. Indeed, shared ethnonym or geographical proximity cannot be an ultimate proof that a given musical parameter is a good marker of history. It does not also allow distinguishing history resulting from common filiation of populations and inter-populations cultural exchange. Nor does it allow identifying musical cultural package of historical coherence [38]. Nevertheless, these considerations reveal interesting leads to explore further.

Conclusion

This study gives a first global overview of Gabonese music. It confirms that Gabon has a singular place among Central Africa: Beat subdivision revealed a clear east-west

organization; there is a wide diversity of musical scales, most of them being heptatonic 691
or hexatonic; repertoires can be shared by various populations; most of repertoires in a 692
given populations are not characterized by a single parameter value. Moreover it helps 693
understand how music could be used to infer history. It suggests that among internal 694
parameters, rhythmic figures may be the best candidate in terms of historical marker. 695

It also opens interesting new questions about music evolution. To better understand 696
how music could be used to infer history, we would need 1) to compare musical data 697
with histories inferred from genetic or linguistic data; 2) to carry detailed regional 698
inter-cultural comparison of known historically related musical material (such as 699
individual pieces), as done by Le Bomin and Bikoma [4] and Boyd et al. [38], which 700
could allow identifying the respective effects of vertical and horizontal transmission on 701
musical material. 702

Supporting information

S1 Protocol – Description of the CRIP method	32
S1 Fig – Proceedings of a recording session with the CRIP technique	33
S2 Table – Description of the populations	34
S2 Fig – Location of the populations studied on the map of Gabon	35
S3 Table – Table of the elementary rhythmic cells	36
S4 Table – Table of the dataset analysed	37
S5 Protocol – Protocol of bootstrap test	48
S5 Table – Result of the bootstrap test of low variability of a parameter inside a repertoire	50
S6 Fig – The three most frequents canonical interval codes sequences (CICS)	51

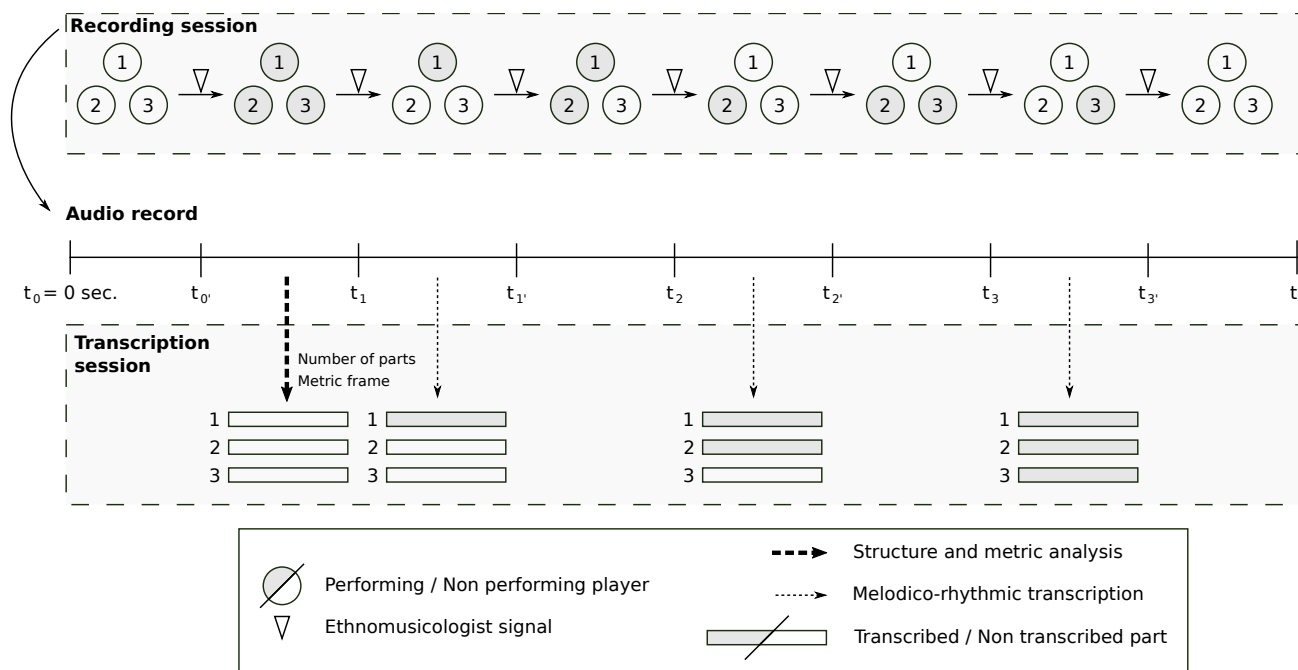
S1 Protocol

Description of the CRIP method. We record music during fieldwork using Continuous Recording of Individual Parts (CRIP). See also S1 Fig for an illustration of the proceedings of such a recording session.

Recording was made by a technique which we refer to as '*continuous recording of individual parts*'. This method is a special kind of analytical recording and is inspired from the '*playback*' recording technique [1, pp. 103–105]. Once the recording device is turned on, performers start to play altogether, as in a regular performance. Then the ethnomusicologist (EM) silently indicates to all performers but one to stop playing, while the last one keeps playing his part. Then the EM indicates to another performer to start playing again, the first one still performing too. When performance is stable enough (which is often very quick, if not immediate), the EM waves the first performer to stop, while the second keeps playing. This process continues until each performer has been recorded separately. By proceeding in this manner, we can get all individual parts of a piece in one recording session (which usually lasts less than 5 minutes), keeping synchronization between parts, thus allowing to accurately transcribe the performance. The EM explains the general purpose of the procedure and the meanings of the signals to the performers beforehand, and a few trial-performances are carried out before recording takes place. Moreover performers are asked to play in the most simple way they can, with as less variation as possible, in order to get a performance which is as close as possible to what Arom [1] would call the '*model*' of the piece.

S1 Fig

Proceedings of a recording session with the CRIP technique. CRIP stands for *Continuous Recording of Individual Parts*. In a single recording session, performers are asked by the ethnomusicologist to successively start and stop playing. Thus each part of the piece is played alone for a while in the record, allowing an accurate transcription of this part. See S1 Protocol for protocol details.



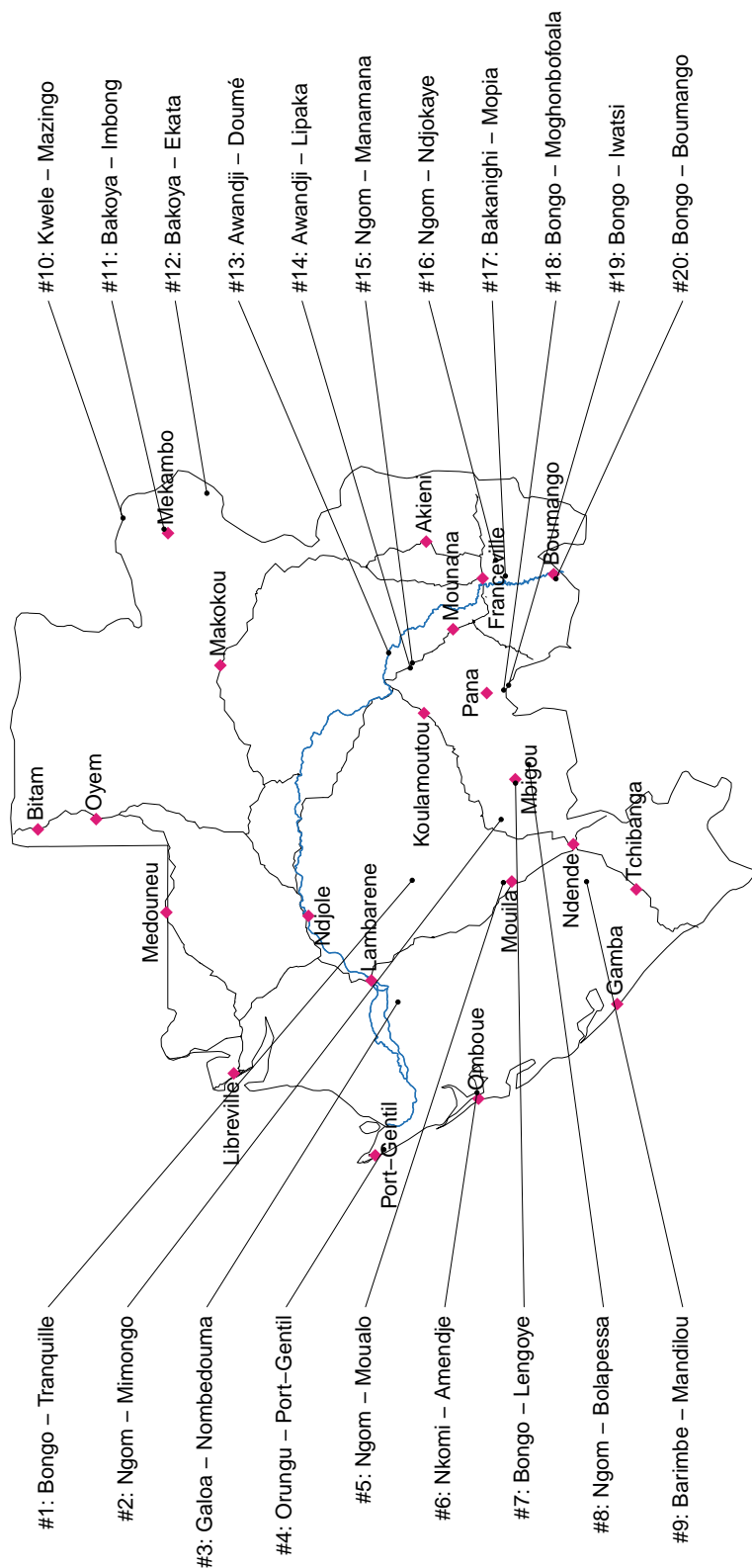
S2 Table

Description of the populations. Table of the populations defined for our analysis. For each population, we give its ethnonym, locality name, longitude and latitude of its locality and province. See S2 Fig for a map.

Number	Ethnonym	Locality	Longitude	Latitude	Province
1	Bongo	Tranquille	11.06	-1.04	Ngounié
2	Ngom	Mimongo	11.57	-1.78	Ngounié
3	Galoa	Nombedouma	10.05	-0.92	Moyen-Ogooué
4	Orungu	Port-Gentil	8.81	-0.80	Ogooué-Maritime
5	Ngom	Moualo	11.04	-1.80	Ngounié
6	Nkomi	Amendje	9.29	-1.58	Ogooué-Maritime
7	Bongo	Lengoye	11.87	-1.90	Ngounié
8	Ngom	Bolapessa	12.03	-2.01	Ngounié
9	Barimbe	Mandilou	11.05	-2.50	Nyanga
10	Kwele	Mazingo	14.09	1.38	Ogooué-Ivindo
11	Bakoya	Imbong	14.00	1.03	Ogooué-Ivindo
12	Bakoya	Ekata	14.30	0.68	Ogooué-Ivindo
13	Awandji	Doumé	12.96	-0.84	Ogooué-Lolo
14	Awandji	Lipaka	12.84	-1.02	Ogooué-Lolo
15	Ngom	Manamana	12.88	-1.04	Ogooué-Lolo
16	Ngom	Ndjokaye	13.74	-1.74	Haut-Ogooué
17	Bakanighi	Mopia	13.60	-1.82	Haut-Ogooué
18	Bongo	Moghonbofoala	12.65	-1.80	Ogooué-Lolo
19	Bongo	Iwatsi	12.69	-1.85	Ogooué-Lolo
20	Bongo	Boumango	13.58	-2.24	Haut-Ogooué

S2 Fig

Location of the populations studied on the map of Gabon. The black points indicate the positions of the populations. Purple diamond: main cities. Black lines: main roads. Blue line: watercourse of the Ogooué river. First part of the name refers to the ethnonym of the population and second part to its location. See S2 Table for details on populations positions.



S3 Table

Table of the elementary rhythmic cells. The table gives for each rhythmic cell the corresponding identifier used in the body of the text. Rhythmic cells used to describe pieces with ternary beat subdivision are colored in blue. Rhythmic cells used to describe pieces with binary beat subdivision are colored in red.

1	18	34	52
2	19	35	53
3	20	36	54
4	21	37	55
5	22	38	56
6	23	41	60
7	24	42	61
8	25	43	62
9	26	44	63
10	27	45	64
12	28	46	65
13	29	47	66
14	30	48	67
15	31	49	68
16	32	50	69
17	33	51	





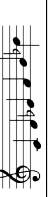
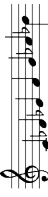

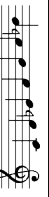









S4 Table








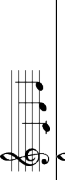









Table of the dataset analysed. It provides the parameters values identified for each recording transcribed and analysed.

We used the following notation to describe scales, inspired from LilyPond format [39]. Each degree of the diatonic scale is represented by its corresponding letter in the English notation system (from ‘a’ to ‘g’) and successive degrees of the scale are separated by a comma. Accidentals ‘#’ and ‘b’, respectively increasing or decreasing the pitch of roughly one semitone, are indicated respectively with ‘is’ or ‘es’ following the degree letter. For example, the scale ‘II’ on the left side of S6 Fig would be written ‘c,des,ees,f,aes,bes’.





















We used the following format to describe the rhythmic organization of the pieces: ‘Instr1:ID1-1,ID1-2,...,ID1-n/.../InstrX:IDX-1,IDX-2,...,IDX-m’. Where ‘Instr1, ..., InstrX’ are the names of the instruments playing the X parts of the rhythmic reference. ‘IDi-j’ are the identifiers (see S3 Table) of the rhythmic cells found in the corresponding part.

See ‘Material and methods’ part for other parameters description.

Piece ID	Population	Repertoire	Metricity	Rhythmic figures	Scale class	CICS	Scale
RAEI065	Awandji - Doumé	<i>Bubegu</i>	16/3	Drum 1: 7,8,22,23	(7-2)	1221222	
RAEI066	Awandji - Doumé	<i>Bubegu</i>	8/3	Drum 1: 7,9,22,23 Drum 2: 2,10,2,65	(7-2)	1221222	
RAEI063	Awandji - Doumé	<i>Doka</i>	8/2	Drum 1: 34,34,46,46 Drum 2: 34,34,34,46	(7-2)	1221222	
RAEI064	Awandji - Doumé	<i>Doka</i>	8/2		(6-1)	122322	
RAEI076	Awandji - Doumé	<i>Issogha</i>	16/2		(6-1)	122322	
RAEI059	Awandji - Doumé	<i>Lisimbu</i>	8/3				
RAEI060	Awandji - Doumé	<i>Lisimbu</i>	8/3	Drum 1: 7,8,22,23	(7-2)	1221222	
RAEI061	Awandji - Doumé	<i>Lisimbu</i>	16/3	Drum 1: 7,8,22,23	(7-2)	1221222	
RAEI068	Awandji - Doumé	<i>Mbadi</i>	16/2	Drum 1: 34,34,34,46 Drum 2: 34,34,46,46	(7-2)	1221222	
RAEI071	Awandji - Doumé	<i>Mbadi</i>	8/2		(6-1)	123222	
RAEI075	Awandji - Doumé	<i>Mbudi</i>	12/2	Drum 1: 34,34,49 Drum 2: 30,46,31,48,34,31	(7-2)	1221222	
RAEI074	Awandji - Doumé	<i>Mungala</i>	8/3	Drum 1: 7,8,22,23 Drum 2: 2,10,4,8			
RAEI072	Awandji - Doumé	<i>Nzobi</i>	8/3	Drum 1: 1,3,3 Drum 2: 2,17,17	(7-2)	1221222	
RAEI236	Awandji - Lipaka	<i>Mbadi</i>	16/2				
RAEI238	Awandji - Lipaka	<i>Mbadi</i>	8/2				
RAEI239	Awandji - Lipaka	<i>Mbadi</i>	16/3				
RAEI030	Bakanighi - Mopia	<i>Lisimbu</i>	12/3	Drum 1: 7,9,22,23 Drum 2: 4,19,6,19	(6-3)	114213	
RAEI031	Bakanighi - Mopia	<i>Lisimbu</i>	12/3	Drum 1: 3,6,4,4 Drum 2: 6,19	(6-1)	122322	
RAEI032	Bakanighi - Mopia	<i>Mopfougou</i>	12/3	Drum: 3,3,25	(4-1)	1522	
RAEI033	Bakanighi - Mopia	<i>Mopfougou</i>	12/3	Drum: 3,21,25,25	(4-0)	2325	
RAEI028	Bakanighi - Mopia	<i>Mungala</i>	12/3	Drum 1: 4,9,22,23 Drum 2: 3,19,6,19	(6-1)	122232	
RAEI029	Bakanighi - Mopia	<i>Mungala</i>	8/3	Drum 1: 4,9,9,23	(7-2)	1221222	

Piece ID	Population	Repertoire	Metricity	Rhythmic figures	Scale class	CICS	Scale
RAEI026	Bakanighi - Mopia	<i>Nzobi</i>	6/3	Drum 2: 2,3,2,4 Drum 1: 3 Drum 2: 6,19,26	(6-1)	122322	
RAEI027	Bakanighi - Mopia	<i>Nzobi</i>	12/3	Drum 1: 3 Drum 2: 6,19,26	(7-2)	1221222	
RAEI125	Bakoya - Elkata	<i>Boumba</i>	16/3	Drum 1: 10,19,23 Drum 2: 3,4,19	(7-2)	1221222	
RAEI136	Bakoya - Elkata	<i>Ibwema</i>	16/3	Drum 1: 2,4,3 Drum 2: 3,12,4	(8-4)	11122122	
RAEI137	Bakoya - Elkata	<i>Ibwema</i>	16/3	Drum 1: 3,4 Drum 2: 3,4,3,23	(6-2)	123213	
RAEI127	Bakoya - Elkata	<i>Lisimbu</i>	16/3	Drum 1: 3,10 Drum 2: 3,19,68,19	(7-2)	1221222	
RAEI126	Bakoya - Elkata	<i>Mungala</i>	6/3	Drum 1: 3 Drum 2: 3			
RAEI138	Bakoya - Elkata	<i>Ngi</i>	12/3	Drum 1: 9,19,63 Drum 2: 3,3,63	(6-2)	122142	
RAEI129	Bakoya - Elkata	<i>Ngodja</i>	8/3	Drum: 3,10,6	(3-0)	224	
RAEI130	Bakoya - Elkata	<i>Ngodja</i>	8/3	Drum 1: 3,19,69,10	(5-1)	12252	
RAEI124	Bakoya - Elkata	<i>Nzokou</i>	20/3	Drum 1: 2,10,3,4 Drum 2: 3,23,4,23	(6-1)	122322	
RAEI128	Bakoya - Elkata	<i>Nzokou</i>	24/3	Drum 1: 2,10,3,4 Drum 2: 3,4,3,10	(8-4)	11122122	
RAEI139	Bakoya - Imbong	<i>Boumba</i>	12/3	Drum 1: 3,4	(7-3)	1123212	
RAEI261	Bakoya - Imbong	<i>Boumba</i>	12/3	Drum 1: 3,4	(7-3)	1123212	
RAEI145	Bakoya - Imbong	<i>Bwiti koji</i>	16/3	Drum 1: 3,21 Drum 2: 3,6	(7-2)	1221222	
RAEI146	Bakoya - Imbong	<i>Bwiti koji</i>	8/3	Drum 1: 19,21 Drum 2: 3,12 Drum 3: 2,3,27,4	(7-3)	1112322	
RAEI131	Bakoya - Imbong	<i>Ibwema</i>	24/3		(8-4)	11121222	
RAEI135	Bakoya - Imbong	<i>Ibwema</i>	8/3		(8-5)	11112123	

















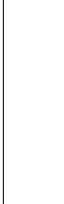
S4 Fig. – 2/10

Piece ID	Population	Repertoire	Metricity	Rhythmic figures	Scale class	CICS	Scale
RAEI147	Bakoya - Imbong	<i>Ibwema</i>	8/3	Drum 1: 3,10	(6-2)	121224	
RAEI148	Bakoya - Imbong	<i>Ibwema</i>			(5-2)	12126	
RAEI141	Bakoya - Imbong	<i>Langu</i>	8/3	Drum 1: 3,19 Drum 2: 3,33	(5-0)	22323	
RAEI142	Bakoya - Imbong	<i>Langu</i>	8/3		(7-2)	1221222	
RAEI143	Bakoya - Imbong	<i>Lisimbu</i>	16/3		(7-2)	1221222	
RAEI144	Bakoya - Imbong	<i>Lisimbu</i>	8/3		(7-2)	1221222	
RAEI152	Bakoya - Imbong	<i>Magnele</i>	8/2		(8-4)	11122212	
RAEI153	Bakoya - Imbong	<i>Magnele</i>	24/2		(8-4)	11122122	
RAEI133	Bakoya - Imbong	<i>Ngodja</i>	12/3		(7-2)	1221222	
RAEI134	Bakoya - Imbong	<i>Ngodja</i>	18/3		(6-1)	122322	
RAEI149	Bakoya - Imbong	<i>Ngodja</i>	12/3		(6-1)	123222	
RAEI150	Bakoya - Imbong	<i>Ngodja</i>	12/3		(6-1)	122322	
RAEI043	Barimbe - Mandilou	<i>Mbumba</i>	8/2	Drum 1: 31,47,46 Drum 2: 43,48,48,46,48,48	(7-2)	1221222	
RAEI041	Barimbe - Mandilou	<i>Mimbwiri</i>	8/3	Drum 1: 2,3 Drum 2: 7,7,14,14	(6-1)	123222	
RAEI042	Barimbe - Mandilou	<i>Ologo</i>	16/2	Drum 1: 30,46	(6-1)	123222	
RAEI034	Bongo - Boumango	<i>Lisimbu</i>	8/3	Drum 1: 2,12,4,3 Drum 2: 2,2,27,3	(7-2)	1221222	
RAEI254	Bongo - Boumango	<i>Lisimbu</i>	8/3	Drum 1: 2,12,4,3 Drum 2: 2,2,27,3	(6-1)	122322	
RAEI035	Bongo - Boumango	<i>Mungala</i>	12/3	Drum 1: 3 Drum 2: 3	(7-2)	1221222	
RAEI036	Bongo - Boumango	<i>Mungala</i>	16/3	Drum 1: 2,4,4,3 Drum 2: 1,1,4,4	(7-2)	1221222	
RAEI039	Bongo - Boumango	<i>Ngwata</i>		Drum 2: 2,25,3,19			
RAEI037	Bongo - Boumango	<i>Wuya</i>	12/3	Drum 1: 2,23,4,3	(6-2)	121422	





















S4 Fig. – 3/10

Piece ID	Population	Repertoire	Metricity	Rhythmic figures	Scale class	CICS	Scale
RAEI038	Bongo - Boumango	<i>Wuya</i>	12/3	Drum 2: 1,10,4,4 Drum 1: 3,3,25 Drum 2: 3	(5-1)	12342	
RAEI008	Bongo - Iwatsi	<i>Laka</i>	8/2	Drum 1: 34,44 Drum 2: 34,45			
RAEI010	Bongo - Iwatsi	<i>Lisimbu</i>	8/3	Drum 1: 3,3,3,6 Drum 2: 3,19,6,19	(6-2)	122124	
RAEI012	Bongo - Iwatsi	<i>Mungala</i>	12/3	Drum 1: 7,23 Drum 2: 3	(6-1)	123222	
RAEI013	Bongo - Iwatsi	<i>Mungala</i>	6/3	Drum 1: 29 Drum 2: 5	(7-2)	1221222	
RAEI006	Bongo - Iwatsi	<i>Ngoye</i>	8/2	Drum 1: 34,43,35 Drum 2: 37	(7-2)	1212222	
RAEI014	Bongo - Iwatsi	<i>Nzobi</i>	6/3	Drum 2: 3	(6-2)	122214	
RAEI015	Bongo - Iwatsi	<i>Nzobi</i>	12/3	Drum 1: 3,19,25 Drum 2: 25	(7-3)	1132122	
RAEI089	Bongo - Lengoye	<i>Ibwema</i>	8/3	Drum 1: 3,25	(7-2)	1221222	
RAEI088	Bongo - Lengoye	<i>Kengé</i>	4/2		(5-0)	22233	
RAEI083	Bongo - Lengoye	<i>Mavasa</i>	8/3	Drum 1: 17,66 Drum 2: 1,13,12,65	(6-1)	122322	
RAEI085	Bongo - Lengoye	<i>Mavasa</i>	12/3	Drum 1: 6,19 Drum 2: 2,13,25,13	(5-1)	12522	
RAEI084	Bongo - Lengoye	<i>Mujaja</i>	16/2	Drum 1: 34,34,46,46 Drum 2: 31,46	(7-2)	1221222	
RAEI092	Bongo - Lengoye	<i>Mujisi</i>	8/3		(7-2)	1221222	
RAEI093	Bongo - Lengoye	<i>Mujisi</i>	16/2	Drum 1: 33	(7-2)	1221222	
RAEI086	Bongo - Lengoye	<i>Mungala</i>	4/2	Drum 1: 51 Drum 2: 31,52,38,53	(6-1)	132222	
RAEI090	Bongo - Lengoye	<i>Mungala</i>	8/3	Drum 1: 3,25,6,10	(7-2)	1221222	
RAEI087	Bongo - Lengoye	<i>Muwiri</i>	4/3		(5-2)	12216	
RAEI091	Bongo - Lengoye	<i>Muwiri</i>	6/3	Drum 1: 63	(7-2)	1221222	














S4 Fig. – 4/10










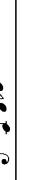


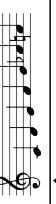

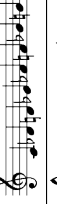



Piece ID	Population	Repertoire	Metricity	Rhythmic figures	Scale class	CICS	Scale
RAEI004	Bongo - Moghombfoala	<i>Laka</i>	12/2	Drum 2: 13,25 Drum 1: 41 Drum 2: 41,42	(7-2)	1221222	
RAEI005	Bongo - Moghombfoala	<i>Laka</i>	8/2	Drum 1: 41 Drum 2: 41,42	(6-1)	122322	
RAEI156	Bongo - Moghombfoala	<i>Laka</i>	8/2	Drum 1: 41 Drum 2: 41,42	(6-1)	122322	
RAEI157	Bongo - Moghombfoala	<i>Laka</i>	16/2	Drum 1: 41 Drum 2: 41,42	(6-1)	122322	
RAEI252	Bongo - Moghombfoala	<i>Lisimbu</i>	8/3	Drum 1: 3,3,3,18 Drum 2: 3,8,3,16	(6-1)	123222	
RAEI258	Bongo - Moghombfoala	<i>Ngoye</i>	6/2		(7-2)	1212222	
RAEI003	Bongo - Moghombfoala	<i>Ngwata</i>	8/2	Drum 1: 34,36 Drum 2: 31			
RAEI154	Bongo - Moghombfoala	<i>Ngwata</i>	4/2		(4-0)	2334	
RAEI155	Bongo - Moghombfoala	<i>Ngwata</i>	8/2	Drum 1: 34,36 Drum 2: 31	(8-4)	11121222	
RAEI260	Bongo - Moghombfoala	<i>Ngwata</i>	4/2				
RAEI002	Bongo - Moghombfoala	<i>Nzobi</i>	6/3	Drum 1: 3,19,20 Drum 2: 3	(6-2)	122214	
RAEI253	Bongo - Moghombfoala	<i>Nzobi</i>	12/3	Drum 1: 3,19,20 Drum 2: 3	(5-0)	22323	
RAEI121	Bongo - Tranquille	<i>Bwiti missoko</i>	16/2	Drum 1: 31,31,32,48 Drum 2: 31,32,48,47	(6-1)	122322	
RAEI119	Bongo - Tranquille	<i>Disumba</i>	12/2	Drum 1: 31,48,46 Drum 2: 31,46,48	(6-1)	123222	
RAEI120	Bongo - Tranquille	<i>Mbumba</i>	8/2	Drum 1: 31,31,34,48 Drum 2: 31,32,48,47	(5-1)	12432	
RAEI122	Bongo - Tranquille	<i>Nyembe</i>	8/2	Drum 1: 31,31,34,48 Drum 2: 31,32,48,47	(4-0)	2235	
RAEI123	Bongo - Tranquille	<i>Nyembe</i>	16/2	Drum 1: 31,31,32,34 Drum 2: 31,32,48,47	(6-1)	122322	
RAEI194	Galao - Nombedouma	<i>Abambo</i>	8/3	Drum: 2,10,3	(5-0)	22233	
RAEI192	Galao - Nombedouma	<i>Mabandji</i>	4/3	Drum: 2,10,3	(5-0)	22233	


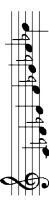
















S4 Fig. – 5/10

Piece ID	Population	Repertoire	Metricity	Rhythmic figures	Scale class	CICS	Scale
RAEI191	Galao - Nombedouma	<i>Mimbwiri</i>	8/2	Drum: 31,54	(7-2)	1221222	
RAEI193	Galao - Nombedouma	<i>Mimbwiri</i>			(7-2)	1221222	
RAEI162	Kwele - Mazingo	<i>Beka</i>	16/2		(7-2)	1221222	
RAEI163	Kwele - Mazingo	<i>Beka</i>	16/3		(7-2)	1221222	
RAEI168	Kwele - Mazingo	<i>Berceuse</i>	16/3		(7-2)	1221222	
RAEI169	Kwele - Mazingo	<i>Berceuse</i>	4/3		(6-1)	123222	
RAEI160	Kwele - Mazingo	<i>Déké</i>	16/3	Drum: 4,19,27	(6-1)	122322	
RAEI161	Kwele - Mazingo	<i>Déké</i>	24/3		(7-2)	1221222	
RAEI167	Kwele - Mazingo	<i>Magnele</i>	16/3	Drum: 3,4	(7-2)	1221222	
RAEI165	Kwele - Mazingo	<i>Mavasa</i>	8/3	Drum: 4,12	(5-2)	12414	
RAEI166	Kwele - Mazingo	<i>Mavasa</i>	16/3	Drum: 4,19	(7-2)	1221222	
RAEI158	Kwele - Mazingo	<i>Ngi</i>	16/3		(7-2)	1221222	
RAEI159	Kwele - Mazingo	<i>Ngi</i>	16/3		(7-2)	1221222	
RAEI164	Kwele - Mazingo	<i>Ngi</i>	8/3	Drum: 2,3,19	(4-1)	1452	
RAEI094	Ngom - Bolapessa	<i>Berceuse</i>	8/2		(5-0)	22224	
RAEI095	Ngom - Bolapessa	<i>Berceuse</i>	12/2		(6-1)	122232	
RAEI096	Ngom - Bolapessa	<i>Dijangu</i>	8/2	Drum 1: 32,46,33,54	(9-6)	11112212	
RAEI101	Ngom - Bolapessa	<i>Langu</i>	8/2		(5-0)	22323	
RAEI103	Ngom - Bolapessa	<i>Langu</i>	12/2		(5-1)	12522	
RAEI104	Ngom - Bolapessa	<i>Lisimbu</i>	12/2	Drum 1: 30,31,46 Drum 2: 31,46,50	(7-2)	1221222	










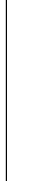
S4 Fig. – 6/10

Piece ID	Population	Repertoire	Metricity	Rhythmic figures	Scale class	CICS	Scale
RAEI105	Ngom - Bolapessa	<i>Lisimbu</i>	12/2	Drum 1: 30,31,46	(7-2)	1221222	
RAEI097	Ngom - Bolapessa	<i>Makouma</i>	4/2	Drum 1: 31,33 Drum 2: 31,46	(5-1)	15222	
RAEI098	Ngom - Bolapessa	<i>Makouma</i>	8/2	Drum 1: 31,48,54	(6-1)	122223	
RAEI100	Ngom - Bolapessa	<i>Makouma</i>	16/2	Drum 1: 33,43 Drum 2: 30,46	(7-2)	1221222	
RAEI099	Ngom - Bolapessa	<i>Mundukue</i>	8/2	Drum 1: 30,48,54,35 Drum 2: 33,48,33,54	(6-3)	111252	
RAEI047	Ngom - Manamana	<i>Bubeyu</i>	8/3	Drum 1: 3,60,6,7 Drum 2: 3,8,10,4	(7-2)	1221222	
RAEI048	Ngom - Manamana	<i>Bubeyu</i>	8/3	Drum 1: 3,23,7,8 Drum 2: 3,23,7,61			
RAEI044	Ngom - Manamana	<i>Lisimbu</i>	8/3	Drum 1: 25 Drum 2: 2,12,3,19	(7-2)	1221222	
RAEI045	Ngom - Manamana	<i>Lisimbu</i>	8/3	Drum 1: 25 Drum 2: 3,19,18,12	(6-2)	122124	
RAEI046	Ngom - Manamana	<i>Lisimbu</i>	8/3	Drum 1: 25 Drum 2: 3,19,22,23			
RAEI049	Ngom - Manamana	<i>Mungala</i>	12/3	Drum 1: 2 Drum 2: 2,19,3	(7-2)	1221222	
RAEI050	Ngom - Manamana	<i>Mungala</i>		Drum 1: 3,9,22,5 Drum 2: 4,4,10,23			
RAEI051	Ngom - Manamana	<i>Mungala</i>	16/3	Drum 1: 2,3,2,4 Drum 2: 2,10,4,9	(7-2)	1221222	
RAEI052	Ngom - Manamana	<i>Ngoye</i>	8/3	Drum 1: 3,3,25,10 Drum 2: 2,19,10,3	(5-0)	22323	
RAEI053	Ngom - Manamana	<i>Ngoye</i>	8/3	Drum 1: 3,3,62,19 Drum 2: 3,21,14,27	(5-1)	12324	
RAEI054	Ngom - Manamana	<i>Ngoye</i>	8/3	Drum 1: 3,3,25,10 Drum 2: 3,19,62,27,3,21,13,27			
RAEI055	Ngom - Manamana	<i>Nzobi</i>	12/3	Drum 1: 63 Drum 2: 2,19,21	(7-2)	1221222	
RAEI056	Ngom - Manamana	<i>Nzobi</i>	12/3	Drum 1: 63 Drum 2: 2,3,3			
RAEI057	Ngom - Manamana	<i>Nzobi</i>		Drum 1: 64			

Piece ID	Population	Repertoire	Metricity	Rhythmic figures	Scale class	CICS	Scale
RAEI188	Ngom - Mimongo	<i>Beka</i>		Drum 2: 2,3,25	(3-0)	253	
RAEI189	Ngom - Mimongo	<i>Beka</i>			(4-0)	2253	
RAEI179	Ngom - Mimongo	<i>Berceuse</i>	8/2		(4-0)	2253	
RAEI175	Ngom - Mimongo	<i>Kunde</i>	8/2	Drum 1: 30,46 Drum 2: 31,46	(6-1)	122322	
RAEI176	Ngom - Mimongo	<i>Kunde</i>	8/2	Drum 1: 30,46	(6-2)	122142	
RAEI170	Ngom - Mimongo	<i>Lisimbu</i>	12/2	Drum: 30,31	(6-1)	122322	
RAEI171	Ngom - Mimongo	<i>Lisimbu</i>	12/2	Drum: 30,31	(7-2)	1221222	
RAEI187	Ngom - Mimongo	<i>Mabulugongwe</i>	8/2		(4-0)	2343	
RAEI185	Ngom - Mimongo	<i>Makouma</i>	16/2	Drum 1: 30,46 Drum 2: 30,33,30,46	(7-3)	1112232	
RAEI186	Ngom - Mimongo	<i>Makouma</i>	16/2		(6-2)	122142	
RAEI172	Ngom - Mimongo	<i>Minonze</i>	8/2	Drum 1: 30,31,31 Drum 2: 32,33,34			
RAEI173	Ngom - Mimongo	<i>Minonze</i>	12/2	Drum 1: 32,33,47	(8-4)	11122212	
RAEI174	Ngom - Mimongo	<i>Minonze</i>	12/2	Drum 1: 31,46,50	(6-1)	123222	
RAEI182	Ngom - Mimongo	<i>Mundukwe</i>	12/2	Drum 1: 30,33,46,31 Drum 2: 31,46,50	(7-3)	1132212	
RAEI183	Ngom - Mimongo	<i>Mundukwe</i>	12/2	Drum 1: 30,31,31,31,31,46 Drum 2: 31,31,31,31,46,50	(7-3)	1123212	
RAEI184	Ngom - Mimongo	<i>Mundukwe</i>	12/2		(8-7)	11111115	
RAEI180	Ngom - Mimongo	<i>Mungala</i>	12/2	Drum 1: 50 Drum 2: 31,31,31,31,31,48	(6-2)	113232	
RAEI181	Ngom - Mimongo	<i>Mungala</i>	12/2	Drum 1: 30,31,31 Drum 2: 31,31,31,31,46,50	(6-1)	122322	
RAEI177	Ngom - Mimongo	<i>Nawul</i>			(4-1)	1522	

Piece ID	Population	Repertoire	Metricity	Rhythmic figures	Scale class	CICS	Scale
RAEI178	Ngom - Mimongo	<i>Nawul</i>			(4-0)	2253	
RAEI106	Ngom - Moualo	<i>Kore</i>	16/2	Drum 1: 31,48,38,48 Drum 2: 31,33,54,50	(7-2)	1221222	
RAEI107	Ngom - Moualo	<i>Kumde</i>	16/2		(5-1)	12234	
RAEI117	Ngom - Moualo	<i>Lisimbu</i>	12/2	Drum 1: 30,31,31,31,31,46 Drum 2: 31,31,31,46,31,34	(7-2)	1221222	
RAEI118	Ngom - Moualo	<i>Lisimbu</i>	12/2	Drum 1: 31,46 Drum 2: 31,34,46	(7-2)	1221222	
RAEI114	Ngom - Moualo	<i>Mabandji</i>	12/2	Drum 1: 31,55,55 Drum 2: 31,55,50	(6-1)	122232	
RAEI111	Ngom - Moualo	<i>Mabundi</i>	16/2	Drum 1: 30,30,31,34 Drum 2: 31,33,32,34	(5-1)	12234	
RAEI109	Ngom - Moualo	<i>Makouma</i>	8/2		(5-1)	12522	
RAEI112	Ngom - Moualo	<i>Makouma</i>	8/2	Drum 1: 30,46 Drum 2: 33,34	(5-1)	12522	
RAEI113	Ngom - Moualo	<i>Mbumba</i>	12/2	Drum 1: 31,33,47 Drum 2: 31,55,55	(8-4)	11212122	
RAEI115	Ngom - Moualo	<i>Minonze</i>	12/2	Drum 1: 31,46,34 Drum 2: 31,47,46	(7-2)	1221222	
RAEI108	Ngom - Moualo	<i>Mungala</i>	12/2	Drum 1: 30 Drum 2: 33,34,34,46	(5-1)	14322	
RAEI110	Ngom - Moualo	<i>Mungala</i>	12/2	Drum 1: 31,31,31,47 Drum 2: 31,46,50,34	(6-1)	122232	
RAEI019	Ngom - Ndjokaye	<i>Lisimbu</i>	8/3	Drum 1: 3,3,23,4 Drum 2: 2,3	(7-2)	1221222	
RAEI020	Ngom - Ndjokaye	<i>Lisimbu</i>	8/3	Drum 1: 3,23,4,9 Drum 2: 2,4,9,3	(6-1)	122322	
RAEI016	Ngom - Ndjokaye	<i>Mungala</i>	6/3	Drum 1: 2,19,21 Drum 2: 2,2,3	(5-1)	12252	
RAEI017	Ngom - Ndjokaye	<i>Mungala</i>	8/3	Drum 1: 3,22,23,23 Drum 2: 2,9,25,24	(6-1)	122322	
RAEI018	Ngom - Ndjokaye	<i>Mungala</i>	12/3				
RAEI023	Ngom - Ndjokaye	<i>Ngoye</i>	8/3				
RAEI021	Ngom - Ndjokaye	<i>Nzobi</i>	12/3	Drum 1: 2,19,21 Drum 2: 3	(6-1)	122322	

S4 Fig. – 9/10

Piece ID	Population	Repertoire	Metricity	Rhythmic figures	Scale class	CICS	Scale
RAEI022	Ngom - Ndjokaye	<i>Nzobi</i>	24/3	Drum 1: 2,21 Drum 2: 2,2,25,3	(6-1)	123222	
RAEI201	Nkomi - Amendje	<i>Mabandji</i>	8/3	Drum: 3,10	(7-2)	1221222	
RAEI202	Nkomi - Amendje	<i>Mabandji</i>	8/3	Drum: 3,10	(6-1)	123222	
RAEI199	Nkomi - Amendje	<i>Mimbwiri</i>	8/2	Drum: 50	(6-1)	122322	
RAEI200	Nkomi - Amendje	<i>Mimbwiri</i>	8/2	Drum: 54,56	(6-1)	122322	
RAEI210	Orungu - Port-Gentil	<i>Agombénéno</i>	8/3	Drum: 4,7	(5-1)	12252	
RAEI215	Orungu - Port-Gentil	<i>Ivanga</i>			(5-1)	14322	
RAEI211	Orungu - Port-Gentil	<i>Mabandji</i>	8/3	Drum: 3,10	(6-2)	121422	
RAEI212	Orungu - Port-Gentil	<i>Mabandji</i>	8/3	Drum: 3,10	(6-1)	123222	
RAEI207	Orungu - Port-Gentil	<i>Mimbwiri</i>	8/3	Drum: 7,17	(6-1)	123222	

S5 Protocol

Protocol of bootstrap test. Protocol describing the bootstrap test used to test for significant low variability of parameters values inside a repertoire. Results are given in S5 Table.

We wanted to test if the diversity of parameters values was significantly lower inside a repertoire among different population than in the whole dataset. We tested it by using bootstrap analysis. For each parameter and each repertoire, we resampled (with replacement) the values found in all pieces. Thus, for each piece this parameter was attributed a random value taken from the actual distribution of values. We then looked at the number of different values found for this repertoire in the simulated datasets. We repeated this 1000 times and looked how many times the number of different values was lower or equal to that the observed value. If it was lower or equal in less than 5% of resampled datasets (i.e. less than 50 times), the diversity of values for this parameter was considered to be significantly lower in this repertoire.

However, we know that parameters values tend to be identical in the pieces of the same repertoire of a given population, which could explain lower diversity in a given repertoire among all populations. Put another way, diversity of rhythmic figures values in *lisimbu* repertoire may be low because in each population, various pieces of this repertoire tend to have similar values (and not because CICS values in a repertoire are shared among different populations). In order to remove the effect of intra-population low variability of repertoires, we proceeded as follow:

For each repertoire of each population (i.e. each 'population-repertoire pair') we checked if all values of the parameter (e.g. CICS) were identical. This allowed us to compute the 'percentage of homogeneous population-repertoire pairs' (pcHPR). pcHPR was 33% for rhythmic figure of the first drum: in 14 over the 42 population-repertoire pairs for which we had rhythmic figure data for the first drum, all pieces had the same rhythmic figure.

Then we resampled the parameters values so that it kept the same low diversity inside a population-repertoire pair. We proceeded as follows: For each population-repertoire pair, we randomly sampled the first element. Then we took a random percentage: if it was lower than the pcHPR, we took the same value for all other pieces of this population-repertoire pair. If it was higher than the pcHPR, we took at least one other value among the other pieces of this population-repertoire pair. We repeated this sampling 1000 times.

Then for each repertoire, we count the number of different values found (among all populations) and compared it the number of different values found for it in each sample (as explained above). The p-value given for each repertoire and each parameter in the table is the percentage of samples where the number of different parameters values found for this repertoire is lower or equal than the observed number in the real dataset.

Here is an example concerning rhythmic figure of the first drum in *lisimbu* repertoire. We found 23 pieces from

lisimbu repertoire, but we identified the rhythmic figure of the first drum for only 20 of them. These 20 pieces come from 11 different populations. In all these pieces from the *lisimbu* repertoire, we found 14 different rhythmic figure values for the first drum. As stated above, 33% of population-repertoire pairs have the same rhythmic figure for the drum 1 (pcHPR = 0.33). We have rhythmic figures data for 11 populations. Let's take one of the those, the Ngom from Manamana #15: we have three pieces of *lisimbu* with rhythmic figure of the first drum. For the first of these pieces, we take a random values of rhythmic figures among those found in all pieces (from all populations and all repertoires). Then we take a number between 0 and 1. If it is lower than 0.33 (pcHPR), we take the same rhythmic figure values for the other two pieces. If it is greater than 0.33, we take (from the full sample) a different value for the second piece and then a random value (identical or different) for the last piece. This ensures that the proportion of population-repertoire pairs where all pieces have the same rhythmic figure is (on average) the same as in the real dataset. We proceed like this for all other populations where we have rhythmic figures for the *lisimbu*. This gives us a random sample of 20 rhythmic figures. We count how many different rhythmic figures we have in this sample. We repeat all the above procedure (sampling and different values counting) 1000 times. Among 1000 samples, there are 15% of samples in which the number of different rhythmic figure values is lower or equal to 14 (i.e. the observed value in our dataset). The corresponding *p-value* is 0.147, which is higher than 0.05: the diversity of rhythmic figures values in *lisimbu* repertoire is not considered significantly low.

S5 Table

Result of the bootstrap test of low variability of a parameter inside a repertoire. For each repertoire and each parameter, this table gives the p-value of bootstrap test, i.e. the percent of random sample where we found a number of different values lower or equal to that really observed. Values are missing when the bootstrap test can not be applied. Significant values are prefixed with a star '*'. See S5 Protocol for more details.

	CICS	Metricity	Nb of puls.	Puls. subdiv.	RF (drum 1)	RF (drum 2)
<i>Abambo</i>						
<i>Agombénéno</i>						
<i>Beka</i>	0.548	1	0.451	1		
<i>Berceuse</i>	1	0.803	0.984	1		
<i>Boumba</i>	0.304	0.573	0.738	0.481	0.346	
<i>Bubeyu</i>	*0.032	0.199	0.476	0.244	1	1
<i>Disumba</i>						
<i>Bwiti koji</i>	1	1	1	0.537	1	1
<i>Bwiti missoko</i>						
<i>Déké</i>	1	1	1	0.504		
<i>Dijangu</i>						
<i>Doka</i>	1	0.148	0.301	0.49		
<i>Ibwema</i>	1	0.177	0.611	*0.037	1	1
<i>Issogha</i>						
<i>Ivanga</i>						
<i>Kengé</i>						
<i>Kore</i>						
<i>Kunde</i>	1	0.348	0.635	0.277	*0.018	
<i>Laka</i>	0.239	0.591	0.865	0.236	*0.001	*0
<i>Langu</i>	0.572	0.85	0.487	1		
<i>Lisimbu</i>	*0	*0.001	*0.033	1	0.147	0.96
<i>Mabandji</i>	0.874	0.299	0.716	1	*0.015	
<i>Mabundi</i>						
<i>Mabulugongwe</i>						
<i>Makouma</i>	0.906	0.201	0.582	0.065	0.124	1
<i>Magnele</i>	1	1	1	1		
<i>Mavasa</i>	1	0.75	0.917	0.253	1	1
<i>Mbadi</i>	1	0.666	0.475	1		
<i>Mbudi</i>						
<i>Mbumba</i>	1	0.368	0.617	0.252	1	1
<i>Mimbwiri</i>	0.139	0.08	*0.008	1	1	
<i>Minonze</i>	1	0.14	0.41	0.142	1	1
<i>Mopfougou</i>	1	0.431	0.433	0.933	1	
<i>Mujaja</i>						
<i>Mujisi</i>	0.115	1	1	1		
<i>Mundukwe</i>	1	0.238	0.509	0.264	1	1
<i>Mungala</i>	0.298	0.284	0.807	1	0.975	0.794
<i>Mwiri</i>	1	1	1	0.535		
<i>Nawul</i>	1					
<i>Ngi</i>	0.522	0.766	0.91	0.266	1	
<i>Ngodja</i>	0.851	0.22	0.672	0.064	1	
<i>Ngoye</i>	0.583	0.241	0.202	1	0.378	1
<i>Nguata</i>	1	0.494	0.664	0.491	*0.014	0.059
<i>Nyembe</i>	1	1	1	0.516	1	0.222
<i>Nzobi</i>	0.429	0.1	0.689	*0.006	0.212	0.059
<i>Nzokou</i>	1	1	1	0.962	*0.02	1
<i>Ologo</i>						
<i>Wuya</i>	1	0.147	0.304	0.547	1	1

S6 Fig

The three most frequent canonical interval codes sequences (CICS). Left: the corresponding scale beginning on 'C'. Right: the corresponding scale with less accidentals (for a better readability). The first interval of the sequence is in bold.



The figure displays two sets of musical notation, each consisting of three staves labeled I, II, and III. The left set shows scales starting on 'C' with various accidentals (flats and naturals). The right set shows the same scales with fewer accidentals for readability. Below each note in the staves are interval codes (1, 2, 3) indicating the intervals between notes. In the right set, the first interval code of each sequence is bolded.

Staff	Left Set (C-starting)	Right Set (Fewer accidentals)
I	1 2 2 1 2 2 2	2 2 1 2 2 2 1
II	1 2 2 3 2 2	2 2 1 2 2 3
III	1 2 3 2 2 2	2 2 1 2 3 2

Acknowledgments

We thank all people from Gabonese villages where we worked during last ten years, for their time and hospitality. The field work benefit logistical support from the Omar Bongo University (UOB) from Libreville and from Luto D.C. laboratory. We gratefully thank Pierre Darlu, Marc Delêtre, Susanne Fűrnis and Évelyne Heyer for their useful comments and advice.

References

1. Arom S. African polyphony and polyrhythm: Musical structure and methodology. Cambridge University Press; 1991.
2. Sallée P. L'arc et la harpe: contribution à l'histoire de la musique du Gabon. Université Paris X; 1985.
3. Le Bomin S. Musiques Bateke. Mpa Atege. Gabon. Saint-Maur-des-Fossés: Sépia; 2004.
4. Le Bomin S, Bikoma F. Musiques Myènè. De Port-Gentil à Lambaréné. Gabon. Saint-Maur-des-Fossés: Sépia; 2005.

5. Arom S, Fernando N, Fürniss S, Le Bomin S, Marandola F, Molino J. La catégorisation des patrimoines musicaux dans les sociétés de tradition orale. In: Alvarez-Pereyre F, editor. *Catégories et catégorisations: une perspective interdisciplinaire*. Peeters; 2008. p. 273–313.
6. Le Bomin S, Lechaux E, Mifune MF. Ce que “faire ensemble” peut vouloir dire en musique. *Cahiers d’ethnomusicologie*. 2008;21:175–204.
7. Blanchard N. *Identité culturelle et patrimoine immatériel: la collection sonore constituée par Herbert Pepper au Gabon (1954-1966)*. Université Lumière Lyon 2; 2011.
8. Sallée P. Un aspect de la musique des Batéké du Gabon: Le Grand Pluriarc Ngomi et sa place dans la danse Onkila: Essai d’analyse formelle d’un document de musique africaine. Libreville: ORSTOM; 1966.
9. Sallée P. *Deux études sur la musique du Gabon*. vol. 85. IRD Editions; 1978.
10. de Ruyter M. *La musique des Bongo du Gabon. Première approche [Master]*. Université de Paris X; 2003.
11. Mifune MF. *Performance et construction identitaire. Une approche interdisciplinaire du culte du bwiti chez les Fang du Gabon*. École des hautes études en sciences sociales (EHESS). Paris; 2012.
12. Lechaux E. *Tisser le fil de la mémoire. Contribution à l’histoire des répertoires musicaux des cérémonies de bwétè chez les Mitsogo du Gabon*. École des hautes études en sciences sociales (EHESS). Paris; 2015.
13. Le Bomin S, Lecointre G, Heyer E. The Evolution of Musical Diversity: The Key Role of Vertical Transmission. *PLOS ONE*. 2016;11(3):e0151570. doi:10.1371/journal.pone.0151570.
14. de Ruyter M. *Musiques et musiciens en « Pays Mobongo » : fondements musicaux et performatifs des dynamiques interethniques dans les monts du Chaillu (Gabon)*. Université de Paris X; 2016.
15. Nettl B. Revisiting Comparison, Comparative Study and Comparative Musicology. In: *Nettl’s elephant: on the history of ethnomusicology*. University of Illinois Press; 2010. p. 70–89.
16. Lomax A. *Folk Song Style and Culture*. Transaction Publishers; 1968.
17. von Hornbostel EM. *Die Probleme der vergleichenden Musikwissenschaft / The Problems of Comparative Musicology*. In: *Hornbostel Opera Omnia*. Martinus Nijhoff; 1975 [1905]. p. 247–270.
18. Nattiez JJ. Inuit Throat-Games and Siberian Throat Singing: A Comparative, Historical, and Semiological Approach. *Ethnomusicology*. 1999;43(3):399–418.

19. Savage PE. Measuring the cultural evolution of music: With case studies of British-American and Japanese folk, art, and popular music. Tokyo University of the Arts; 2016.
20. Grauer VA. Echoes of Our Forgotten Ancestors. *The World of Music*. 2006;2:5–58.
21. Fürniss S. The adoption of the circumcision ritual *bèkà* by the Baka-Pygmies in Southeast Cameroon. *African Music*. 2008;8(2):92–113.
22. Le Bomin S. Pygmies and Farmers Music in Gabon. In: *Recent Research among Congo Bassin Hunter-Gatherers and Farmers*. Albuquerque, United States; 2010.
23. Arom S. L'organisation du temps musical: essai de typologie. In: Nattiez JJ, editor. *Musiques. Une encyclopédie pour XXI^e siècle*. Actes Sud / Cité de la Musique; 2007. p. 928–944.
24. Arom S. Un synthétiseur dans la brousse. In: *La boîte à outils d'un ethnomusicologue*. Les Presses de l'université de Montréal; 2007. p. 367–378.
25. Arom S, Fernando N, Marandola F. Une méthode innovante pour l'étude des échelles musicales africaines: aspects cognitifs et techniques. In: *La boîte à outils d'un ethnomusicologue*. Les Presses de l'université de Montréal; 2007. p. 379–403.
26. Cuthbert MS. type; 2015. Available from: <http://web.mit.edu/music21/>. v2.2.1.
27. Team RC. R: A Language and Environment for Statistical Computing; 2016. Available from: <http://www.R-project.org>.
28. Eugster MJA, Schlesinger T. osmar: OpenStreetMap and R. *R Journal*. 2013;5(1):53–63.
29. Bellosta CJG. rPython: Package Allowing R to Call Python; 2015. Available from: <https://CRAN.R-project.org/package=rPython>.
30. Arom S. L'étude des échelles dans les musiques traditionnelles: une approche interactive. *Analyse Musicale*. 1991;23:21–25.
31. Fürniss S. Recherches scalaires chez les Pygmées *Aka* de Centrafrique. *Analyse Musicale*. 1991;23:31–35.
32. Arom S. Polyphonies et polyrythmies instrumentales d'Afrique Centrale: Structure et méthodologie. vol. 1 of *Ethnomusicologie*. Société d'Études Linguistiques et Anthropologiques de France (SELAF); 1985.
33. Le Bomin S. Le patrimoine musical des Banda Gbambiya (République Centrafricaine). *Catégorisation — Systématique de la musique pour xylophones*. École des hautes études en sciences sociales (EHESS); 2000.

34. Dehoux V. Chants à penser Gbaya (Centrafrique). vol. 2 of *Ethnomusicologie*. Paris: Société d'Études Linguistiques et Anthropologiques de France (SELAF); 1986.
35. Fūrniſſ-Yacoubi S. *Approche interdisciplinaire des musiques pygmées*. Université de paris x ed.; 2007.
36. Pagel M, Atkinson QD, Meade A. Frequency of Word-Use Predicts Rates of Lexical Evolution Throughout Indo-European History. *Nature*. 2007;449(7163):717–720. doi:10.1038/nature06176.
37. Arom S, Khalfa J. Une raison en acte: pensée formelle et systématique musicale dans les sociétés de tradition orale. *Revue de musicologie*. 1998;84(1):5–17.
38. Boyd R, Borgerhoff-Mulder M, Durham WH, Richerson PJ. Are Cultural Phylogenies Possible. In: Weingart P, Mitchell SD, Richerson PJ, Maasen S, editors. *Human By Nature: Between Biology and the Social Sciences*. Mahwah, New Jersey: Lawrence Erlbaum Associates; 1997. p. 355–386.
39. Nienhuys HW, Nieuwenhuizen J, Reuter J, Zedeler R. type; 2002. Available from: <http://lilypond.org>.
40. Arom S. *La boîte à outils d'un ethnomusicologue*. Les Presses de l'université de Montréal; 2007.

6.3 Conclusion et perspectives

Principaux résultats. On a pu constater dans l'analyse présentée ci-dessus l'importante diversité des valeurs des principaux **paramètres musicaux intrinsèques**, aussi bien au sein d'une qu'entre ces populations.

La métrique est relativement peu variable mais marquée par une nette organisation géographique. Ainsi la pulsation est subdivisée majoritairement de manière binaire dans l'ouest du pays et de manière ternaire dans l'est. De plus les cycles longs sont rencontrés plutôt dans le nord-est du pays, quand les cycles courts proviennent plutôt du sud.

L'organisation scalaire révèle une étonnante diversité comparé à ce qui est observé dans les autres pays d'Afrique centrale. Trois types d'échelles prédominent aussi bien au niveau du pays qu'à celui des populations : une échelle heptatonique et deux échelles hexatoniques. À côté de ces trois échelles, trouvées dans la moitié des pièces étudiées, on trouve une multitude d'échelles différentes et rares rencontrées le plus souvent dans une seule pièce.

L'organisation rythmique révèle un lien plus fort avec le répertoire que les autres paramètres. En outre le partage des **formules rythmiques** entre différentes populations reflète presque systématiquement une proximité géographique. Ceci suggère que les formules rythmiques pourraient, mieux que les autres paramètres musicaux, refléter l'histoire de ces musiques et peut être celle des populations.

Perspectives. Mais l'article présenté ne suffit pas à établir un lien entre paramètres musicaux et histoire, loin s'en faut. Si cette description minutieuse de la diversité des musiques du Gabon s'avérait une première étape nécessaire, l'inférence historique à partir de données musicales manque aussi d'une base théorique plus solide : Comment évaluer avec plus de certitude si un paramètre peut être un bon indicateur de l'histoire ? Quelles histoires la musique est-elle susceptible de révéler ? Ce sont ces questions que je chercherai à défricher dans le chapitre suivant.

Chapitre 7

Quelles données, quelle(s) histoire(s) ?

7.1	Positionnement épistémologique	249
	Sciences nomothétiques et idéographiques	249
	Instances et types	249
	Inférence historique et mécanismes évolutifs	250
	Une approche historique	250
7.2	Croisements d'histoires	251
7.2.1	Histoires musicale et génétique.	251
7.2.1.1	Histoires linguistiques et génétiques.	251
7.2.1.2	Revue détaillée de quelques études	252
	Les travaux de Grauer (2006)	252
	Les travaux de Pamjav et coll. (2012)	252
	Les travaux de Brown et coll. (2014)	253
	Les travaux de Savage et coll. (2015)	254
7.2.1.3	Que nous apprennent ces comparaisons?	255
	Corrélation gènes/musiques et qualité du marqueur historique . . .	255
	Gènes et musiques comme sources historiques complémentaires . .	258
7.2.2	Histoires musicale et culturelle	260
7.2.2.1	Différentes histoires culturelles.	260
7.2.2.2	Musiques et cultures	262
7.2.3	Histoires musicales multiples.	262
7.2.3.1	Exemples génétiques et culturels	263
	Hétérogénéité du génome	263
	Hétérogénéité culturelle	263
	En linguistique	263
	En anthropologie culturelle	263
7.2.3.2	Hétérogénéité du musical	264
7.2.4	Histoire musicale et histoire des populations	266
	Ethnie, groupe, nation, race, identité, communauté... population? .	266
	La musique, élément de définition de la population diachronique .	267
7.3	Les indicateurs de choix des marqueurs historiques	268
7.3.1	La notion de potentiel historiographique	268
7.3.2	Indicateurs <i>a priori</i>	269

7.3.2.1	Savoirs théoriques	269
7.3.2.2	Savoirs empiriques	270
	Stabilité des paramètres musicaux	270
	Corrélation avec une histoire connue	271
7.3.3	Indicateurs <i>internes</i>	272
7.3.4	Indicateurs <i>a posteriori</i>	273
7.3.5	Comparaison et limites de ces indicateurs	273
	Indicateurs <i>a posteriori</i>	273
	Indicateurs internes	273
	Indicateurs <i>a priori</i>	273
7.4	Conclusion du chapitre	274

Comme on l’a vu dans le chapitre 4, la musique porte les marques de son histoire et de celles des personnes la pratiquant, et peut à ce titre être utilisée pour étudier le passé. Mais le terme de *musique* est vague, constitué d’une multitude de paramètres distincts, décrits succinctement au cours du chapitre 3. Le mot *histoire* n’est pas plus précis et il faut définir quelle est l’histoire que l’on cherche à retrouver : celle de la musique elle-même, celle des populations qui la véhiculent, ou encore celle des individu-e-s qui la pratiquent ? Ce chapitre cherche donc à débroussailler cette question, à l’aide de l’expérience de l’ethnomusicologie, mais aussi d’autres disciplines historiographiques. Je proposerai ensuite quelques critères de sélection permettant de choisir certains paramètres musicaux plutôt que d’autres lors de l’inférence historique.

7.1 Positionnement épistémologique

Sciences nomothétiques et idéographiques. Il est nécessaire, tout d’abord, de clarifier les objectifs et les méthodes qui m’intéressent ici et pour cela de détailler quelques éléments d’épistémologie concernant les différents types de sciences. Du point de vue épistémologique, les disciplines scientifiques ont longtemps été définies, suivant Hegel, en fonction des objets étudiés [424, pp. 275–276]. Ainsi la biologie s’intéresse-t-elle à l’étude des organismes vivants, l’ethnomusicologie à l’étude de la musique, la linguistique à celle des langues. L’école néo-kantienne propose de différencier les sciences en fonction de leur but et par conséquent des moyens déployés dans ce sens [424, p. 276]. Windelband [A443]¹ différencie en 1894 des sciences dites *nomothétiques*² qui cherchent des lois générales et des sciences *idéographiques* qui décrivent des événements particuliers [424, p. 276].

Instances et types. On peut relier cette distinction aux notions d’*instance* et de *type* proposées par Peirce³. Les *instances*⁴ sont des éléments concrets et matériels rencontrés en une situation spatio-temporelle unique, alors que *types* sont des catégories abstraites n’ayant pas d’existence matérielle mais pouvant être réalisés zéro, une ou plusieurs fois [424, p. 279] sous forme d’*instance*. Ainsi la réalisation de la pièce *Igogo nuga* par un ensemble de personnes en un endroit et à un moment donné est une *instance*, alors que la *pièce* *Igogo nuga*, en tant que concept regroupant un ensemble de réalisations musicales, est un *type*. Chaque réalisation de cette pièce est une instance du type *Igogo nuga*⁵. On peut considérer que les sciences idéographiques sont celles qui s’intéressent aux événements particuliers, les instances. Elles se distinguent alors

1. ↑ Citée par Tucker [424, p. 276].

2. ↑ Du grec ancien νόμος (*nomos*), loi.

3. ↑ Tucker [424, p. 289].

4. ↑ *Token*, en anglais.

5. ↑ On peut rapprocher ces notions d’*instance* et de *type* respectivement à celles de *version de pièce* et d’*entité* proposées en ethnomusicologie par Arom et coll. [21]. Ma seule réserve, comme développé dans la partie 3.3 page 118, concerne la nature ambiguë de l’*entité* chez Arom et coll. [21] : la notion d’*entité* ne peut correspondre à celle de *type* que si elle est clairement affirmée comme une catégorie construite pour les besoins de l’analyse scientifique (conception matérialiste) et non pas comme une réalité objective qu’il s’agirait de décrire (conception essentialiste).

des sciences nomothétiques qui s'intéressent aux lois générales et se concentrent sur la notion de types. Ce changement d'angle conduit Tucker [424, p. 294] à distinguer plutôt des *sciences historiques* qui s'intéressent à la description des instances particulières et des *sciences théoriques* qui cherchent à construire des types d'événements permettant de comprendre et d'appréhender la réalité⁶.

Inférence historique et mécanismes évolutifs. La distinction entre science historique et science théorique traverse la plupart des disciplines traditionnelles. Ainsi la **biologie de l'évolution** s'attache à la fois à inférer et décrire l'histoire du vivant telle qu'elle s'est déroulée (science historique, décrivant des instances) et à décrire les mécanismes généraux expliquant l'évolution des espèces (science théorique, décrivant des types). L'ethnomusicologie cherche à modéliser la **systématique musicale** régissant la pratique musicale (et donc à formuler les types que sont les règles qui président à la réalisation musicale), mais s'intéresse aussi à décrire des instances particulières, par exemple un événement d'emprunt musical entre deux populations. De manière générale, dès lors que l'on s'intéresse à des événements passés, on peut le faire soit pour essayer de retrouver l'histoire telle qu'elle s'est réellement déroulée, soit pour comprendre certains mécanismes généraux permettant d'expliquer et de modéliser un ensemble de faits historiques. Les buts de ces deux approches, historique ou théorique, et donc leurs méthodes, sont radicalement distincts. Ces deux aspects sont pourtant abordés par les mêmes disciplines et souvent par les mêmes scientifiques. Ils peuvent même être traités conjointement et indifféremment au cours d'une même analyse. Ceci vient sans doute en partie du fait qu'on s'intéresse dans les deux cas à un même objet (par exemple l'histoire et l'évolution des lépidoptères) : les personnes spécialistes de cet objet abordent donc les différents aspects de son étude⁹. Mais cela est aussi dû au fait que les deux types d'approches interagissent, même si comme le soutient Tucker [424, pp. 291–293] il n'est pas souhaitable d'essayer de fusionner ces deux approches. En effet les résultats des sciences historiques servent à illustrer ceux des sciences théoriques ou à inspirer de nouvelles théories [424, pp. 292–293]. En retour les résultats des sciences théoriques fournissent des théories permettant de comprendre comment se transmet et se transforme l'information historique au cours du temps et d'évaluer les hypothèses sur lesquelles se fondent l'inférence historique, ces deux aspects participant à améliorer les possibilités de décrire les événements passés.

Une approche historique. L'approche poursuivie dans cette thèse est résolument historique : il s'agit de chercher à inférer des événements historiques et donc de réfléchir aux méthodes per-

6. ↑ Les notions respectives de *science théorique* et de *science historique* de Tucker ne correspondent pas exactement à celles de *science nomothétique* et de *science idéographique* de Windelband. En effet Tucker définit le but des sciences idéographiques comme « la description globale des événements »⁷ alors que celui des sciences historiques est l'« inférence de causes ou d'origines communes qui sont des instances »⁸. Comme la nuance n'est pas, selon moi, nécessaire dans le présent travail, je ne ferai pas de distinction entre sciences *idéographique* et *historique*, ni entre sciences *nomothétique* et *théorique*.

7. « the ideographic sciences aim at a comprehensive description of events » [424, p. 276].

8. « historical sciences are concerned with inferring token common causes or origins » [424, p. 294].

9. ↑ Cela est donc dû entre autres à ce que nombre d'institutions scientifiques sont organisées autour des disciplines « traditionnelles », définies par l'objet et pas toujours en fonction du but poursuivi.

mettant d'utiliser les **paramètres** musicaux pour y accéder. Le but n'est donc pas celui des *sciences théoriques* de Tucker : je ne cherche pas à identifier les mécanismes généraux de l'évolution de la musique. Je ne présente donc pas la bibliographie abondante publiée ces dernières dizaines d'années concernant les théories de l'évolution culturelle¹⁰. Inférer l'histoire nécessite cependant de comprendre comment l'information historique est transmise et modifiée au cours du temps. J'aurai donc besoin de discuter des traces historiques laissées dans le matériau musical. Je développerai pour cela, dans la partie 7¹¹, la notion de **potentiel historiographique**.

7.2 Croisements d'histoires

Une « population » humaine est souvent pensée comme un ensemble d'individus qui présente à la fois une certaine homogénéité génétique et culturelle, comparativement aux individus appartenant à d'autres populations. En supposant que l'on puisse adopter ce point de vue, la recherche historique pourrait se focaliser sur l'histoire de ces populations, les informations procurées par l'étude des données génétiques et culturelles étant censées converger en une même histoire. La réalité est bien entendu plus compliquée puisque les populations définies selon différents critères, génétiques ou culturels, ne coïncident pas nécessairement. De plus il existe des échanges à la fois génétiques et culturels, pas nécessairement liés, entre ces catégories que l'on nomme « populations ». En m'appuyant sur les apports de différentes disciplines, je tâcherai ici de réfléchir à quel type d'histoire le matériau musical peut apporter sa contribution. Je discuterai tout d'abord du lien entre histoire musicale et histoire génétique, puis entre l'histoire musicale et l'histoire issue de l'analyse d'autres données culturelles, et enfin entre les histoires racontées par différents types de données d'ordre musical.

7.2.1 Histoires musicale et génétique.

7.2.1.1 Histoires linguistiques et génétiques.

En 1988, Cavalli-Sforza et coll. [83] publient un article très influent dans lequel sont comparés un arbre issu de l'analyse phylogénétique des gènes de 42 populations du monde entier et une classification linguistique des langues parlées par celles-ci. Ce travail, bien que très critiqué¹², conforte l'idée que gènes et langues évoluent de manière corrélée et seraient donc susceptibles de révéler la même histoire. Comme l'explique Campbell [76] dans un état de l'art détaillé sur la question du lien entre ascendance linguistique et ascendance biologique, de nombreux cas de discordance entre histoire génétique et histoire linguistique ont été décrits et les cas de concordance parfaite sont plutôt l'exception que la règle.

10. ↑ On pourra se référer à la récente synthèse faite par Mesoudi [292] sur l'étude de l'évolution culturelle.

11. ↑ Pages 268–??.

12. ↑ La première de ces critiques est sans doute celle de Bateman et coll. [A36]¹³. On pourra se référer (par exemple) à la thèse de doctorat de Atkinson [26, p. 7.24] pour une revue des critiques de l'article de Cavalli-Sforza et coll. (1988) [83].

13. Citée par Nunn [320, p. 14].

7.2.1.2 Revue détaillée de quelques études

Les études mettant en parallèle données génétiques et données musicales pour étudier l'histoire des populations et de leur musiques se développent depuis la fin des années 2000. Je propose ici de présenter les principales en détail¹⁴ afin de discuter, dans la partie suivante, de ce qu'elles nous apprennent concernant l'utilisation des données musicales dans l'inférence historique.

Les travaux de Grauer (2006). Comme on l'a vu¹⁶, Grauer [165] confronte les informations ethnomusicologiques avec les apports de la génétique¹⁷ pour tenter d'établir l'histoire des musiques du monde et de montrer qu'elles suivent les migrations des peuples. Dans une autre étude¹⁹, Grauer [164] persévère dans sa démonstration d'un lien historique entre Bochimans Ju|'hoansi et Pygmées en montrant la similarité à la fois musicale et génétique de ces populations. Ces études, aussi critiquables et critiquées qu'elles soient [397] [312] [A259]²², sont les premières à s'intéresser au lien entre données génétiques et musicales pour étudier l'histoire. Cependant, pour Grauer et contrairement aux études qui suivent, il ne s'agit pas de montrer que les caractéristiques génétiques et musicales des populations sont corrélées, ce qui justifierait la possibilité d'utiliser la musique comme source historique. Grauer défend, sur des bases théoriques et indépendamment des données génétiques, que le musical peut être utilisé pour l'inférence historique [165, pp. 10–13, pp. 23–26]. Il se sert ensuite des analyses génétiques et musicales comme des arguments parallèles et indépendants justifiant sa théorie.

Les travaux de Pamjav et coll. (2012). Pamjav et coll. [334] proposent pour la première fois une étude quantifiée du lien entre diversités génétique et musicale. L'étude définit 31 cultures musicales eurasiennes²³ et regroupe pour chacune d'entre elles entre 1000 et 2500 transcriptions de pièces issues de collections de musiques populaires. Le contour mélodique de chaque transcription est représenté par un vecteur²⁴. Une méthode basée sur les réseaux artificiels de neurones est ensuite utilisée pour analyser l'ensemble de ces représentations des contours mélodiques d'une

14. ↑ À ma connaissance, le seul autre article publié sur cette thématique et cité dans des publications scientifiques¹⁵ est l'article *Music is in Our Genes* publié par Ewen Callaway dans *Nature News*. Je ne présente pas cette étude ici car ce journal est une revue de vulgarisation, l'article en question reprenant simplement quelques éléments de l'article de Grauer [165].

15. Par exemple par Brown et coll. [65, p. 5] ou Savage et coll. [370, p. 2].

16. ↑ L'analyse de Grauer [165] est présentée plus en détail dans la partie 462 page 142.

17. ↑ Il s'appuie en particulier sur les travaux de Oppenheimer [A328]¹⁸, que Grauer reconnaît lui-même comme spéculatifs [165, p. 15, p. 26].

18. Citée par Grauer [165].

19. ↑ Mais cette étude n'apporte pas beaucoup d'éléments nouveaux par rapport à la précédente [165] (si ce n'est une analyse un peu plus détaillée de la proportion de pièces caractérisées par la présence de *jodel* et d'« enchevêtrement²⁰ » dans le projet *Cantometric* [164, pp. 6–9]) puisqu'elle se fonde sur les mêmes matériaux musicaux (projet *Cantometric*) et les mêmes analyses génétiques (Oppenheimer [A328]²¹ et Chen et coll. [A90]²¹).

20. « *interlock* »

21. Citées par Grauer (2006) [165] et Grauer (2007) [164].

22. ↑ Citée par Savage et Brown [369, p. 173].

23. ↑ Les critères de définition de ce qu'est une « culture » ne sont malheureusement pas précisés.

24. ↑ Appelé *vecteur de contour* (*contour vector*, en anglais).

culture musicale. Ceci permet d'obtenir un ensemble de *types de contours mélodiques*²⁵ censé représenter au mieux cette culture musicale. Les cultures musicales sont comparées deux à deux en calculant le degré de similarités de l'ensemble de leurs types de contours respectifs²⁶. Ceci leur permet de construire une *matrice de distances* musicales.

Parallèlement les données génétiques des populations correspondant à ces cultures musicales²⁷, issues de l'ADNmt²⁸ et du *chromosome Y*, sont récupérées lorsqu'elles sont disponibles dans les bases de données génétiques. Ces séquences génétiques servent alors à calculer la distance génétique entre chaque paire de populations, et d'obtenir deux *matrices de distances* (une pour l'ADNmt et une pour le chromosome Y).

Chacune de ces trois matrices de distances (une musicale et deux génétiques) est enfin représentée sur un graphique²⁹ permettant de visualiser les populations les plus proches soit génétiquement soit musicalement. En effet, les populations séparées par des distances faibles dans une matrice se trouvent regroupées sur la carte correspondante. Pamjav et coll. [334] définissent sur chacune de ces cartes des groupes de populations proches, puis discutent la similarité des groupes formés sur les cartes génétiques et musicales.

Pamjav et coll. [334] proposent aussi une manière de définir si deux cultures musicales ont une relation génétique proche. Si au moins une des distances génétiques entre groupes appartenant à chacune de ces cultures est inférieure à 5%, ces cultures sont dites génétiquement proches. Cela leur permet de compter le nombre de paires de populations où il existe une proximité génétique, le nombre où il y a une proximité musicale et enfin le nombre où il y a une proximité à la fois génétique et musicale. Il est ainsi estimé d'une part que dans 28% des cas où il y a une proximité génétique il y a aussi une proximité musicale, d'autre part que dans 82% des cas où il y a une proximité musicale, il y a une proximité génétique. Les auteurs en concluent que la proximité musicale est un bon marqueur de la proximité génétique (mais pas l'inverse), en particulier celle basée sur l'ADNmt.

Les travaux de Brown et coll. (2014). Dans une perspective semblable, Brown et coll. [65] s'intéressent à la manière dont sont organisées et distribuées la diversité musicale, la diversité des langues et celle des gènes, puis plus particulièrement aux liens pouvant exister entre ces différents aspects. Ainsi les distances musicales, génétiques, linguistiques et géographiques sont calculées deux à deux entre 9 populations taïwanaises puis les matrices de distances obtenues sont comparées afin d'analyser le lien entre ces différents paramètres.

Pour chaque population, entre 18 et 30 pièces musicales sont décrites par 26 paramètres structurels (définis par le projet *CantoCore* [A371]³⁰) et 15 paramètres performatifs (définis par

25. ↑ *Contour type vectors*, en anglais.

26. ↑ La mesure utilisée n'est pas symétrique : aussi la distance calculée entre une culture musicale A et une culture B (d_{AB}) n'est pas égale à la distance entre B et A (d_{BA}).

27. ↑ Ne sont pas non plus donnés les critères utilisés pour choisir quelles séquences génétiques correspondent à la population génétique censée être équivalente à la culture musicale étudiée.

28. ↑ *ADN mitochondrial*.

29. ↑ Par une technique de positionnement multidimensionnel (*multidimensional scaling* ou *MDS*, en anglais).

30. ↑ Citée par Brown et coll. [65, p. 2].

le projet *Cantometric*³¹ [265]). Pour chaque paramètre, une distance est calculée entre chaque paire de pièces³². La distance globale entre deux pièces est alors définie comme la moyenne de la distance entre ces deux pièces calculée pour chaque paramètre. Le calcul de la distance globale entre chaque paire de pièce permet donc de construire une **matrice de distances** entre pièces musicales. Celle-ci est utilisée pour calculer la distance musicale entre chaque paire de populations par analyse de la variance moléculaire (AMOVA). La distance génétique entre populations est calculée par AMOVA à partir de séquences d'ADN mitochondrial et la distance linguistique est reprise d'une publication antérieure [166]³³.

Les trois matrices de distances (génétiques, musicales et linguistiques) sont analysées par la méthode de **réseau de voisins**³⁴, qui permet de tracer un réseau de relations à partir des distances et par ce moyen d'évaluer si une matrice de distances peut être représentée de manière adéquate par un arbre. La matrice de distances linguistiques est bien représentée par un arbre, mais la forme en étoile de celui-ci montre l'éloignement de toutes les langues les unes par rapport aux autres et ne révèle donc presque aucune relation de proximité. Les matrices de distances musicales et génétiques se prêtent en revanche moins bien à la représentation en arbre.

Les trois matrices de distances sont aussi comparées les unes aux autres grâce au **test de Mantel** afin de détecter une corrélation entre elles. Cependant une corrélation entre distances génétiques et musicales³⁵ peut venir non pas d'une ascendance commune mais d'un mélange culturel et génétique récent dû à la proximité géographique entre les populations. La matrice de distances géographique est ainsi utilisée pour tester la corrélation entre distances génétiques et musicales³⁵ tout en prenant en compte la proximité géographique des populations, grâce au **test partiel de Mantel**. Les distances musicale et génétique se révèlent être corrélées, tout comme les distances linguistique et génétique. Cependant les distances musicales et linguistiques ne sont pas corrélées.

Les travaux de Savage et coll. (2015). Dans la continuité de l'étude de Brown et coll. [65], Savage et coll. [370] s'intéressent à l'histoire de la population aïnoue³⁶ à partir de l'étude de leurs musiques et de leurs gènes. Les distances entre 680 pièces provenant de 35 populations du Japon, de l'est de l'Asie et de toute la région circumpolaire³⁷ sont calculées par une méthode similaire à celle utilisée dans leur précédente étude [65] et décrite ci-dessus. Grâce à l'analyse de la **matrice de distances** obtenue via une méthode de partitionnement de données^{38,39}, les

31. ↑ Ce projet est présenté dans la partie 4.2.3.2 page 134.

32. ↑ La méthode de calcul de la distance entre deux valeurs de paramètres, selon qu'il soit ordinal ou cardinal, est décrite dans l'article de Rzeszutek et coll. [363, supplément S2].

33. ↑ Citée par Brown et coll. [65, p. 3].

34. ↑ *NeighborNet*, en anglais.

35. ↑ Par exemple. Le raisonnement est le même pour les distances génétiques et linguistiques, ou linguistiques et musicales.

36. ↑ Les Aïnou-e-s vivent dans la partie nord du Japon et dans l'extrême est de la Russie.

37. ↑ La région circumpolaire est celle qui est située aux alentours d'un pôle terrestre. Il s'agit ici de toutes les terres aux abords du pôle Nord, c'est-à-dire le nord des continents eurasiatique et américain ainsi que le Nord du Groënland et les îles arctiques.

38. ↑ *Data clustering*, en anglais.

39. ↑ La méthode du partitionnement en *k*-moyennes [184] (*k-means*, en anglais).

pièces sont regroupées en cinq⁴⁰ ensembles, appelés *cantogroups*. Ces groupes ont une certaine cohérence géographique⁴³ et sont donc nommés en fonction de trois régions géographiques (Asie de l'est, circumpolaire et Aïnou⁴⁴). Par exemple les cantogroupes E1 et E2 rassemblent des pièces provenant majoritairement de l'Asie de l'est et sont donc respectivement nommés « East Asian 1 » et « 2 ».

Parmi les pièces de la population aïnoue, la proportion appartenant aux cantogroupes associés à l'Asie de l'est (7%), à la région circumpolaire (42%) et à la région aïnoue (52%) est calculée. Ces proportions sont comparées aux pourcentages de types de séquences d'ADNmt de la population aïnoue caractéristique de ces trois mêmes régions, respectivement 18%, 28% et 51%⁴⁵, calculés dans une précédente étude de génétique des populations [A406]⁴⁶. La distribution de ces pourcentages, jugés similaires entre les données musicales d'une part et les données génétiques d'autre part⁴⁷ est interprétée en termes historiques comme pouvant être due à la migration conjointe des populations et de leurs musiques, mais aussi à des échanges génétiques et musicaux récents permis par la proximité géographique.

7.2.1.3 Que nous apprennent ces comparaisons ?

Dans l'ensemble des études présentées la finalité reste d'étudier l'histoire (entre autres) à partir de la comparaison de données musicales. Cependant l'angle et le but précis sont légèrement différents d'une étude à l'autre. Je discute donc ici de ce que peuvent nous apprendre ces comparaisons entre musiques et gènes sur le potentiel des données musicales comme source d'information historique⁴⁸.

Corrélation gènes/musiques et qualité du marqueur historique. L'étude de Brown et coll. [65] prétend montrer, par l'existence d'une corrélation entre distances génétiques et distances musicales, que la musique est un marqueur utile pour étudier l'histoire :

« These results provide, to our knowledge, the first quantitative evidence that music

40. ↑ Le fait de diviser l'ensemble des pièces en exactement cinq ensembles (et pas en un nombre supérieur ou inférieur) n'est pas justifié dans l'article. Les auteur-e-s semblent présenter ce nombre comme étant un résultat de la méthode utilisée : « Cluster analysis identified five cantogroups [...] as the optimal number of clusters. »⁴¹ [370, p. 8]. Pourtant la méthode des *k*-moyennes utilisée [184]⁴² requiert de préciser explicitement le nombre *k* de groupes recherchés [184, p. 100]. Comme aucun détail n'est donné quant à l'implémentation de la méthode, on ne peut savoir s'il s'agit d'un choix arbitraire des auteur-e-s ou une spécificité de l'implémentation utilisée.

41. « L'analyse de partitionnement a identifié cinq cantogroupes [...] comme nombre optimal de groupes. »

42. Citée par Savage et coll. [370, p. 8].

43. ↑ La corrélation entre les cantogroupes et les régions qui y sont associées ne sont cependant pas testées statistiquement.

44. ↑ Le terme « Aïnou » ne renvoie pas directement à une région géographique, mais il faut comprendre ici « la région où vivent les Aïnou-e-s ».

45. ↑ Plus 4% de types de séquences caractéristiques de l'Asie du Sud-Est.

46. ↑ Citée par Savage et coll. [370, pp. 12–13].

47. ↑ La corrélation entre ces distributions n'est, là encore, pas testée statistiquement.

48. ↑ Les études analysées présentent toutes d'importants problèmes méthodologiques (sans doute inévitables dans des travaux pionniers comme ceux-ci). J'en ai présenté certains en notes de bas de page dans la partie 7.2.1.2 les présentant (page 252) et je ne reviendrai pas dessus. La discussion de cette partie se fera donc à partir des résultats présentés par ces travaux, en les supposant fiables, même s'il y a de bonnes raisons d'en douter.

might be useful as a novel marker to study human population history by demonstrating statistically significant correlations between musical and genetic diversity for nine indigenous populations of Taiwan.»⁴⁹ Brown et coll. (2014) [65, p. 4]

Dans une moindre mesure, l'étude de Pamjav et coll. [334] suggère une conclusion similaire :

« In other words, genetic relations can be predicted from the musical relations with a hit probability of 82%. »^{51,52} Pamjav et coll. (2012) [334, p. 245]

Cette formule beaucoup plus modeste attire l'attention sur le fait que la conclusion de Brown et coll. [65] est loin d'être aussi directe qu'elle ne le semble. En effet le lien entre la corrélation des distances musicales et génétiques d'une part et l'utilisation de la musique comme marqueur de l'histoire des populations de l'autre suppose plusieurs niveaux de déduction :

1. La distance génétique est un marqueur de l'histoire des populations, ce qui suppose que :
 - (a) la distance génétique est un marqueur de l'histoire des gènes ;
 - (b) l'histoire des populations peut être déduite de l'histoire des gènes.
2. La distance musicale permet de prédire la distance génétique.

Je n'aborderai pas le premier point (loin d'être trivial), car il est hors de la portée du présent travail⁵³. Le deuxième point nous rappelle bien sûr la formulation de Pamjav et coll. [334] citée ci-dessus et aborde trois questions différentes :

Peut-on faire une prédiction à partir d'une corrélation ? La corrélation entre deux variables n'est qu'une simple mesure du degré d'association des valeurs de ces variables et ne peut être directement utilisée en termes prédictifs. Certes il est possible de prédire à partir d'une analyse statistique sans pour autant comprendre les mécanismes explicatifs sous-jacents [368, p. 506]. Cependant la corrélation seule n'est qu'une description des données. Pour avoir un pouvoir prédictif (et surtout pour pouvoir l'estimer) il faut *a minima* tester la prédiction sur un jeu de données indépendant de celui sur lequel a été construit la corrélation [A96]⁵⁴.

49. ↑ « Ces résultats constituent, à notre connaissance, le premier témoignage quantitatif que la musique peut être utile comme nouveau marqueur pour étudier l'histoire des populations humaines⁵⁰ en démontrant la corrélation statistiquement significative entre les diversités musicale et génétique pour neufs populations autochtones de Taïwan. »

50. Je traduit « human population history » par « histoire des populations humaines » malgré le singulier utilisé par Brown et coll. [65] qui me paraît plutôt une tournure de la langue anglaise que la réelle référence à une population regroupant l'ensemble des êtres humains.

51. ↑ « En d'autres termes, les relations génétiques peuvent être prédites des relations musicales avec une probabilité de réussite de 82%. »

52. ↑ Cette formule me semble plutôt une manière d'exprimer leur résultat plutôt qu'une conclusion de l'article. Je la relève néanmoins ici à la fois pour illustrer mon propos et pour lever tout ambiguïté.

53. ↑ J'évoquerai la question des différentes histoires génétiques dans la partie 7.2.3.1 page 263.

54. ↑ Cité sur <https://neuroneurotic.net/2015/11/30/does-correlation-imply-prediction/> (consulté le 29 octobre 2016).

Quelle niveau de prédiction peut-on attendre de la corrélation ? En d'autres termes, il s'agit d'évaluer l'erreur commise lorsque l'on prédit la distance génétique à partir de la distance musicale. Le carré du coefficient de corrélation (r^2) est souvent interprété comme le pourcentage de variation expliqué par la variable explicative [368, p. 416]. Dans l'étude de Brown et coll. [65, p. 4], la statistique de corrélation (r) entre les distances musicales et génétiques vaut⁵⁵ 0,385, donc $r^2 = 0,148$. On peut l'interpréter ainsi : 14,8% de la variance des distances génétiques est expliquée par les distances musicales. Si ce chiffre est sans doute élevé compte tenu de la complexité du sujet d'étude, il peut faire douter quant à notre capacité à en tirer des conclusions sur l'histoire des gènes. Les possibilités de déduire l'histoire des populations sont encore plus réduites, puisque chaque étape d'inférence (1a, 1b et 2) de la liste ci-dessus est imparfaite et diminue le pouvoir inférentiel⁵⁶.

Quelle généralité peut avoir cette prédiction ? Cette prédiction selon laquelle la distance génétique peut être inférée à partir de la distance musicale a été faite dans un cas particulier. Pour pouvoir en tirer une conclusion plus générale, il faudrait tout d'abord pouvoir corroborer cette observation dans d'autres situations et vérifier que ce cas n'est pas isolé. Les différentes études liant données musicales et génétiques suggèrent une association entre ces deux variables mais les méthodes restent très différentes⁵⁷ et ne permettent pas d'en assurer la généralité pour le moment. Quand bien même l'observation aurait-elle été reproduite dans un nombre important de cas, il serait judicieux si l'on veut s'en servir dans un cas nouveau d'estimer à quel point ce cas est susceptible de sortir de la norme établie. La compréhension des mécanismes explicatifs de la corrélation pourrait alors être d'une aide précieuse. Ceci nécessite donc d'arriver à séparer les différentes causes possibles de cette corrélation entre distances génétiques et musicales et notamment entre celles évoquées par Savage et coll. [370, pp. 13–14] : évolution conjointe ou isolément par la distance.

On voit que déduire les potentialités de la musique comme marqueur historique de sa corrélation avec des données génétiques semble difficile et que la puissance prédictive de cette approche paraît faible. Brown et coll. [65, pp. 4–5] remarquent en particulier que distances musicales et génétiques sont significativement corrélées tout comme les distances linguistiques et génétiques, mais que la corrélation entre distances musicales et linguistiques n'est pas significative et l'interprètent comme le fait que « these two cultural markers might be capturing partially distinct components of human population history »⁵⁸ [65, p. 4]. Dès lors, si ces marqueurs enregistrent des histoires différentes, mais que tous deux sont corrélés aux marqueurs génétiques, il est évident que la corrélation ne peut servir à montrer que la musique ou la linguistique est un marqueur de

55. ↑ Dans le test partiel de Mantel intégrant les distances géographiques, voir la partie 563 page 253.

56. ↑ Mais passer par ces trois étapes suppose qu'inférer l'histoire des populations à partir des données musicales passe nécessairement par l'histoire génétique. On verra par la suite qu'il est peut-être plus judicieux de voir la similarité musicale, tout comme la similarité génétique, comme deux critères (pas nécessairement convergents) de *définition* de la population.

57. ↑ Et contestables, comme on l'a déjà dit.

58. ↑ « ces deux marqueurs culturels pourraient enregistrer des composantes en partie distinctes de l'histoire des populations humaines. »

l'histoire génétique. Le marqueur génétique capture donc vraisemblablement lui aussi une composante distincte de l'histoire des populations. C'est sous cet angle, plus réaliste, que les autres études présentées abordent l'utilisation conjointe des données musicales et génétiques.

Gènes et musiques comme sources historiques complémentaires. Dans les études de Grauer [164, 165], de Savage et coll. [370], mais aussi de Pamjav et coll. [334], données musicales et génétiques sont analysées séparément et les indices historiques qui en sont tirés sont finalement réunis pour discuter de l'histoire des populations étudiées. Cette approche, sans doute prometteuse, ne peut cependant faire l'économie de la question à laquelle tente de répondre l'article de Brown et coll. [65] : la musique est-elle un marqueur de l'histoire des populations ? En effet, utiliser la musique comme marqueur historique, aux côtés de la génétique, suppose :

1. Que l'histoire des populations reste inscrite au sein du matériau musical (rémanence des traces historiques) ;
2. que la méthode permette de déduire de manière relativement fiable cette histoire à partir des données musicales.

Le premier point ne pose pas de problèmes au premier abord et semble assez consensuel. En effet il paraît évident⁵⁹ que les caractéristiques musicales perdurent au moins en partie au cours du temps et que la transmission du musical soit liée au moins en partie à l'histoire des populations. Il paraîtrait même difficile de défendre que l'histoire ne se reflète jamais d'une quelconque manière dans la musique. Mais si ce principe peut-être facilement accepté c'est par son imprécision et c'est en termes de degré qu'il s'agirait de le discuter. De quel histoire parle-t-on ? Dans quelle profondeur temporelle s'inscrit-elle (rémanence des traces historiques) ? Quelle est le degré d'influence de l'histoire sur le contenu du musical (saillance des traces historiques) ? Mais ces questions de degré ne peuvent se poser indépendamment de la question de la méthode⁶⁰.

Concernant le deuxième point, de nombreuses méthodes ont été développées pour inférer l'histoire (y compris des populations) par comparaison de certaines de leurs caractéristiques⁶¹ et en particulier de leurs caractéristiques musicales⁶². La capacité de ces différentes méthodes à inférer l'histoire à partir du musical n'a pas été testée. Je m'intéresse surtout ici à la manière dont est justifiée l'utilisation de la musique comme marqueur de l'histoire des populations dans les études présentées dans la section précédente. Une première approche, que je qualifierais plutôt d'heuristique, est basée sur le sentiment que la similarité observable entre les différentes musiques étudiées ne peut être due au hasard et peut être expliquée par une histoire commune. On la retrouve dans les arguments avancés par Pamjav et coll. [334] et Savage et coll. [370] :

59. ↑ Mais méfions-nous toujours des évidences. Je n'ai pas le temps d'appuyer ce paragraphe sur des arguments plus solides et je me repose donc sur les convictions exposées ici sans plus de preuve. On retiendra que c'est un point de faiblesse de l'argumentaire qu'il s'agirait de développer et d'étayer.

60. ↑ La notion de « trace historique » ne peut de manière générale pas être dissociée des méthodes permettant de les révéler.

61. ↑ Voir chapitre 1 pages 15 à 47.

62. ↑ Voir partie 4.3 pages 139 à 143.

« One can hardly imagine a different explanation for such common musical style than the existence of a common “parent language” as an initial crystallization point of more or less independent musical evolutions. »⁶³

Pamjav et coll. (2012) [334, p. 347]

« It seems logical that these parallels could have arisen through the joint migration of people and their songs [...] but other explanations are also plausible, such as musical borrowing and genetic ‘isolation by distance’ taking place independently due to geographic proximity [...] »⁶⁴

Savage et coll. (2015) [370, pp. 13–14]

Je considère ce raisonnement comme heuristique car il n'explique pas les méthodes et l'ensemble des étapes qui permettent de lier l'observation de la similarité musicale à l'inférence historique. La deuxième approche, très proche dans la logique, se distingue cependant par la méthode. Ainsi Grauer [165, pp. 23–26] décrit en détail, dans la 8^e section de son article, le raisonnement le conduisant à proposer que la musique de la population d'*Homo sapiens* sortie d'Afrique il y a 100 000 ans aurait été proche de ce qu'il nomme le style *Pygmée/Bochimán* (P/B). La logique consiste à décrire l'ensemble des hypothèses permettant d'expliquer la présence⁶⁵ de variantes vocales et instrumentales du style P/B dans les différentes parties du monde et de discuter chacune d'entre elles puis de ne garder que la plus vraisemblable. Par exemple l'hypothèse de la nature innée de ce style est rejetée car elle n'explique pas l'absence du style P/B dans de nombreuses régions du monde. Si l'approche se veut rigoureuse, la démesure du projet ne permet pas d'être convaincu par l'hypothèse retenue (pas plus d'ailleurs que par celles exclues) tant il est difficile de démêler les liens logiques permettant de tester les hypothèses, compte tenu de nos faibles connaissances sur les modifications du musical à de larges échelles temporelles et spatiales.

La troisième approche est similaire du point de vue de la logique hypothético-déductive, mais s'en distingue par l'ambition plus limitée et par là même par une possibilité de rigueur accrue. Savage et coll. [370, pp. 13–15] étudient conjointement des données musicales et génétiques concernant les populations aïnoues et comparent deux modèles expliquant l'histoire de ces populations : un modèle dit de *double influence*⁶⁶ et un modèle dit de *triple influence*⁶⁷. D'après cette étude, les proportions des types de pièces musicales et des types de séquences d'ADNmt observées sont mieux expliquées par le modèle de triple influence que par celui de double influence. La faiblesse du modèle choisi est néanmoins reconnue au vu de l'histoire complexe des populations de cette région et le modèle constitue plus une étape de compréhension destinée à être affinée qu'une conclusion définitive [370, p. 15]. Il paraît important de multiplier de telles

63. ↑ « On peut difficilement imaginer une autre explication à une telle similarité de style musical que l'existence d'une “langue mère” commune en tant que point de cristallisation initiale d'évolutions musicales plus ou moins indépendantes. »

64. ↑ « Il semble logique que ces parallèles puissent provenir de la migration conjointe des populations et de leurs chants [...] mais d'autres explications sont aussi possibles, comme l'emprunt musical et l'“isolement par la distance” intervenus indépendamment en raison de la proximité géographique [...] »

65. ↑ Selon l'interprétation de Grauer [165].

66. ↑ *Dual structure model*, en anglais.

67. ↑ *Triple structure model*, en anglais.

approches particulières et régionales avant de s'attaquer à l'origine et à l'évolution de la musique depuis l'apparition d'*Homo sapiens*, aussi bien pour construire d'abord une compréhension fine à petite échelle que pour affûter les outils épistémologiques et méthodologiques de l'inférence historique.

7.2.2 Histoires musicale et culturelle

7.2.2.1 Différentes histoires culturelles.

Comme l'ont remarqué Boyd et coll. [61], « la » culture ne se transmet et ne se transforme pas nécessairement comme un bloc homogène, mais pourrait être vue comme plusieurs ensembles cohérents, chacun de ces blocs ayant une histoire bien distincte. Ainsi les histoires inférées à partir de l'analyse de différents types de données culturelles pourraient se révéler totalement divergentes. Boyd et coll. [61] donnent des exemples où ce pourrait être le cas, comme chez les Tiriki (Afrique de l'est) qui partagent le système de classes d'âge des Teriki, mais pas la langue, proche des langues bantoues du groupe luyia.

De nombreuses études creusent les liens entre les apports de différents types de données culturelles. Il s'agit le plus souvent d'analyses mettant en regard données linguistiques et autres données culturelles, par exemple issues de l'anthropologie de la culture matérielle. En effet les arbres de parenté linguistique⁶⁸ sont fréquemment utilisés pour modéliser l'histoire des populations⁶⁹. Un certain nombre d'études cherchent ainsi à retracer l'évolution de caractères culturels⁷⁰ le long d'arbres linguistiques. Ces analyses, basées sur des **méthodes comparatives phylogénétiquement informées**, ne s'intéressent la plupart du temps qu'à un petit nombre de caractères culturels, mais certaines d'entre elles comparent des ensembles culturels plus larges.

Ainsi Jordan et Shennan [217] s'intéressent aux techniques de vanneries, aux langues, aux zones écologiques et aux distances entre différentes populations amérindiennes de Californie. Plusieurs types d'analyses (**tests de Mantel**, analyse factorielle des correspondances, comparaison d'arbres) permettent de comparer (entre autres) la structuration des techniques de vanneries et des langues au sein de ces populations. Tout en révélant la complexité de l'ensemble de données étudié, ils constatent⁷¹ que la proximité linguistique n'explique que peu la similarité des techniques de vannerie, au moins à grande échelle.

Jordan et Mace [215] s'attaquent directement à la question soulevée par Boyd et coll. [61] au sujet de la cohérence historique des ensembles culturels en étudiant les cultures de 17 populations amérindiennes de la côte nord-ouest des États-Unis. Constatant par une analyse **cladistique** que le jeu de données est mal expliqué par une structure **arborescente**, ils le divisent en six sous-jeux

68. ↑ Voir partie 2.1.3 page 61.

69. ↑ Voir par exemple les analyses de Gray et coll. [167, p. 3929] ou de Calude et Verkerk [73].

70. ↑ Par exemple le système de descendance [A197, A212, A213, A327], le système de résidence [A327], les pratiques maritales [147], la complexité sociale [A433], le pastoralisme [A197] ou encore la digestion du lactose [A198].

71. ↑ Bien que les résultats dépendent largement de la classification linguistique utilisée.

de données⁷² qu'ils soumettent de nouveau à une analyse cladistique. Ils construisent aussi un arbre linguistique à partir d'une classification existante et un arbre de proximité géographique à partir des distances à vol d'oiseau. La similarité entre la topologie de ces arbres est évaluée par le *test de Kishino-Hasegawa*, ainsi que par mesure de la distance des triplets. Ces analyses montrent d'une part qu'il n'y a pas d'association entre les arbres culturels et linguistiques, d'autre part que certains groupes de données culturelles (mais pas tous) montrent une corrélation dans la structure de leurs arbres. Ceci suggère l'existence d'un noyau de transmission culturelle dont les contours sont assez flous. Mais il semble aussi que certains ensembles culturels soient transmis de manière plus cohérente, comme celui constitué des structures des bâtiments, des pratiques maritales et des vêtements et bijoux. Jordan et O'Neill [216] poursuivent cette analyse en étudiant plus en détail l'évolution conjointe d'un certain type d'habitation et des langues dans cette région. Ils montrent que la structure des données architecturales est contrainte par l'arbre linguistique, même si les topologies des deux types d'arbres est largement différente.

Graça da Silva et Tehrani [162] s'intéressent à l'ancienneté de 275 contes collectés dans 50 populations locutrices de langues indo-européennes. L'étude s'appuie sur un arbre des langues indo-européennes pour tester si la transmission de chaque conte est majoritairement verticale, c'est-à-dire qu'elle suit majoritairement la structure de cet arbre. Comme la proximité spatiale peut biaiser l'analyse, un test complémentaire permet de vérifier l'existence d'un signal vertical en prenant en compte l'impact de la proximité géographique. Le signal vertical se révèle significatif pour 100 contes, parmi lesquels 76 semblent montrer une transmission plus fortement verticale qu'horizontale. La présence ou l'absence de chacun de ces 76 contes aux différents nœuds ancestraux est alors inférée par analyse bayésienne. On voit là encore que si la transmission des contes n'est pas complètement indépendante de l'histoire linguistique, elle ne lui correspond pas non plus complètement, seule une minorité de contes montrant une adéquation forte avec l'arbre linguistique.

Ces quelques exemples montrent que l'histoire de différents éléments culturels (langues, culture matérielle, traditions orales, ...) ne converge pas nécessairement en une histoire unique et cohérente. Bien que Boyd et coll. [61] envisagent la possibilité théorique d'une culture formant un tout intégré dont la transmission se ferait de manière parfaitement cohérente, ils remarquent néanmoins :

« [...] there is such overwhelming evidence for substantial diffusion and for rapid evolution in many components of culture that we believe it unlikely that any tenable empirical defense of a completely holistic cultures-as-species position can be offered. »⁷³

Boyd et coll. (1997) [61]

72. ↑ Structure des bâtiments, vêtements et bijoux, instruments de musique, *potlachs*, pratiques maritales et rituels féminins.

73. ↑ « [...] il y a des preuves si écrasantes de la présence d'une diffusion substantielle et d'une évolution rapide dans de nombreuses composantes de la culture que nous pensons improbable qu'une quelconque défense, empirique et soutenable, d'une position complètement holistique de la culture-espèce puisse être proposée. »

7.2.2.2 Musiques et cultures

Les informations historiques que peuvent nous apporter la musique n'ont que rarement été confrontées à celles issues d'autres données culturelles. Les deux études qui, à ma connaissance, le font ont déjà été présentées et je ne ferai que les rappeler. Ainsi on a déjà vu que Brown et coll. [65] ne trouvent pas de corrélation entre les matrices de distances musicales et linguistiques chez les populations taïwanaises. Dans leur recherche de paquets culturels cohérents, Jordan et Mace [215] s'intéressent aussi aux instruments de musique des amérindiens de la côte nord-ouest des États-Unis d'Amérique. Comme pour les 5 autres types de données culturelles, ils effectuent une analyse cladistique à partir des données concernant ces instruments⁷⁴. Il est d'ailleurs intéressant de noter que l'arbre obtenu est celui dont l'indice de cohérence est le plus élevé parmi les différents types d'éléments culturels, bien que sa valeur (0,57) soit relativement faible⁷⁵. Comme les autres types d'éléments culturels, le test de Kishino-Hasegawa ne permet pas de trouver de lien étroit entre la structure de l'arbre construit à partir des instruments et n'importe lequel des autres arbres (linguistique, géographique ou culturel). Néanmoins, la distance entre l'arbre des instruments et celui des rites féminins est significativement inférieure à celle mesurée entre des arbres aléatoires⁷⁶ et elle est relativement faible avec chacun des arbres des autres données culturelles⁷⁷. L'arbre des instruments n'est cependant pas significativement similaire à l'arbre des distances et des langues⁷⁸.

Ces études ne révèlent donc pas une association forte des données musicales avec certains types de données culturelles (linguistiques ou autres), même si il existe parfois des liens significatifs. Bien que le faible nombre d'études sur le sujet ne permet pas d'établir de conclusion, il paraît raisonnable (au vu de ces travaux préliminaires, mais aussi de ce qui a été fait dans d'autres domaines de l'anthropologie culturelle) de suggérer que la musique peut avoir une histoire largement indépendante à la fois des gènes des personnes qui les réalisent et des autres pratiques culturelles, tout en n'étant probablement indépendante ni de l'une ni de l'autre. Cette conclusion nécessiterait d'être étudiée plus en détail en ce qui concerne les degrés d'indépendance et de liaison existants entre histoires musicales, culturelles et génétiques.

7.2.3 Histoires musicales multiples.

Ainsi histoires génétiques et culturelles peuvent diverger, tout comme les histoires issues de différents types d'éléments culturels. Aussi l'histoire de la musique n'est pas nécessairement celle des gènes ni celle d'autres éléments culturels. Mais que signifie *l'histoire de la musique*? Loin

74. ↑ Malheureusement la description des caractères et la procédure utilisée ne sont pas détaillées et il est donc difficile d'en prendre toutes les conséquences.

75. ↑ La moyenne observée par Collard et coll. [94] pour cet indice est de 0,61 dans les jeux de données biologiques et de 0,59 dans les jeux de données culturels. La valeur trouvée ici est cependant bien plus élevée que celles obtenues par Temkin et Eldredge [413, p. 147, p. 149] dans une analyse similaire (0,31 et 0,4).

76. ↑ La distance de triplets mesurée entre ces deux arbres est de 79, alors que la distance minimale entre les arbres aléatoires est de 174.

77. ↑ Les distances entre l'arbre des instruments et celui 1. des vêtements et bijoux, 2. des structures de bâtiments, 3. des pratiques maritales et 4. des *potlatches* sont respectivement de 110, 130, 131, 159.

78. ↑ Distances respectivement de 248 et 240.

d'être un tout homogène, « la » musique est composée d'éléments disparates dont l'étude peut elle aussi révéler des histoires distinctes. Pour l'illustrer, je présenterai tout d'abord quelques situations venues de la biologie et de l'anthropologie culturelle, avant de discuter du cas musical.

7.2.3.1 Exemples génétiques et culturels

Hétérogénéité du génome. De la même manière que la notion de *population*, la notion de *génom* en tant qu'unité cohérente du point de vue historique a été remise en cause depuis quelques dizaines d'années. Ainsi différentes parties d'un génome n'ont pas nécessairement la même histoire. Le cas le plus emblématique est peut-être celui des mitochondries, organites présents dans les cellules des eucaryotes⁷⁹ et issues de l'inclusion (appelée *endosymbiose*) d'une bactérie⁸⁰ dans une cellule eucaryote il y a environ 2 milliards d'années. Ainsi la cellule eucaryote et la mitochondrie ont conservé leur génome propre lesquels ont évolué indépendamment. Si l'histoire des mitochondries et des eucaryotes est intimement liée depuis l'endosymbiose originelle, leurs histoires antérieures sont distinctes et l'inférence phylogénétique à partir de l'ADNmt rapproche les mitochondries des α -protéobactéries. De la même manière, tous les gènes d'un même génome n'ont pas nécessairement la même histoire, certains ayant pu être dupliqués ou acquis par transfert horizontal⁸¹. Le génome peut donc être vu comme un ensemble hétérogène pouvant révéler différentes histoires⁸².

Hétérogénéité culturelle.

En linguistique. De la même manière un domaine de la culture ne peut être supposé *a priori* comme un ensemble historiquement homogène. En ce qui concerne la linguistique, Rexová et coll. [346] effectuent une analyse cladistique des langues bantoues à partir de deux types de caractères linguistiques : lexicaux et grammaticaux. Même si l'analyse combine ces deux types dans un seul jeu de données et que l'on ne peut donc comparer leurs apports respectifs, les auteurs notent qu'il existe une incongruence significative entre les deux catégories de données [346, p. 191]. Dans deux groupes de langues (indoeuropéennes et austronésiennes) Greenhill et coll. [A171]⁸³ comparent les réseaux obtenus par réseau de voisins à partir de deux types de données linguistiques : typologiques d'une part et lexicales d'autre part. Le réseau typologique se révèle bien moins arborescent que le réseau lexical, suggérant que les caractéristiques typologiques seraient plus sujettes à la transmission horizontale que les caractéristiques lexicales.

79. ↑ Les eucaryotes sont un ensemble d'êtres vivant caractérisés notamment par la présence d'un noyau et de mitochondries dans leurs cellules.

80. ↑ Une α -protéobactérie.

81. ↑ Page et Charleston [330, p. 386].

82. ↑ Boyd et coll. [61].

83. ↑ Citée par Gray et coll. [167, p. 3928].

En anthropologie culturelle. En ce qui concerne l'anthropologie culturelle, Jordan et Mace [215]⁸⁴ comparent les informations obtenues par analyse cladistique à partir de six types de données différentes, trois issues de la culture matérielle et trois de la culture sociale. Même si on constate une association entre certains types de données, les topologies des différents arbres obtenus sont largement divergentes.

Rogers et coll. [358]⁸⁵ cherchent à inférer les relations entre des populations du Pacifique à partir de l'étude des caractères de leurs canoës, divisés en deux types : caractères fonctionnels et caractères symboliques⁸⁶. Gray et coll. [167, pp. 3929–3931] poursuivent cette étude en construisant un arbre des langues des populations concernées. Ils cherchent ensuite à savoir à quel point chacun des deux types de caractères des canoës sont cohérents avec cet arbre linguistique. Ils constatent qu'aucun des deux ne s'accorde bien avec l'arbre, mais que les caractères fonctionnels s'y accordent encore plus mal que les caractères symboliques, suggérant que ces deux types de caractères pourraient subir des influences différentes.

Au cours de son étude des psaltériens baltiques, Temkin [412]⁸⁷ analyse deux catégories de caractères : structurels et décoratifs, opérant ainsi une distinction semblable à celle de Rogers et coll. [358]⁸⁵. Il remarque qu'en augmentant le poids attribué aux caractères structurels la topologie obtenue est inchangée, alors qu'en augmentant celui des caractères décoratifs la topologie n'est presque pas résolue. Les éléments décoratifs seraient ainsi plus affectés par les emprunts et les innovations alors que les éléments fonctionnels seraient plus stables au cours du temps [412, p. 227].

7.2.3.2 Hétérogénéité du musical

Constatant cette hétérogénéité dans d'autres domaines de l'anthropologie culturelle, on ne peut que s'interroger sur la cohérence du musical et donc sur l'unicité des traces historiques qu'il porte. La musique est un concept polysémique, largement utilisé hors du seul domaine scientifique. Pour cette raison, l'acceptation de ce terme en ethnomusicologie est en partie héritée de l'organisation sociale du sonore dans les sociétés occidentales. L'influence des définitions du musical ne provenant pas du milieu scientifique persiste dans le domaine scientifique, notamment à travers la manière dont sont définies les institutions de recherche. Fürniss [151, p. 11] réfléchit à ce que comprend la notion de musique :

«La musique» est une notion vaste qui recouvre des domaines variés mais complémentaires, voire concomitants. C'est, tout d'abord, la systématique musicale, c'est-à-dire le langage musical avec ses règles qui définissent l'agencement des sons dans le temps. La mise en sons est effectuée soit par des instruments de musique, soit par la voix, et pour ces deux types de sources sonores, il existe plusieurs techniques de réalisation. La matérialité des instruments diffère en termes de morphologie, de facture et

84. ↑ Cette analyse a déjà été présentée dans la partie 602 page 260.

85. ↑ Citée par Gray et coll. [167, p. 3929].

86. ↑ Voir aussi la partie 214 page 69.

87. ↑ Ces travaux sont présentés plus en détail dans la partie 2.1.5.1 page 70.

de matières utilisées, qui sont autant de paramètres observables ayant une incidence directe sur le son. Enfin, la musique n'existe pas sans sa réalisation en un lieu et à un moment précis. Associée ou non à la danse et à une mise en espace particulière, cette réalisation lie étroitement la notion de musique à celle de la performance. »

Fürniss (2014) [151, p. 11]

Ainsi « la musique » d'une population peut désigner tout ou partie des éléments suivants⁸⁸ :

- Les éléments matériels permettant la réalisation musicale (instruments) ;
- Les techniques vocales et instrumentales utilisées au cours de la réalisation ;
- Le contexte social dans lequel intervient la réalisation musicale ;
- La perception et les effets de la réalisation chez les différentes personnes concernées par celle-ci ;
- Le phénomène sonore réalisé ;
- Les caractéristiques musicologiques issues de l'analyse du matériel sonore (caractéristiques métriques, rythmiques, mélodiques, procédés polyphoniques, formes responsoriales, ...) ;
- La **systématique musicale** de chacun des éléments ci-dessus.

Peut-on supposer que tous ces éléments se perpétuent de manière parallèle au cours du temps ? Si c'est le cas, étudier l'histoire de n'importe quel sous-groupe de ces éléments reviendrait à étudier l'histoire de « la musique ». C'est l'idée sous-jacente à l'ensemble du travail de Grauer [165] lorsqu'il propose de reconstruire « the early music »⁸⁹ [312, pp. 62–66]. Même en supposant qu'il puisse être pertinent et réaliste, compte tenu de nos connaissances sur les mécanismes de transmission et de transformation des musiques, de reconstruire des caractéristiques musicales à une époque si lointaine, rechercher « une » musique originelle paraît de toute façon illusoire :

« I think we are probably safest to consider the origins of music as a series of separate events, derived from a variety of social needs and processes, united only by their use of sound distinguished from ordinary speech. »⁹⁰ Nettl (2006) [312, p. 66]

De nombreux exemples nous apprennent que différents éléments du musical peuvent être transmis différemment et donc avoir des histoires différentes. Ainsi Savage et Brown [369, p. 173]⁹¹ prennent l'exemple des musiques d'Océanie montrant simultanément des caractéristiques de musiques populaires contemporaines, d'hymnes chrétiens et de musiques préexistantes à l'arrivée

88. ↑ Cette liste ne cherche pas l'exhaustivité ni la rigueur catégorielle, mais vise simplement à illustrer la diversité des éléments rattachés de près ou de loin à la notion de « musique ».

89. ↑ « la musique ancienne »

90. ↑ « Je pense que le plus sûr est probablement de considérer les origines de la musique comme une série d'événements séparés, issus d'une variété de besoins et de processus sociaux, unis seulement par une utilisation du sonore différente du parlé ordinaire. »

91. ↑ Citant l'encyclopédie de Nettl et coll. [A315].

des occidentaux. Les méthodes d'**inférence historique** présentées dans le chapitre 2 sont cependant basées pour la plupart sur l'analyse statistique des données. Elles n'ambitionnent pas de démêler l'écheveau complexe des histoires imbriquées, mais de découvrir les trames historiques les plus grossières par la découverte d'un nombre suffisant de données congruentes ayant suivi cette trame. Ces études, montrant la divergence ponctuelle de certains éléments musicaux, n'éliminent donc pas complètement la possibilité d'intégrer l'ensemble des données musicales lors de telles inférences. L'hétérogénéité du concept de *musique* oblige cependant à se demander si différents types de données musicales peuvent être historiquement congruents.

Il n'y a pas, à ma connaissance, de comparaison systématique des apports historiques de différents types de paramètres musicaux⁹², mais la question est abordée dans le travail de Le Bomin et coll. [240]⁹⁴. Leur analyse cladistique des patrimoines musicaux du Gabon s'appuie sur trois types de paramètres musicaux différents : des répertoires, des paramètres performatifs et des paramètres intrinsèques. L'étude détaillée des indices de **cohérence** (IC) et de **réention** (IR) montre une forte hétérogénéité selon le type de données : l'IC moyen des paramètres intrinsèques est de 0,889 alors que celui de la plupart des autres types de paramètres est compris entre 0,4 et 0,6 et celui des types polyphoniques est de 0,2 [240, pp. 11–12]. De plus l'étude des trois types de caractères par des analyses cladistiques séparées produit des arbres différents, ne montrant pas la séparation entre populations matrilineaires et patrilineaires [240, p. 11]. Ces résultats suggèrent d'une part que différents types de caractères musicaux peuvent avoir une histoire distincte, mais aussi que les paramètres internes semblent avoir une transmission cohérente, correspondant peut-être à la notion de *paquet culturel* de Boyd et coll. [61].

7.2.4 Histoire musicale et histoire des populations

Il peut être tentant, comme le fait par exemple Grauer [165], d'utiliser les données musicales pour mieux comprendre l'histoire des populations actuelles. Encore faut-il bien s'entendre sur la signification du mot *population*. Il convient, à mon avis, de se garder d'une vision essentialiste de cette notion, c'est-à-dire de ne pas considérer la population comme une réalité matérielle et objective mais bien comme un concept scientifique construit pour mieux appréhender la réalité. Je propose donc ici de considérer la musique comme élément contribuant à définir une population au sens *diachronique*⁹⁵ du terme.

Ethnie, groupe, nation, race, identité, communauté... population ? Les termes pour désigner des ensembles d'êtres humains partageant un certain nombre de caractéristiques, bio-

92. ↑ Je rattache plutôt l'exemple des instruments de musique étudiés par Temkin [412] donné ci-dessus⁹³ à l'anthropologie de la culture matérielle, bien que le cas soit intermédiaire.

93. Voir paragraphe *En anthropologie culturelle* page 263.

94. ↑ Voir *partie 2.1.5.3* page 73.

95. ↑ C'est-à-dire une population définie pour être pertinente non pas nécessairement à un moment donnée, mais entre différentes époques.

logiques ou culturelles, sont nombreux et leurs définitions sont loin d'être clairement établies⁹⁶. Le terme d'*ethnie* en particulier a joué un rôle central en anthropologie, notamment à travers la discipline à laquelle il a donné son nom. Introduit à la fin du XIX^{ème} siècle, il est d'abord utilisé pour définir et catégoriser les êtres humains selon des critères objectifs, dans une perspective largement essentialiste : il existerait des ethnies identifiables par des caractéristiques biologiques (en particulier anatomiques) et culturelles (linguistiques, musicales, ...), chaque être humain appartenant à une et une seule d'entre elles. Cette perspective a été complètement renversée à la fin du XX^{ème} siècle dans une optique postmoderniste et postcoloniale, en particulier par les travaux de Barth [A35]⁹⁷ et Amselle et M'Bokolo [A5]⁹⁷ : l'ethnie devient une catégorie d'inscription caractérisée par le maintien d'une frontière perméable et constamment renouvelée. C'est une entité construite et labile, issue d'une histoire particulière, rompant la logique selon laquelle à une ethnie correspondrait une langue, un ethnonyme et une origine biologique. On substitue aussi à la notion d'identité celle d'*identité plurielle*, chaque personne se rattachant de par son histoire à différents groupes culturels.

Le concept d'*ethnie* a alors parfois été remplacé par celui de *population*, par exemple dans les travaux de Le Bomin [239–243] et dans le présent travail. On peut penser que cette utilisation provient de son acceptation en biologie (en particulier en génétique des populations), elle-même issue de sa signification statistique : « toute collection d'objets à étudier ayant des propriétés communes » [368, p. xxvi]. En biologie la pensée populationnelle s'oppose résolument à la pensée *typologique* et donc à toute vision essentialiste [282] : la population est un concept scientifique construit *a posteriori* pour décrire et expliquer nos observations et en aucun cas une catégorie transcendante ayant une existence propre.

La musique, élément de définition de la population diachronique. Cependant, dès lors que le nombre de propriétés étudiées augmente, la chance de voir toutes les propriétés définir les mêmes groupes s'amenuise, chaque personne s'inscrivant dans une identité plurielle. Ceci est d'autant plus vrai quand on s'intéresse à la dimension historique. Si l'on définit un groupe caractérisé aujourd'hui par un certain nombre de propriétés, pourra-t-on retrouver ces mêmes propriétés dans un groupe passé ? Cela aura-t-il un sens ? Comme on l'a vu dans les parties précédentes⁹⁸, les histoires de différentes caractéristiques (biologiques ou culturelles) associées aux êtres humains divergent et se croisent.

Aussi, on ne peut supposer *a priori* qu'une population humaine définie de manière *synchrone* aura une pertinence dans une dimension *diachronique*. À l'inverse, on pourrait décrire la transmission et l'évolution de différentes caractéristiques (gènes, langues, rythmes, ...) observées. Si l'on observe des associations au cours de ces différentes histoires, alors seulement pourra-t-on définir une *population diachronique*. Dans cette perspective, il ne s'agit plus d'utiliser la musique

96. ↑ Ce paragraphe est issu en grande partie de ma compréhension de l'exposé sur ce thème présenté le 8 décembre 2016 par Jonhatan Benabou au musée de l'homme à l'occasion de l'*atelier terrain-théorie en anthropologie* (ATTA).

97. ↑ Citées par Jonhatan Benabou le 8 décembre 2016 (voir note 96).

98. ↑ Voir parties 7.2.1 page 251, 7.2.2 page 260 et 7.2.3 page 262.

pour étudier l'histoire de populations supposées avoir une pertinence historique, mais d'étudier l'histoire de différents paramètres, musicaux ou non (culturels, biologiques, ...) dont la congruence pourrait servir à *définir* une population.

7.3 Les indicateurs de choix des marqueurs historiques

Comme on l'a vu, l'inférence historique par comparaison des données d'ordre musical est complexe car différentes informations musicales peuvent révéler différentes histoires et ne doivent pas nécessairement être utilisées ensemble, certaines pouvant même ne pas être pertinentes à inclure dans l'analyse. Le peu de connaissance de mécanismes généraux (ou même valables dans certains cas particuliers) concernant les processus de transmission et de modification du musical fournit peu d'appui théorique sur lequel fonder l'analyse. De plus la divergence au moins partielle entre le musical et d'autres sources d'information historique (génétique, linguistique, ...) rend difficile l'évaluation de la qualité des données musicales comme marqueur historique⁹⁹.

Cette partie vise donc à partager quelques réflexions sur la manière d'évaluer à quel point un certain type de données musicales est utilisable pour inférer l'histoire : la comparaison des échelles musicales ou des cellules rythmiques peut-elle apporter des connaissances sur le passé ? Si oui, apportent-elles les mêmes connaissances ? Devraient-elles être utilisées conjointement ou séparément ? J'apporterai quelques outils pour y réfléchir. Après avoir expliqué ce que j'entends par *potentiel historiographique*, je présenterai les différents types d'arguments qui peuvent servir à discuter de la valeur historiographique d'un type de données et leur valeur épistémologique.

7.3.1 La notion de potentiel historiographique

J'appelle ici **historiographie** la représentation de l'**histoire** construite de manière scientifique, dans une démarche rationnelle. Suivant cet usage, j'utilise le terme de **potentiel historiographique** d'un jeu de données pour désigner les possibilités d'inférence historiques offerte par celles-ci. Il permet donc d'évaluer à quel point un ensemble de données peut nous renseigner sur l'histoire.

Cette notion n'est pas une mesure quantifiable, mais plutôt une qualité que l'on peut discuter. En effet, pour un même ensemble de données, ce potentiel peut être différent selon l'histoire à laquelle on s'intéresse ou l'échelle de l'analyse. En biologie, par exemple, le gène codant pour l'ARN ribosomique 18S évolue très lentement, aussi sera-t-il parfaitement identique dans toutes les espèces proches et ne permettra donc pas d'inférer les relations de parentés entre celles-ci. Mais cette faible vitesse d'évolution et sa fonction vitale fait aussi qu'il est présent dans l'ensemble des eucaryotes¹⁰⁰ et permettra donc de comparer des **taxons** très éloignés et qui partagent généralement peu de caractères permettant de les comparer. Ce gène a donc un potentiel historiographique faible ou nul pour l'histoire récente, mais élevé pour l'histoire ancienne.

99. ↑ Voir partie 580 page 255.

100. ↑ Et même chez les eubactéries et les archées, si l'on considère que l'ARN ribosomique 16S de ces taxons lui est **homologue**. Cet gène est donc présent chez l'ensemble des êtres vivants, excepté les virus.

Je parle de *potentiel* car rien ne permet de savoir avec certitude que cet indice se révélera pertinent pour l'analyse projetée. Il permet d'estimer *a priori* l'utilité de tel ou tel type de données, mais cette estimation provenant nécessairement d'un contexte différent, elle peut se révéler inadéquate. De plus, la notion d'*information historique* ne peut être séparée des méthodes permettant d'utiliser cette information : l'information est contenue tout autant dans les données analysées que dans la méthode utilisée. Ce n'est pas une propriété matérielle et objective des données, mais un outil scientifique permettant de modéliser le processus de connaissance historique. Aussi l'information disponible est-elle dépendante de la méthode utilisée. Cependant différentes méthodes, même très distinctes, appliquées aux mêmes données peuvent donner des conclusions historiques du même ordre bien que le détail diffère et c'est cette partie imputable aux données que le potentiel historiographique cherche à capter. Il n'en reste pas moins qu'il dépend de la méthode utilisée.

Le but de cette partie est de donner des pistes pour des recherches futures qui puissent servir les objectifs de l'inférence historique à partir de données musicales. Celle-ci étant limitée par notre faible connaissances des processus de transmission et de transformation des données musicales, les recherches sur ces sujets pourraient être d'un grand apport. Encore faut-il qu'elles aient été réfléchies dans cette perspective afin d'apporter des connaissances aussi utiles que possible de ce point de vue. Une manière d'aborder la question est de chercher à mieux connaître le potentiel historiographique de différents types de données musicales, mais celui-ci me semble, dans l'état actuel des connaissances, difficile à évaluer. Il s'agit donc en premier lieu de conduire des études permettant de nous renseigner sur le potentiel historiographique de ces données.

Je discute donc dans les parties qui viennent de différents indices permettant d'évaluer le potentiel historiographique d'un ensemble de données. Comme on le verra ces indices n'ont pas tous le même statut épistémologique : certains ne permettent qu'une estimation incertaine quand d'autres fournissent une évaluation relativement robuste. Ils n'ont pas non plus la même accessibilité empirique : certains indices requièrent beaucoup plus de données de terrain et de travail d'analyse que d'autres. Je distinguerai trois types d'indicateurs permettant d'évaluer le potentiel historiographique. Les indicateurs *a priori* (théoriques ou empiriques) proviennent des connaissances précédemment établies sur les musiques et les processus historiques auxquels elles sont soumises. Les indicateurs *internes* sont ceux qui, pour estimer le potentiel historiographique d'un jeu de données, proviennent du jeu de données lui-même. Les indicateurs *a posteriori* sont ceux disponibles seulement une fois l'inférence historique effectuée.

7.3.2 Indicateurs *a priori*

7.3.2.1 Savoirs théoriques

Une méthode pour estimer le potentiel historiographique d'un ensemble de données est de se fonder sur le savoir théorique établi sur la transmission et la transformation des musiques. Dans une étude sur les cultures circumpolaires, Nattiez [304] compare les jeux et chants de gorge

dans trois cultures. Il montre qu'elles ont vraisemblablement une origine commune¹⁰¹ mais aussi que la technique vocale est mieux préservée au cours du temps que la signification sociale de ces pratiques. Il propose ainsi une théorie plus générale concernant la transmission :

« From this situation, we may draw broader conclusions of interest for general musicology and semiology. In sonorous symbolic forms, the form, the signifier, best resists transformations through time. However, the signified, the religious signification of the animal and nature imitations associated with these forms, are evanescent. »¹⁰²

Nattiez (1999) [304, p. 414]

Cette théorie proposée par Nattiez à partir d'un cas particulier nécessiterait bien entendu d'être étayée et testée sur des cas particuliers. Mais à supposer que l'on puisse avoir suffisamment de confiance dans l'efficacité et la généralité du mécanisme avancé, celui-ci pourrait être utilisé pour évaluer le potentiel historiographique de différents types de données. En prenant l'exemple des musiques du Gabon, on pourrait alors considérer que les caractéristiques symboliques associées à un répertoire (par exemple l'*Elengé*, un répertoire de possession des femmes chez les Myènè [238, p. 63]) sont préservées moins longtemps que les caractéristiques formelles (par exemple la *formule rythmique* caractéristique de ce répertoire [238, p. 67]). Les éléments formels seraient donc supposés pouvoir révéler une histoire plus ancienne que les éléments symboliques^{103,104}.

On pourrait prendre comme autre exemple la théorie présentée par Le Bomin [234] dans sa thèse de doctorat¹⁰⁵. Comme on l'a vu, elle propose que certaines références mentales individuelles (c'est-à-dire propres à un musicien) puisse être adoptées plus généralement et devenir des références mentales collectives (partagées par toute une population). On peut y voir la proposition d'un mécanisme d'évolution des références mentales, où un énoncé individuel est « cristallisé » [234, p. 278] en un énoncé collectif. Si elle était confirmée, au moins au niveau local, une telle théorie pourrait là encore servir à informer sur le potentiel historiographique du musical.

7.3.2.2 Savoirs empiriques

Stabilité des paramètres musicaux. Au cours de leur étude des musiques Myènè, Le Bomin et Bikoma [238] s'intéressent à trois pièces (*Ologo ma me*, *Ologo mbela* et *Mbumba osaka*), chacune trouvée dans au moins deux des populations étudiées (Orungu, Galoa et Nkomi). La métrique et l'échelle musicale utilisées dans chacune des versions de pièce analysées sont présentées dans

101. ↑ Voir la partie 458 page 141.

102. ↑ « À partir de ce cas, on pourrait tirer des conclusions plus larges utiles à la musicologie et à la sémiologie générales. Dans les formes symboliques du sonore, la forme, le signifiant, résiste le mieux aux transformations au cours du temps. En revanche, le signifié, la signification religieuse des imitations des animaux et de la nature associées à ces formes, sont évanescences. »

103. ↑ Ce lien entre pérennité des traces et profondeur historique potentielle de l'inférence évoque l'exemple de l'ADN mitochondrial présenté dans la partie 7.2.3.1 page 263.

104. ↑ La théorie de Nattiez, proposant un lien entre fonction et pérennité, peut rappeler un concept de biologie de l'évolution utilisé au cours de l'inférence historique. En effet les éléments fonctionnels de l'ADN (parties codantes pour un gène par exemple) sont considérées être plus stables que les parties « non fonctionnelles » (portions non codantes et sans fonction connue). On note cependant que le rapport est inversé.

105. ↑ Voir partie 3.3.1.4 page 122 ainsi que la citation C25 page 384.

Pièce Paramètre	<i>Ologo ma me</i>		<i>Ologo mbela</i>		<i>Mbumba osaka</i>	
	Métricité	Échelle	Métricité	Échelle	Métricité	Échelle
Galoa	8/2	heptatonique	8/3	hexatonique	8/2	pentatonique
Orungu			8/3	pentatonique		
Nkomi	8/2	pentatonique			8/2	heptatonique

TAB. 7.1 : **Métriques et échelles musicales de trois pièces dans différentes populations.** Chacune de ces trois pièces a été trouvée dans deux populations Myènè du Gabon. Leurs métriques et leurs échelles sont celles identifiées par Le Bomin et Bikoma [238, pp. 68–75] et décrites dans leur livre sur les musiques Myènè.




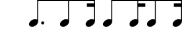





le tableau 7.1. La comparaison approfondie de ces pièces permet de penser avec bonne confiance que ces pièces ont une origine commune. Pour chacune de ces trois pièces, on observe que la métrique est identique dans les deux populations, alors que l'échelle musicale est différente. Ceci suggère que, dans ces populations, la métrique est plus stable que l'échelle utilisée, laquelle est plus variable. Cette observation nécessiterait cependant d'être approfondie, et c'est dans ce sens que nous avons entamé un travail de collecte de pièces ayant vraisemblablement une origine commune dans différentes populations du Gabon¹⁰⁶.

Corrélation avec une histoire connue. Dans certains cas des informations provenant de sources indépendantes peuvent nous renseigner sur l'histoire des populations étudiées et l'on peut chercher les paramètres musicaux qui correspondent le mieux à cette histoire. Le tableau 7.2 montre les **formules rythmiques** caractérisant trois répertoires de populations Galoa, Orungu et Nkomi décrites par Le Bomin et Bikoma [238, p. 56, p. 67]. L'analyse de l'histoire orale de ces populations suggère que les Galoa et les Orungu sont issus d'un même groupe ancestral quand la proximité avec les Nkomi vient de flux migratoires ultérieurs [238, pp. 12–20]. Or on voit dans le tableau 7.2 que dans les répertoires *Ologo* et *Imbwiri ighono*, les formules rythmiques des Galoa et des Orungu sont identiques mais qu'elles diffèrent de celles de Nkomi. Cette convergence entre histoire orale et systématique musicale peut être un argument pour utiliser la similarité des formules rythmiques comme marqueur historique.

On voit cependant que dans le répertoire *Imbwiri intche* les formules rythmiques sont différentes dans les trois populations. Ceci montre que la correspondance n'est pas systématique. Certes les méthodes d'**inférence historique** par comparaisons statistiques permettent justement d'établir des liens dans les cas où les correspondances ne sont pas strictes. Cependant la généralité de l'association observée entre **formule rythmique** et histoire est très limitée. En effet elle est issue de l'observation de deux répertoires seulement, qui plus est dans un cas très particulier. Cette association devrait donc être corroborée plus largement avant de pouvoir être utilisée comme argument solide¹⁰⁷.

106. ↑ Ce projet est rapidement présenté dans la partie 4.3.2 page 141.

107. ↑ C'est dans cette perspective que nous avons entamé le travail de description statistique des musiques du Gabon présenté dans la partie 6 pages 187 à 245.

Répertoire	<i>Imbwiri intche</i>	<i>Imbwiri ighono</i>	<i>Ologo</i>
Galoa			
Orungu			
Nkomi			

TAB. 7.2 : **Formules rythmiques caractéristiques de trois répertoires dans trois populations Myènè.** Pour chacune des trois populations Myènè (Galoa, Orungu et Nkomi) Le Bomin et Bikoma [238, p. 56, p. 67] ont décrit les **formules rythmiques** caractéristiques de trois répertoires, reproduites dans ce tableau.

7.3.3 Indicateurs *internes*

Les indicateurs internes présentés ici sont des caractéristiques des données que l'on veut utiliser pour inférer l'histoire. Ils ne proviennent pas de connaissances extérieures et doivent donc être utilisés avec précaution pour éviter la circularité du raisonnement. Je présenterai deux exemples de tels indicateurs : la mesure de la variabilité des données et les indices traditionnellement utilisés en cladistique.

Lorsque l'on étudie un ensemble de groupes (par exemple des populations), les paramètres analysés varient aussi bien au sein du groupe (entre différents individu-e-s, différentes réalisations, différents contextes, ...) qu'entre les groupes (entre différentes populations). On peut donc parler de variabilité intra-groupes ou inter-groupes. Le rapport entre ces variabilités est un élément pour estimer l'apport des paramètres utilisés dans l'inférence historique. En effet si la variabilité inter-groupes d'un paramètre est très importante, aucune similarité ne pourra être observée entre différents groupes. De même si la variabilité inter-groupes est très faible, les valeurs des paramètres seront identiques dans tous les groupes ce qui ne permettra pas non plus de rapprocher certains groupes par opposition à d'autres. Ainsi, dans le jeu de données utilisé dans l'analyse de Le Bomin et coll. [240], le caractère « présence de formes vocales antiphonales » (caractère 320) vaut 0 dans toutes les populations sauf une (les Awandji de Lipaka). Le caractère « présence du procédé polyphonique homorythmique » (caractère 318) vaut 1 dans toutes les populations où il a pu être étudié. On voit donc que ces deux paramètres sont peu informatifs pour l'inférence historique. De plus la variabilité intra-groupes doit être relativement faible et significativement plus faible que la variabilité inter-groupes de manière à ce que la valeur du paramètre dans un groupe puisse légitimement être représentative de ce groupe.

D'autres indicateurs statistiques peuvent être utilisés pour évaluer le potentiel historiographique, comme les **indices de cohérence** (ICs) et de **réention** (IRs)¹⁰⁸. Comme on l'a vu ci-dessus¹⁰⁹, les valeurs des ICs et IRs dans l'étude de Le Bomin et coll. [240] suggèrent par exemple que différents paramètres musicaux (les paramètres intrinsèques et les procédés polyphoniques) ont une histoire distincte.

108. ↑ Ces indices ont été présentés dans le paragraphe *Indices de cohérence et de réention* page 79. Les entrées du glossaire les concernant donnent plus de détail sur la manière dont ils sont calculés.

109. ↑ Voir partie 7.2.3.2 page 264.

7.3.4 Indicateurs *a posteriori*

Les indicateurs *a posteriori* sont disponibles une fois l'inférence historique effectuée. Il s'agit principalement de la comparaison des résultats obtenus avec des sources indépendantes concernant la même histoire que celle à laquelle on s'intéresse. Afin de tester la validité de leur analyse cladistique de l'histoire d'un instrument de musique (le psaltérion baltique), Temkin et Eldredge [413] effectuent aussi une analyse à partir des descriptions de cornets à pistons, instrument dont l'histoire est relativement bien connue, par différentes **méthodes issues de l'inférence phylogénétique**. Aucune des méthodes ne se révèle capable d'inférer avec justesse l'histoire établie jusque là. Cette observation peut remettre en cause la pertinence de l'utilisation des descriptions des instruments pour établir leur histoire.

7.3.5 Comparaison et limites de ces indicateurs

Indicateurs *a posteriori*. Ces indicateurs, basés sur des sources historiques indépendantes, ne permettent pas d'évaluer finement le potentiel historiographique des différents paramètres. En effet, ils donnent une réponse globale sur l'adéquation entre les résultats produits et ceux issus d'analyses indépendantes. Si le test est concluant, les données utilisées sont vraisemblablement adéquates pour l'inférence historique effectuée. Si en revanche il ne l'est pas, on peut incriminer les paramètres utilisés (mais sans dire si certains le sont plus que d'autres) mais aussi la méthode d'inférence employée. De plus chaque source historique a ses propres caractéristiques et ses limites. Ainsi un désaccord dans l'analyse de différentes sources peut très bien refléter des histoires distinctes¹¹⁰ ou encore un problème dans l'histoire indépendante à laquelle on se réfère. De manière générale, si on souscrit à l'avis de Forber et Griffith [143] selon lequel la fiabilité d'une inférence historique repose sur la confiance dans un faisceau de sources aussi indépendantes que possibles, on ne peut se servir d'une source historique pour en valider une autre puisque cela réduit l'indépendance entre elles.

Indicateurs internes. Comme on l'a vu, ces indicateurs sont déduits des mêmes données que celles utilisées pour l'analyse inférentielle. Cette double utilisation des données (à la fois pour évaluer le potentiel historiographique des paramètres et pour effectuer l'inférence) peut engendrer une circularité du raisonnement et doit donc être utilisée avec précaution. Par ailleurs la connaissance ainsi établie concernant le potentiel historiographique des différents paramètres musicaux a une portée limitée. En effet elle est intimement liée au jeu de données utilisé et ne pourra pas nécessairement être appliquée à une autre analyse.

Indicateurs *a priori*. Ces arguments, théoriques ou empiriques, sont construits par l'ethnomusicologie à partir de données et de questionnements indépendants. Ils ne sont pas directement liés à l'inférence historique, ce qui confère une confiance accrue dans leur utilisation. Cependant

110. ↑ Cette question est discutée plus longuement dans la partie 7.2 pages 251 à 268.

la connaissance ainsi acquise doit être suffisamment large et générale pour pouvoir être appliquée à l'ensemble des données étudiées, aussi bien du point de vue géographique que du point de vue culturel. Reprenons l'exemple ci-dessus suggérant que la métricité pourrait être un meilleur marqueur historique que l'échelle musicale. Cette proposition a été établie à partir de la seule étude de la côte est du Gabon. Elle ne sera cependant pas forcément valide dans la partie nord du Gabon puisque les mécanismes de transmission ou les jugements esthétiques peuvent y être différents et affecter la manière dont la musique est transmise et modifiée au cours du temps. La validité du choix de la métricité comme marqueur historique pour l'ensemble du Gabon est donc limitée par la portée de la connaissance acquise. Ce type d'arguments paraît néanmoins, de par son indépendance, le plus à même d'améliorer notre connaissance du potentiel historiographique des paramètres musicaux et c'est sans doute en explorant cette voie que les études ethnomusicologiques futures pourront améliorer nos possibilités d'inférence historique à partir des données musicales.

7.4 Conclusion du chapitre

Dans ce chapitre j'ai tâché de réfléchir, grâce à l'apport d'autres disciplines (en particulier la linguistique historique, la systématique biologique et l'anthropologie de la culture matérielle) aux différentes histoires que peuvent nous apporter les données culturelles en général et musicales en particulier. J'ai montré que les données génétiques et culturelles n'apportent pas nécessairement une information cohérente, pas plus que différents types de données culturelles (par exemple linguistique et archéologique). En outre les données appartenant à un même domaine peuvent transmettre des histoires différentes, aussi bien dans le cas de données génétiques que linguistiques ou de la culture matérielle.

La musique ne semble pas échapper à ces observations et il ne peut être supposé *a priori* que l'ensemble d'un patrimoine musical révèle une seule et même histoire. Ceci montre la nécessité d'identifier des ensembles de paramètres musicaux ayant une histoire commune, par exemple parce qu'ils présentent une cohérence du point de vue de leur transmission. Par ailleurs tous les paramètres musicaux ne semblent pas avoir la même capacité à retenir et révéler les traces de l'histoire. Il s'agit donc de déterminer dans quelle mesure un paramètre est intéressant pour inférer l'histoire, c'est-à-dire évaluer son *potentiel historiographique*. J'ai proposé quelques critères permettant d'évaluer ce potentiel et j'ai discuté de leurs limites respectives. Je n'ai cependant pas encore eu l'occasion d'appliquer ces critères à mes propres données et ce sont donc pour l'instant de simples propositions théoriques qui n'ont pas encore eu l'occasion d'affronter l'épreuve de la pratique.

Conclusion de la deuxième partie

J'ai présenté dans cette seconde partie de thèse les principales recherches que j'ai effectuées au cours de mon doctorat. Le chapitre 5 a décrit l'utilisation d'une méthode d'**inférence phylo-géographique** pour étudier l'histoire d'une soixantaine de **patrimoines musicaux** du Gabon¹¹¹. À partir de l'**arbre** de ces patrimoines musicaux issu de l'article de Le Bomin et coll. [240] et des positions géographiques où ils ont été collectés, il s'agissait d'inférer la position ancestrale d'une vingtaine de **nœuds internes** de l'arbre ayant une certaine pertinence géographique ou culturelle grâce à une méthode d'**inférence bayésienne**. Au-delà des améliorations qui pourraient y être apportées, cette analyse comporte des problèmes majeurs : L'arbre utilisé ne peut être considéré comme une modélisation de l'histoire des patrimoines musicaux et les fondements théoriques et méthodologiques sont mal définis. Cette constatation m'a conduit à orienter mes recherches d'une toute autre manière.

Tout d'abord cette analyse préliminaire a révélé un manque de recul sur les données que j'avais à analyser. Sylvie Le Bomin a fait une description minutieuse de patrimoines musicaux issus de différentes parties du territoire gabonais, mais aucune vue d'ensemble n'en a été donnée. Il m'est donc apparu nécessaire, avant d'utiliser ces données pour inférer l'histoire de ces musiques, d'en donner un aperçu général afin de comprendre la manière dont était organisée, culturellement et géographiquement, la diversité des musiques du Gabon. C'est ce à quoi nous nous sommes attaché-e-s dans le chapitre 6¹¹², qui reproduit un article récemment soumis par Sylvie Le Bomin et moi-même à la revue *Plos One*. Nous nous sommes intéressé-e-s plus particulièrement à l'étude des **paramètres internes** de ces musiques, c'est-à-dire aux organisations **rythmique**, **métrique** et **scalaire** des **pièces** étudiées. Nous avons pu constater que la manière dont est organisée, géographiquement et culturellement, la diversité des **valeurs** observées est très différente d'un **paramètre** à l'autre. Les musiques du Gabon montrent aussi certaines particularités relativement aux autres musiques d'Afrique centrale, notamment une moins forte propension des paramètres musicaux (et en particulier rythmiques) à caractériser un **répertoire** dans une population. Enfin le partage de formules rythmiques par différentes populations semblent refléter, plus que celui des autres paramètres musicaux, une proximité géographique laissant penser que les formules rythmiques pourraient être de bons marqueurs historiques.

111. ↑ Voir **chapitre 5** pages 155 à 185.

112. ↑ Voir **chapitre 6** pages 187 à 245.

Par ailleurs l'analyse phylogéographique du chapitre 5 souffre de lacunes méthodologiques et théoriques et a montré le besoin de fondements plus solides avant de continuer à rechercher l'histoire à travers la comparaison des musiques. En effet quelle histoire peut nous révéler la musique? Et quelles données musicales peuvent nous permettre de l'explorer? En réfléchissant à ces questions à la lumière de l'expérience d'autres disciplines historiques, j'ai suggéré dans le chapitre 7 que l'histoire de différents types de données (biologiques, linguistiques, anthropologiques et musicales) ne convergent pas nécessairement en une histoire unique, pas plus que celles que l'on peut déduire de différents types de données musicales (instruments, paramètres internes, ...) ¹¹³. Ceci montre la nécessité de pouvoir identifier des ensembles de paramètres musicaux (et éventuellement culturels) ayant une cohérence historique. J'ai aussi proposé quelques critères, à travers la notion de *potentiel historiographique*, pour évaluer la confiance que l'on peut avoir en un paramètre musical lorsque l'on cherche à en inférer l'histoire.

Ces deux derniers chapitres m'ont permis d'affiner les pistes qu'il serait intéressant de poursuivre si l'on veut élaborer des méthodes d'inférence historique par comparaison de données musicales qui soient à la fois pertinentes et efficaces. Je présenterai quelques unes de ces perspectives dans le chapitre 8, proposant d'approfondir l'étude de la transformation et de la transmission des musiques ainsi que d'améliorer nos méthodes de comparaison des formules rythmiques ¹¹⁴. Mais le cheminement sinueux de ce doctorat m'a aussi conduit à remettre en question mes pratiques de recherches et à m'interroger sur les problèmes et enjeux du travail interdisciplinaires. Je partagerai donc ¹¹⁵ quelques réflexions issues de cette expérience dans le chapitre 9, avant de conclure cette thèse.

113. ↑ Voir chapitre 7 pages 247 à 274.

114. ↑ Voir chapitre 8 pages 279 à 292.

115. ↑ Voir chapitre 9 pages 295 à 313.

Troisième partie

Discussion

Chapitre 8

Pistes de recherches

8.1	Comprendre la transmission et la transformation des musiques	280
8.1.1	Observer le changement des pratiques musicales	281
	Enregistrements d'archives	281
	Problèmes des archives	281
	Collectage <i>ad hoc</i>	282
8.1.2	Comprendre les dynamiques de la transmission musicale	283
8.1.2.1	Nécessités de s'intéresser à la transmission	283
	Comprendre la transmission pour mieux inférer l'histoire	283
	Des transmissions différentes selon les paramètres	283
	Influence du type de transmission sur le musical	284
8.1.2.2	Deux exemples d'études de la transmission	284
	Étudier la transmission au sein d'une culture musicale	284
	L'emprunt entre différentes cultures musicales	285
8.1.3	Étudier la généralité et la portée de ces mécanismes	285
8.1.4	Relier l'observé au passé	286
8.2	Approfondir l'étude des formules rythmiques	287
8.2.1	Accroître le jeu de données	288
8.2.2	Affiner la comparaison des formules rythmiques	288
8.2.2.1	Quantifier la ressemblance	288
	Cellules rythmiques contenues	289
	Similarité entre cellules rythmiques	290
8.2.2.2	Décrire la rythmique indépendamment de la métrique	291
	Une même pulsation avec plusieurs subdivisions possibles	291
	Plusieurs pulsations possibles	292

J'ai montré dans la première partie¹ comment la comparaison de structures biologiques, linguistiques et culturelles pouvait apporter des informations sur leur histoire. On a vu en particulier en quoi les pratiques musicales semblent se prêter à une telle comparaison à visée historique. La seconde partie², présentant les axes et résultats de mon travail de recherche, a cependant permis de constater les difficultés posées par l'**inférence historique** à partir de la comparaison d'un grand nombre de données musicales. En effet, il ne s'agit pas simplement de transposer des méthodes venues d'autres disciplines : avant toute chose il est nécessaire d'une part de bien comprendre les données dont on dispose et ce qu'elles peuvent nous apprendre et, d'autre part, d'établir les bases théoriques et épistémologiques permettant d'avoir confiance dans de telles inférences. On voit que, concernant ces deux points, d'importantes recherches doivent encore être poursuivies. Si des développements théoriques sont sans doute nécessaires, il me semble que ce sont les connaissances empiriques concernant la transmission et la transformation du musical au cours du temps qui font aujourd'hui défaut et qui seules permettront de proposer des méthodes et adaptations méthodologiques pertinentes.

Ce chapitre expose quelques pistes dans cette direction. Je présente d'abord quelques réflexions qui pourraient guider des recherches futures. Je propose ensuite d'approfondir l'étude des **formules rythmiques**, qui semblent se prêter plus particulièrement à l'inférence historique, en suggérant quelques améliorations possibles dans les méthodes pour les comparer.

8.1 Comprendre la transmission et la transformation des musiques

Dans le chapitre précédent³, j'ai pu identifier plusieurs types d'indices pouvant permettre de développer des méthodes robustes pour traiter les données musicales en termes d'inférence historique. On peut par exemple comparer si les inférences obtenues à partir de données musicales corrélient avec celles issues d'autres types de données, par exemple génétiques ou linguistiques. Mais une inférence historique ne peut être considérée comme robuste que si elle repose sur un faisceau de preuves indépendantes. Utiliser des données externes pour valider cette inférence permet tout au plus de conforter un peu la méthode elle-même mais pas de renforcer l'inférence réalisée. Il est donc nécessaire d'approfondir notre compréhension de la transmission et de la transformation des musiques au cours du temps, afin de mieux cerner quels types de données liées à la musique peuvent nous informer sur leur histoire et si oui par quelles méthodes.

Comme prélude indispensable à la compréhension de l'**évolution du musical**, il faut sans doute commencer par documenter la transformation des musiques⁴. Cette étape doit être complétée par l'étude des mécanismes sociaux et individuels qui président à la transmission des pratiques

1. ↑ Pages 15–146.

2. ↑ Pages 155–276.

3. ↑ Voir **partie 7.3.5** pages 273 à 274.

4. ↑ Voir **partie 8.1.1** page 281.

musicales⁵. Les connaissances ainsi acquises ne pourront cependant être utiles à l'inférence historique que si l'on peut montrer leur généralité à l'échelle d'une zone géoculturelle suffisamment large⁶. Une fois de telles informations recueillies et utilisables pour l'inférence historique, il faudra pouvoir tester les méthodes obtenues avec d'autres sources historiques : il convient donc dès maintenant de rechercher l'histoire des musiques par d'autres biais, en particulier à travers les témoignages ethnohistoriques⁷.

8.1.1 Observer le changement des pratiques musicales

Enregistrements d'archives. La première approche est d'étudier en détail la réalité de la transformation des pratiques musicales au cours du temps. Lacombe [232] compare par exemple une quarantaine d'enregistrements de musiques pour xylophones issues de populations Fang du Gabon, du Cameroun et de Guinée Équatoriale. Elle met ainsi en évidence des éléments stables dans ces pratiques musicales (le contexte d'exécution, l'accord des instruments ou encore les modalités de jeu) ainsi que des éléments variables comme la dénomination des pièces ou les rôles respectifs des instruments. Dans la même perspective, j'ai déjà évoqué le travail d'Émeline Lechaux [249]⁸ qui compare de manière beaucoup plus approfondie différentes versions de cérémonies du *bwétè* chez les Tsogho du Gabon.

De telles études peuvent nous renseigner sur les paramètres musicaux les plus sujets à la transformation au cours du temps. Elles nous informent aussi sur les liens qui peuvent exister entre les transformations de différents paramètres musicaux. Ainsi Lacombe [232, p. 99] explique qu'un même nom ne désigne pas nécessairement la même pièce à différentes époques. Elle suggère donc que la dénomination endogène des pièces puisse être transmise de manière indépendante du matériel musical. De telles constatations peuvent nous renseigner sur les données qui devraient être utilisées conjointement (ou au contraire celles qui devraient être analysées séparément) lorsque l'on cherche à tirer des informations historiques.

Après plus d'un siècle de collectage, il existe aujourd'hui de nombreux enregistrements des musiques du Gabon. En plus des fonds d'archives constitués et inventoriés, tels que le fond Herbert Pepper [45] ou le fond Pierre Sallée [248], de nombreux enregistrements réalisés depuis une vingtaine d'années par Sylvie Le Bomin pourraient être analysés. Ce sont autant de sources d'informations potentielles sur la manière dont les musiques se transforment au cours du temps.

Problèmes des archives. L'utilisation des données d'archives pour informer sur l'évolution des musiques soulève cependant certaines difficultés. Dans le cas d'une analyse basée sur une mission de terrain, les enregistrements sont effectués spécifiquement pour répondre à la question de recherche que l'on s'est posée et les lieux et contextes d'enregistrement sont choisis, dans la mesure du possible, en fonction de cet objectif. Dans le cas de l'utilisation de données d'archives,

5. ↑ Voir [partie 8.1.2](#) page 283.

6. ↑ Voir [partie 8.1.3](#) page 285.

7. ↑ Voir [partie 8.1.4](#) page 286.

8. ↑ Voir paragraphe [La comparaison diachronique](#) page 136.

il faut composer avec les enregistrements disponibles. Il n'est alors évidemment pas possible de maîtriser les lieux et contextes de production des musiques étudiées, la seule marge de manœuvre résidant dans le choix des enregistrements à inclure dans l'analyse. Ces contraintes imposent un important biais d'échantillonnage. Lacombe [232, p. 99] compare par exemple le nombre d'enregistrements de musiques pour xylophones portatifs avec le nombre d'enregistrement de musiques pour xylophones sur troncs de bananier présents dans les différents corpus étudiés. Comment être certain que les proportions respectives de ces enregistrements sont bien représentatives des pratiques réelles ? Le fait qu'elle n'ait trouvée, dans les enregistrements de Tessmann de 1908, aucun enregistrement de musiques pour xylophones portatifs signifie-t-il pour autant que ces instruments n'étaient pas présents à ce moment ?

Par ailleurs, les enregistrements comparés proviennent généralement de populations de même ethnonyme, mais issues de diverses localités. Comment savoir si les différences relevées témoignent d'une *évolution* temporelle, ou simplement d'une variabilité géographique ? Une comparaison entre deux ensembles ne peut être informative que si l'on a une idée de la diversité interne à chacun de ces ensembles. Il faudrait donc connaître la diversité des paramètres musicaux étudiés aux époques comparées, afin de pouvoir juger si les différences observées sont imputables à un changement *diachronique*. L'idéal serait sans doute de pouvoir comparer les musiques enregistrées à différentes époques dans une région très circonscrite, par exemple un même village. Mais, selon les paramètres étudiés, il peut exister une variabilité interne au sein d'un village, dues par exemple à des spécificités d'interprètes ou simplement aux marges d'exécutions culturellement autorisées entre plusieurs exécutions d'une même pièce. C'est l'intérêt de la notion de *modèle*⁹, qui permet de donner une représentation de la pièce faisant abstraction des variations culturelles. Mais il n'est pas toujours possible d'accéder à ce modèle à partir d'enregistrements d'archives, car ils n'ont pas nécessairement été réalisés dans ce but. On voit que les différences observées entre deux époques doivent être comparées à celles observées entre les variantes à une même époque, afin de pouvoir estimer si elles sont significatives. Les données disponibles ne permettent généralement pas de satisfaire à ces exigences et les informations qu'on en tire restent alors limitées.

Collectage *ad hoc*. La comparaison d'enregistrements issus de différentes époques et de collectes aux perspectives variées permet donc de donner une idée de l'évolution des musiques, mais bien souvent sans permettre de décrire de manière précise cette évolution. Une autre approche est donc de constituer un corpus d'enregistrements spécifiquement conçu pour étudier l'évolution musicale. On peut par exemple collecter à différentes époques les mêmes pièces pratiquées par les mêmes populations afin d'en décrire les éléments inchangés et ceux qui ont été modifiés. Mais là encore la diversité observée pourrait être simplement le reflet de la variabilité interne présente à chacune des époques. De plus, puisque la démarche est alors explicitement comparative, on pourra anticiper cet aspect et prévoir des enregistrements pour étudier la variabilité existante dans la population étudiée (notamment via le modèle des pièces considérées). Ainsi, lors du collec-

9. ↑ Voir [partie 3.1.3](#) pages 109 à 113.

tage que nous avons effectué dans la perspective d'une comparaison **synchronique** de différentes pièces¹⁰, Sylvie Le Bomin et moi-même avons enregistré plusieurs versions différentes de la pièce *Igogo nuga* chez les Sekiani du quartier d'Alibadeng (Libreville) ainsi que chez les Benga du Cap Estérias. Ceci nous permettra d'estimer la variabilité qui peut exister entre les variantes d'une même pièce et d'estimer si les différences observées entre les enregistrements faits dans plusieurs populations sont pertinentes.

Cette approche, sans doute beaucoup plus informative, a cependant elle aussi ses limites. Elle demande un investissement important dans le collectage et sur le long terme. En outre, le financement des projets de recherche s'inscrit bien souvent dans des durées très brèves (de l'ordre de quelques années au maximum) et la pérennité d'une telle problématique ne peut venir que de la persévérance de l'ethnomusicologue¹¹. Elle ne permet par ailleurs qu'une comparaison sur des durées relativement courtes, de quelques dizaines d'années soit la durée d'une carrière de recherche.

8.1.2 Comprendre les dynamiques de la transmission musicale

8.1.2.1 Nécessités de s'intéresser à la transmission

Comprendre la transmission pour mieux inférer l'histoire. Observer la transformation du matériau musical au cours du temps peut nous renseigner sur la vitesse des changements qu'il subit et donc sur la profondeur historique sur laquelle ce matériau peut nous informer. Mais l'intérêt reste limité si l'on ne s'intéresse pas en même temps aux processus de transmission des musiques et à la manière dont ceux-ci affectent le matériau musical. Comparer deux musiques dans une perspectives diachronique suppose que celles-ci ont un lien à travers l'histoire, qui peut être interprété comme des formes de transmission, qu'il s'agit d'étudier et de comprendre.

Des transmissions différentes selon les paramètres. Il est évident que tous les paramètres musicaux ne sont pas transmis de la même manière. Ainsi les mécanismes individuels et collectifs d'apprentissage et de reproduction d'une pièce musicale sont vraisemblablement très différents de ceux de transmission de la fabrication des instruments de musique. Décrire les mécanismes de transmission peut permettre de mieux comprendre pourquoi différents paramètres musicaux n'ont pas la même histoire. Cela peut aider à identifier des « paquets » de paramètres musicaux soumis aux mêmes processus de transmission et donc susceptibles d'avoir connu une histoire semblable. Il devient alors légitime d'étudier l'histoire de chacun de ces paquets¹² et éventuellement d'analyser les histoires croisées de différents paquets. On peut comprendre ainsi le choix de Temkin [412] de distinguer, dans son analyse **cladistique** des psaltérions baltiques, les **caractères** décoratifs d'une part et les caractères structurels d'autre part. Ces deux types de caractères ayant des fonctions

10. ↑ Voir paragraphe *Histoire de pièces* page 141.

11. ↑ Cette persévérance est grandement facilitée par l'obtention d'un poste permanent de chercheur·e, permettant de s'engager à long terme sur un même terrain, dans une même perspective.

12. ↑ Par exemple par une **méthode issue de l'inférence phylogénétique**.

sociales bien distinctes (esthétique ou fonctionnel), on peut s'attendre à ce que les mécanismes culturels de transmission de ces caractères soient différents¹³.

Influence du type de transmission sur le musical. La transmission peut bien entendu être comprise comme la manière, au sein d'une culture donnée, de perpétuer ses pratiques musicales. Mais la transmission concerne de manière plus générale toute reproduction du matériau musical, éventuellement entre plusieurs populations. Là encore ces deux mécanismes de transmission du musical, au sein d'une population ou entre populations distinctes, mettent en jeu des processus sociaux très différents et il est probable que le matériau musical en soit affecté de manière diverse. On peut ainsi chercher à comprendre comment un répertoire musical peut-être emprunté et comment le matériau musical est modifié au cours de cet emprunt. C'est par exemple le travail, déjà évoqué¹⁴, réalisé par Susanne Fűrnis [151] concernant l'emprunt du répertoire du *bèkà* au Cameroun. Aussi peut-on imaginer que différents processus de transmission imposent différents types de modifications au musical. Cette connaissance peut servir en retour à émettre des hypothèses, à partir des caractéristique observées sur le matériau musical, concernant les types de transmission qui l'ont affecté (par exemple un emprunt). Cette approche est couramment utilisée en linguistique historique pour écarter les emprunts, identifiés par expertise linguistique¹⁵. Il s'agit donc de s'intéresser aux mécanismes de transmission du musical aussi bien au sein d'une population qu'entre des populations différentes et de décrire leurs impacts sur le matériau musical.

8.1.2.2 Deux exemples d'études de la transmission

Étudier la transmission au sein d'une culture musicale. Concrètement, il s'agit tout d'abord de comprendre comment les différents paramètres musicaux sont transmis d'une génération à l'autre au sein d'une population. Sylvie Le Bomin étudie ainsi en détail le processus d'apprentissage des musiques pour xylophones chez les Banda Gbambiya d'Afrique centrale [234, 235, 237]. Elle montre que la transmission d'une pièce musicale, en particulier des parties solistes, passe par l'apprentissage de différents énoncés musicaux de référence (les matrices) de difficulté croissante [237, p. 181]. La matrice 1, qui est le premier énoncé musical appris, n'est en fait jamais réalisé hors de ces situations d'apprentissage [237, p. 182]. C'est pourtant cette première matrice qui est au fondement de la transmission et qui devrait donc être comparée [240, p. 182]. Cette compréhension de la transmission musicale est au fondement de la comparaison des patrimoines musicaux de différentes populations du Gabon faite par Le Bomin et coll. [240] : les paramètres musicaux sont identifiés non pas d'après les énoncés musicaux réalisés en contexte, mais d'après les énoncés minimaux réalisés à la demande de l'ethnomusicologue, simulant en quelque sorte les énoncés utilisés pendant la transmission. On voit donc comment la compréh-

13. ↑ Voir à propos de la transmission de différents paquets culturels la partie 7.2.2 pages 260 à 262.

14. ↑ Voir paragraphe *Traces historiques entre cultures proches* page 142.

15. ↑ Voir aussi l'étude de Gaillard-Corvaglia et coll. [A155]¹⁶, présentée dans le paragraphe *Méthodes cladistiques* page 63.

16. Citée par Darlu [106, pp. 67–68].

sion des mécanismes de transmission peut nourrir les méthodes de l'inférence historique à partir de données musicales.

L'emprunt entre différentes cultures musicales. S'intéresser à la transmission c'est aussi comprendre comment un même paramètre musical peut être transmis d'une culture à une autre. Il s'agit donc comme le fait Susanne Fürniss [151], d'étudier en détail des cas concrets d'emprunt. Il n'est souvent pas difficile d'identifier un répertoire emprunté, car les interprètes sont généralement en mesure d'explicitier ces échanges entre groupes. Il est cependant parfois plus difficile de déterminer le parcours réel suivi par le répertoire au cours des emprunts successifs, les discours sur ces emprunts pouvant être lacunaires ou sujets à des manipulations identitaires [151, p. 36].

L'unité d'étude de l'emprunt ne doit pas nécessairement être le répertoire. Nous avons pu observer qu'au Gabon des pièces pouvaient être réalisées dans des répertoires différents d'une population à l'autre, mais aussi au sein d'une même population. Ceci suggère qu'une pièce seule pourrait être empruntée, indépendamment de son contexte d'exécution. L'impact de l'emprunt sur le musical peut donc être étudié au niveau du répertoire mais aussi au niveau de la pièce. Par ailleurs d'autres paramètres peuvent être empruntés selon des processus différents, comme les instruments de musique. Là encore leur origine fait souvent l'objet d'un discours endogène qui peut servir comme point de départ pour étudier les mécanismes de transmissions.

8.1.3 Étudier la généralité et la portée de ces mécanismes

Il serait alors possible d'utiliser les mécanismes ainsi décrits et identifiés comme hypothèses de départ sur lesquelles fonder les inférences historiques. Il est cependant nécessaire que ces mécanismes soient valables et opérationnels pour l'ensemble des musiques étudiées et comparées. En effet, on ne peut supposer *a priori* qu'un mécanisme de transmission et ses effets décrits en une localité précise aient une portée plus générale. La reproduction des pratiques musicales est issue des structures sociales de transmission et dépend donc des cultures considérées. Ainsi les études conduites dans différentes parties d'Afrique centrale, en particulier en République centrafricaine, au Gabon et au Cameroun, ont montré que cette fonction centrale des formules rythmiques est partagée sur une aire culturelle assez large. Ceci pourrait servir à justifier l'utilisation des formules rythmiques pour inférer l'histoire dans cette région, mais ce ne serait pas nécessairement le cas pour d'autres zones culturelles où l'organisation rythmique et ses modalités de transmission sont différentes.

Cette précaution élargit d'une certaine manière la préconisation de Mantle Hood selon laquelle il est nécessaire, avant de les comparer, de décrire et de connaître les cultures musicales étudiées¹⁷. Si la comparaison est effectuée dans une perspective historique, il faut aussi s'intéresser aux mécanismes de transmission des pratiques musicales. Cependant, il n'est peut-être pas nécessaire de s'interdire toute comparaison tant qu'une connaissance fine à la fois des musiques comparées,

17. ↑ Voir paragraphe *Critiques et rejets de la comparaison* page 131.

mais aussi des mécanismes de leur transmission, n'a pas été acquise. En effet, s'attaquer empiriquement à la comparaison permet de défricher les problèmes qu'elle pose. Ainsi, comme on l'a vu, l'analyse des musiques d'Afrique centrale a permis de dégager des connaissances sur la transmission des musiques [151, 236, 295] qui nourrissent la comparaison empirique de musiques du Gabon faite par Le Bomin et coll. [240]. Cette étude en retour permet de soulever d'autres questions concernant les mécanismes de transmission, par exemple concernant le problème de la **congruence** des caractères musicaux¹⁸, discuté dans le chapitre 7.

Les études empiriques conduisent donc à réfléchir aux recherches qui pourraient accroître le sens de futures comparaisons. Aussi il me paraît légitime d'effectuer ces comparaisons, fondées sur des connaissances parfois imparfaites, dès lors que l'on peut juger et discuter ces bases. Les résultats proposés peuvent alors être évalués à leur juste valeur : non pas des résultats solidement établis et définitifs, mais des pistes exploratoires, des indices sur la direction à donner à nos recherches et sur la manière d'améliorer nos méthodes. Ainsi Le Bomin et coll. [240] ne prétendent pas reconstruire les musiques passées mais proposent plutôt de détecter la présence d'un signal de transmission verticale. Ceci permet en retour de donner des pistes pour d'autres recherches, puisque cette étude m'a conduit à m'intéresser plus en détails aux **paramètres intrinsèques**, dont l'**indice de cohérence** mesuré est particulièrement élevé¹⁹.

8.1.4 Relier l'observé au passé

Comme je l'ai soutenu²⁰, confronter les résultats de l'inférence historique par comparaison **synchronique** des musiques avec les connaissances provenant d'autres sources historiques ne permet pas, compte tenu du côté expérimental de l'utilisation de ces méthodes d'inférence, d'affiner nos connaissances concernant l'histoire. Cependant cela permet de tester la validité de ces méthodes, en prouvant leur capacité à obtenir des résultats établis par ailleurs. Encore faut-il que ces méthodes soient dotées d'un fondement théorique cohérent et robuste qui puisse nous permettre de penser légitimement que la validation obtenue ne soit pas un hasard. Ceci justifie alors l'utilisation de cette méthode pour proposer des hypothèses historiques (et plus seulement à des fins heuristiques ou de développement méthodologique). On voit donc pourquoi il est important de comprendre la dynamique de transformation du matériau musical, comme exposé ci-dessus.

Mais après avoir établi ces connaissances sur l'évolution du musical, les mécanismes de transmission et la généralité de ces connaissances dans les cultures étudiées, il me semble intéressant de tester nos méthodes d'inférences historiques. C'est lors de cette étape de contrôle que la confrontation des résultats obtenus avec ceux issus d'autres sources historiques me paraît réellement utile. C'est ce que font par exemple Temkin et Eldredge [413] dans leur étude des cornets à piston. Après avoir inféré par différentes **méthodes issues de l'inférence phylogénétique** les relations (sous forme d'**arbre** ou de **réseau**) entre différents cornets, ils comparent leurs résultats à

18. ↑ Voir **partie 7.2** pages 251 à 268.

19. ↑ Voir 622 page 266.

20. ↑ Voir **partie 7.3.5** page 273.

l'histoire établie de l'évolution du cornet, bien documentée par ailleurs. Constatant les différences importantes obtenues, ils estiment que l'échec de ces méthodes vient entre autres des différences de mécanismes entre transmissions biologiques et culturelles [413, p. 150]. Ils considèrent en particulier que deux états d'un « caractère » organologique (comme le type de piston) ne peuvent souvent être considérés comme des états successifs, l'un ne pouvant être caractérisé comme l'état ancestral et l'autre comme l'état dérivé. L'inadéquation de leurs résultats avec ceux établis les conduit ainsi à réfléchir en retour sur les mécanismes de transmission spécifiques qui affectent les instruments de musique afin d'améliorer la méthode d'inférence.

Pouvoir tester une méthode requiert donc d'avoir d'autres informations historiques sur lesquelles se baser pour pouvoir l'évaluer. Dans le cas des musiques du Gabon, on ne connaît que très peu de choses de leur histoire, ce qui rend difficile cette vérification. Une stratégie pourrait être de partir des indices historiques qui ont pu être recueillis jusque là. On dispose par exemple de divers indices concernant l'origine de la danse féminine *Ivanga* des Myènè, comme retracé par Le Bomin et Bikoma [238, pp. 81–84]. Ceci pourrait servir de départ à une étude plus approfondie de l'histoire (en particulier via les sources ethnohistoriques) de cette danse d'une part et à la comparaison des musiques aujourd'hui jouées pendant l'*Ivanga* d'autre part avant de mettre en relations ces deux types d'informations. Compte tenu de la difficulté d'obtenir des informations historiques et de nos faibles connaissances sur la transformation et la transmission du musical au cours du temps, il peut paraître sage de commencer par des études d'une profondeur historique très faible (quelques dizaines d'années) dont l'histoire peut sans doute être plus facilement retracée. Ces difficultés soulignent en tout cas une nouvelle fois la nécessité de documenter les musiques d'une zone géographique restreinte à différentes époques, comme suggéré ci-dessus²¹.

8.2 Approfondir l'étude des formules rythmiques

Comme on l'a vu dans le chapitre 6, il semble assez rare qu'une même formule rythmique soit retrouvée dans deux populations différentes par un simple hasard (c'est-à-dire par convergence), beaucoup plus rare en tout cas que pour d'autres paramètres musicaux. Ceci suggère que le partage d'une formule rythmique reflète une parenté historique entre ces deux formules rythmiques. Les formules rythmiques pourraient donc être des indicateurs de parenté entre des musiques voire entre des populations.

Il paraît donc intéressant d'explorer plus en détail la diversité des formules rythmiques dans les musiques du Gabon. En plus des perspectives déjà évoquées plus haut visant à mieux comprendre la transmission et la transformation des paramètres musicaux en général, on peut explorer deux autres pistes concernant les formules rythmiques. Tout d'abord augmenter la taille du jeu de données utilisé. D'autre part affiner les méthodes de comparaison entre formules rythmiques.

21. ↑ Voir partie 8.1.1 page 281.

8.2.1 Accroître le jeu de données

Dans l'analyse présentée dans le chapitre 6, la formule rythmique a été identifiée pour au moins l'une des parties de tambour dans 137 pièces. Cependant de nombreux enregistrements réalisés par Sylvie Le Bomin au cours de ses différentes missions n'ont pas été exploités. En effet, certaines pièces n'ont pas encore été transcrites (notamment parce qu'elles ont été enregistrées récemment) et d'autres l'ont été mais n'ont pas été incluses dans les analyses. Au total, en ce qui concerne l'organisation rythmique, il est possible que 200 pièces collectées par **enregistrement analytique** n'aient pas encore été exploités. Compte tenu de l'intérêt que semble avoir la comparaison de formules rythmiques pour retrouver l'histoire des musiques, il pourrait être intéressant de transcrire en priorité les formules rythmiques de ces pièces. Ceci permettrait d'avoir un jeu de données plus large et explorer plus en détail les hypothèses proposées dans le chapitre 6.

8.2.2 Affiner la comparaison des formules rythmiques

La comparaison des **formules rythmiques** pourrait être affinée, ce qui permettrait d'améliorer les hypothèses historiques qui en sont issues. En effet, la méthode de comparaison présentée dans le chapitre 6 pose plusieurs problèmes pour lesquels je propose des pistes de réflexion ci-dessous. D'une part on peut envisager des méthodes plus fines de comparaison des formules rythmiques. D'autre part la méthode de codage choisie est dépendante de la **métricité** de la pièce, ce qui peut parfois occulter certaines similarités.

8.2.2.1 Quantifier la ressemblance

Le travail de Le Bomin et coll. [240] compare directement les **cellules rythmiques**, ce qui donne une information plus fine que la seule comparaison des formules rythmiques. Cependant les cellules rythmiques sont considérées de manière indépendante, ce qui ne permet pas de prendre en compte les propriétés de la formule rythmique comme un tout, qui a une cohérence propre et qui est reconnue comme telle par les interprètes. Nous avons d'ailleurs montré²² que la comparaison des seules cellules rythmiques est beaucoup moins spécifique que celle des formules rythmiques. À l'inverse, dans les analyses du chapitre 6, la comparaison entre les formules rythmiques est assez rudimentaire puisque deux formules ne peuvent être que identiques ou différentes. Aucune nuance n'est possible entre ces deux extrêmes.

Il existe pourtant des degrés dans les différences entre formules. On voit par exemple dans le tableau 8.1 que les formules *I.a* et *I.b* sont beaucoup plus semblables que les formules *I.a* et *I.c*²³. On pourrait donc affiner la comparaison des formules rythmiques, par exemple en quantifiant la ressemblance entre deux formules, c'est-à-dire en leur attribuant une mesure indiquant à quel point elles sont proches. On voit dans l'exemple du tableau 8.1 qu'au moins deux paramètres peuvent permettre de mieux prendre en compte la similarité entre formules rythmiques : les

22. ↑ Voir la partie *Discussion* (pages 213 et suivantes) du chapitre 6.

23. ↑ Ou que les formules *I.b* et *I.c*.

Exemple	Pièces	Formule rythmique			
I.a	RAEI019	2	3	2	3
	RAEI041				
I.b	RAEI029	2	3	2	4
	RAEI051				
I.c	RAEI010	3	19	6	19
	RAEI028				







TAB. 8.1 : **Similarité des formules rythmiques en fonction des cellules rythmiques contenues.** Chaque ligne du tableau correspond à une **formule rythmique** trouvée dans une ou plusieurs pièces (listé-e-s dans la colonne « Pièces »). Pour chaque formule, les différentes **cellules rythmiques** sont notées en notation solfégique (partie basse de la ligne) et le code de la cellule (présenté dans le tableau des cellules rythmiques, chapitre 6, page 226) est donné au-dessus. On voit que les compositions en cellules rythmiques des formules *I.a* et *I.b* sont beaucoup plus similaires que celles des cellules *I.a* et *I.c* ou celles des cellules *I.b* et *I.c*.

cellules rythmiques contenues dans chaque formule et la similarité entre les cellules rythmiques elles-mêmes.

Cellules rythmiques contenues. On peut comparer deux formules rythmiques en considérant leur composition en termes de cellules rythmiques. De ce point de vue, on voit dans le tableau 8.1 que la formule rythmique *I.a*, composée des cellules « 2, 3, 2, 3 » est très proche de la formule *I.b* composée des cellules « 2, 3, 2, 4 » puisque trois des quatre cellules contenues sont identiques, seule la dernière étant différente (respectivement « 3 » ou « 4 »). En revanche, elle est très différente de la formule *I.c* composée des cellules « 3, 19, 6, 19 », puisqu'une seule cellule se retrouve dans les deux formules (la cellule « 3 »). On pourrait ainsi quantifier une ressemblance de $\frac{3}{4} = 75\%$ entre *I.a* et *I.b* et de $\frac{1}{4} = 25\%$ entre *I.a* et *I.c*.

Cette modalité de calcul est sans doute elle-même un peu simpliste et demanderait à être encore affinée. Ainsi faudrait-il prendre en compte l'ordre des cellules dans la formule rythmique. Ainsi dans l'exemple du tableau 8.2 les deux formules contiennent exactement les mêmes cellules (« 31 », « 34 » et « 46 ») mais dans un ordre différent. Si l'on compare seulement les cellules rythmiques contenues (sans prendre en compte leur ordre), la similarité serait de $\frac{3}{3} = 100\%$. Les formules seraient donc considérées identiques, ce qui occulterait la différence observable et qui fait, comme on l'a dit, la spécificité d'une formule rythmique, à savoir l'ordre de succession des cellules²⁴. On pourrait donc ne comparer que les cellules situées à la même position : la première cellule de la formule *II.a* est identique à celle de la formule *II.b* (« 31 ») ; la seconde de *II.a* (« 34 ») est différente de la seconde de *II.b* (« 46 ») tout comme la troisième de *II.a* (« 46 ») et de *II.b* (« 34 »). La similarité serait alors de $\frac{1}{3} = 33\%$. Les formules seraient donc considérées comme plutôt dissemblables, ne prenant pas en compte la forte similarité en ce qui concerne les cellules



24. ↑ Si la formule rythmique ne contient qu'une ou deux cellules, il n'y a qu'un ordre possible (puisque les pièces considérées sont cycliques). La prise en compte de l'ordre des cellules rythmiques est pertinente dès lors qu'il y a au moins trois cellules, ce qui est fréquemment le cas, comme on l'a vu dans le chapitre 6 (pages 191 à 245, voir en particulier le tableau des données pages 228 et suivantes).

Exemple	Pièce	Formule rythmique		
II.a	RAEI118	31	34	46
				
II.b	RAEI115	31	46	34
				





TAB. 8.2 : **Similarité des formules rythmiques en fonction de l'ordre des cellules.** Ces deux pièces viennent de la même population (Ngom - Moualo #5) mais de deux répertoires différents (respectivement *lisimbu* et *minonze*).

rythmiques qui les composent.

Une piste pour remédier à ce problème pourrait être d'utiliser une mesure s'inspirant des *distances d'édition*^{25,26}. En effet ce type de distances mesure le nombre de modifications nécessaires pour passer d'une séquence à une autre. Ces modifications pourraient être, dans le cas présenté ici, des ajouts de cellules rythmiques, des suppression de cellules ou des changements de position²⁷ dans la séquence. Ainsi en effectuant une seule modification sur la formule *II.a* (intervertir la deuxième et la troisième cellule), on obtient la formule *II.b*. La distance entre ces deux formules serait donc de 1. Une telle mesure, puisqu'elle permet des ajouts ou des suppressions de cellules rythmiques, pourrait permettre de gérer aussi le problème que pose la comparaison de formules rythmiques qui n'ont pas le même nombre de cellules.

Similarité entre cellules rythmiques. Les mesures de similarité entre formules rythmiques proposées jusqu'ici comparent les numéros des cellules rythmiques contenues dans les diverses formules. Elles décrivent donc simplement si ces cellules sont identiques ou différentes, là encore sans nuance entre ces extrêmes. On voit pourtant dans le tableau 8.1 que la dernière cellule de la formule *I.a* (« 3 », ) et celle de la formule *I.b* (« 4 », ) sont assez similaires. Lorsque l'on compare deux formules, si une cellule diffère, la mesure de distance pourrait prendre en compte le degré de similarité entre ces deux cellules.

En effet, on sait que certaines cellules rythmiques peuvent *commuter* entre elles, c'est-à-dire qu'elles peuvent être utilisées indifféremment dans un énoncé rythmique donné sans que celui-ci ne soit jugé différent par les interprètes. Ce jugement culturel d'équivalence, s'il est partagé à une échelle géographique suffisamment large, pourrait être utilisé pour déterminer la similarité entre différentes cellules rythmiques.

Ainsi, dans leur étude des musiques Myènè, Le Bomin et Bikoma [238] proposent que le passage entre la cellule  et la cellule  puisse se faire via des cellules intermédiaires, comme illustré dans la figure 8.1. Cette équivalence entre cellules rythmiques semble partagée de manière assez large (au moins dans les différentes populations Myènè). Elle pourrait donc servir à définir une distance entre deux cellules. Par exemple les cellules  et  auraient une distance de 1,

25. ↑ *Edit-distance*, en anglais.

26. ↑ Toussaint [422].

27. ↑ C'est ce que propose la distance de Hamming relâchée (*fuzzy Hamming distance*) [A48]²⁸.

28. Citée par Toussaint [422].

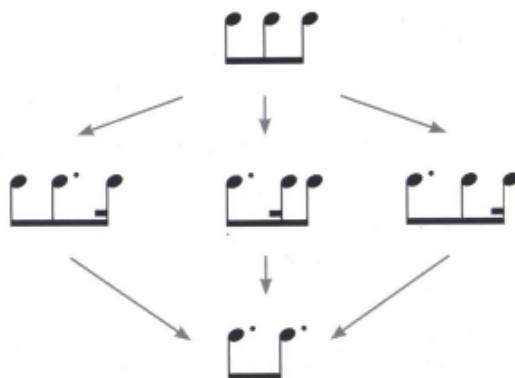


FIG. 8.1 : **Exemple de passage entre deux cellules rythmiques.** Dans les musiques Myènè, le passage entre la cellule rythmique ♪♪♪ (en haut) à tendance ternaire à la cellule ♪♪ (en bas) à tendance binaire peut se faire par l'intermédiaire de cellules intermédiaires maintenant l'ambiguïté entre ces deux modalités de subdivision de la pulsation : « Ces cellules rythmiques binaires et ternaires ne sont cependant pas très éloignées les unes des autres et le passage d'une forme à l'autre peut se faire par des modifications subtiles des valeurs. » [238, p. 43]. Cette figure est issue de l'ouvrage de Le Bomin et Bikoma [238, p. 44].

puisque l'on passe directement de l'une à l'autre, tout comme les cellules ♪♪♪ et ♪♪ . Les cellules ♪♪♪ et ♪♪ auraient une distance de 2, puisqu'il faut deux étapes pour passer de l'une à l'autre, par l'intermédiaire de la cellule ♪♪♪ (ou de l'une des deux autres).

8.2.2.2 Décrire la rythmique indépendamment de la métrique

Dans le codage utilisé²⁹, le code attribué à une cellule rythmique est différent selon que la pièce d'où elle provient a une pulsation de subdivision binaire ou de subdivision ternaire. Ce choix peut cependant être problématique dans le cas des musiques du Gabon où il n'est pas toujours facile ni possible d'identifier une modalité unique de subdivision de la pulsation, comme illustré sur la figure 8.2 et détaillé ci-dessous.

Une même pulsation avec plusieurs subdivisions possibles. Une première raison à cela vient de ce qu'une même pulsation peut être interprétée de manière binaire ou ternaire si l'on ne considère que les parties de tambour. Comme exposé plus haut, il est possible dans les musiques Myènè que « des cellules rythmiques de caractères binaire ou ternaire se succèdent plus ou moins régulièrement dans la réalisation d'une phrase » [238, p. 43], ce qui permet de « maintenir l'ambiguïté entre binaire et ternaire » [238, p. 44]. Ainsi, en s'intéressant aux formules rythmiques les plus fréquentes chez les Myènè, ♪♪♪♪ et ♪♪♪♪ , Le Bomin et Bikoma [238, p. 45] constatent que « [c]es deux formules n'en constituent qu'une seule en faisant alterner un premier temps divisé par trois et un second par deux ». De la même manière, on voit dans l'exemple de la figure 8.2a qu'une même formule rythmique peut être codée par des cellules rythmiques différentes si la

29. ↑ Voir le tableau des cellules rythmiques dans le chapitre 6, page 226.

pulsation est interprétée de manière binaire ou ternaire. Dans ce cas, c'est généralement l'analyse des parties chantées qui permet de lever l'ambiguïté et d'identifier la modalité de subdivision de la pulsation [238, p. 45]. Susanne Fürniss [151] a cependant montré que, lors d'un emprunt, une formule rythmique pouvait rester inchangée alors même qu'elle est introduite dans un nouveau cadre métrique. En codant différemment les cellules rythmiques selon que la pulsation soit binaire ou ternaire, il est impossible d'identifier un tel emprunt. Dans une perspective comparative, il serait donc intéressant de réfléchir à un système de codage pouvant permettre ce rapprochement.

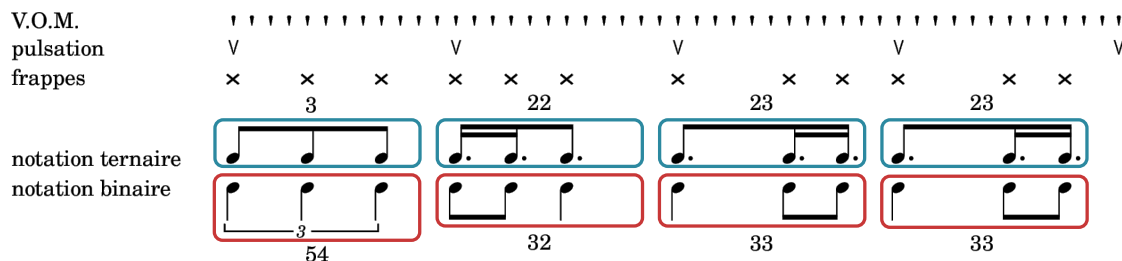
Plusieurs pulsations possibles. Il arrive aussi que, pour une même pièce, différentes pulsations puissent être indifféremment marquées. Cela a été décrit dans les musiques des Téké et des Tsogho par Le Bomin [236, pp. 22–23] :

« Au Gabon, à la différence de ce qui a été décrit en République centrafricaine, j'ai pu observer que pour un même chant une subdivision binaire et une subdivision ternaire de la pulsation peuvent se superposer pour une même partie de la polyphonie ; c'est le cas pour certains chants du *bwiti* des Tsogho. Chez les Teke, cette superposition a lieu entre plusieurs parties dans différents répertoires. »

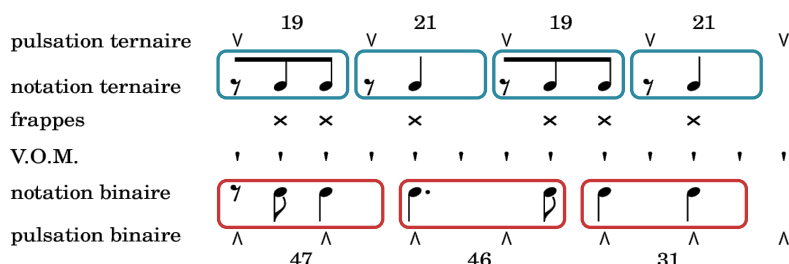
Le Bomin (2004) [236, p. 22]

Cette observation a été reproduite par Lechaux [249, pp. 349–350] dans sa thèse de doctorat sur les musiques tsogho. Sylvie Le Bomin et moi-même avons pu de même vérifier la possibilité de cette polymétrie chez ces mêmes populations tsogho. Lors des enregistrements en contexte, les deux pulsations sont réalisées simultanément. Lors des **enregistrements analytiques**, il est en effet possible de demander à la personne qui frappe la pulsation de changer ses frappes pour obtenir une pulsation soit de subdivision ternaire, soit de subdivision binaire. Il faut noter que ce changement de battue ne lui demande pas le moindre effort et ne pose aucun problème aux interprètes réalisant les parties de tambour ou les parties chantées. Ces deux pulsations sont considérées être « la même chose » par les interprètes, suggérant qu'elles constituent ensemble le cadre **métrique** des pièces étudiées.

L'exemple de la figure 8.2b montre que le codage d'une même formule rythmique sera là encore différent selon la pulsation frappée. Si la pulsation est frappée toutes les trois **valeurs opérationnelles minimales**, on obtient la formule « 19, 21, 19, 21 ». Si la pulsation est frappée toutes les deux valeurs opérationnelles minimales, on obtient la formule « 47, 36, 31 ». Ces deux formules sont présentes dans notre jeu de données, la première venant des Bakoya de Imbong et la seconde des Ngom de Moualo. Il serait intéressant de pouvoir demander aux interprètes de ces deux cultures musicales si ces formules peuvent être considérées comme équivalentes ou non.



(a) **Exemple de codage d'une formule rythmique selon deux interprétations d'une même pulsation.** Dans cet exemple, une seule **pulsation** est possible mais sa subdivision peut être interprétée soit de manière **binaire**, soit de manière **ternaire**. La notation rythmique d'une même série de frappes sera alors différente. En conséquences, le codage de la formule rythmique sera lui aussi différent : « 3, 22, 23, 23 » si la pulsation est interprétée de manière ternaire (en haut) ou « 54, 32, 33, 33 » si elle est interprétée de manière binaire (en bas).



(b) **Exemple de codage d'une formule rythmique selon deux pulsations différentes.** Dans cet exemple, deux pulsations différentes peuvent être marquées (indiquées par des chevrons), l'une de subdivision **ternaire** constituée de 3 **V.O.M.** (en haut), l'autre de subdivision **binaire** constituée de 2 **V.O.M.** (en bas). Dans le cas d'une pulsation subdivisée de manière ternaire, la formule rythmique sera codée « 19, 21, 19, 21 », comme dans le cas de la pièce RAEI146 (Bakoya - Imbong #11). Dans le cas d'une pulsation subdivisée de manière binaire, la formule rythmique sera codée « 47, 46, 31 », comme dans le cas de la pièce RAEI115 (Ngom - Moualo #5).

FIG. 8.2 : **Influence de la subdivision de la pulsation sur le codage des formules rythmiques.** Dans chacun de ces deux exemples, la manière dont est subdivisée la pulsation peut influencer le codage de la **formule rythmique** (c'est-à-dire la suite de codes des **cellules rythmiques**) décrivant une même série d'événements rythmiques. Les cellules rythmiques subdivisées de manière ternaire et binaire sont encadrées respectivement en bleu et en rouge. Les codages correspondants dans le tableau des cellules rythmiques (chapitre 6, page 226) sont indiqués par les chiffres positionnés respectivement au-dessus et au-dessous du cadre. La ligne « frappes » indique la position des événements rythmiques (dans notre cas des frappes de tambour). La ligne « V.O.M. » indique la position des **valeurs opérationnelles minimales**.

Chapitre 9

Réflexions sur la pratique de la recherche

9.1	La relation aux données	297
9.1.1	Quand commence l'archive?	297
	Un jeu de données homogène	297
	Des données de seconde main	298
9.1.2	De nombreux implicites	298
	9.1.2.1 Implicite car évident	298
	9.1.2.2 Collecter pour soi	299
	9.1.2.3 Implicite car ineffable	299
9.1.3	Travailler avec l'implicite	299
	9.1.3.1 De l'ignorance à la méfiance	299
	9.1.3.2 Directrice et informatrice	300
9.2	Le rôle d'analyste	301
9.2.1	Analyser des données préparées par d'autres	301
	Une relation interdisciplinaire	301
	Une interdisciplinarité asymétrique	301
9.2.2	Un analyste déconnecté du terrain	302
	9.2.2.1 Scinder l'étape d'analyse?	302
	La transcription est analyse	302
	L'analyse doit être globalement cohérente	302
	9.2.2.2 Séparer collecte et analyse?	303
	Spécialisation de l'analyste	303
	L'analyse liée au terrain	303
9.3	Limites de l'interdisciplinarité	304
9.3.1	Les « points aveugles » de la coopération interdisciplinaire	304
	Des zones délaissées	304
	Occultées par les différences disciplinaires	304
9.3.2	Le besoin d'unifier	305
	De la coopération à l'intégration	305
	Des ressources pour le travail interdisciplinaire	305

9.3.3	L'individualisation de la discipline	306
9.4	La transposition de méthodes	306
9.4.1	Légitimité de la transposition	307
9.4.1.1	Des raisons	307
	L'analogie a une valeur heuristique	307
	Sur les conditions de naissances de l'analogie	308
9.4.1.2	Des justifications	309
9.4.2	Évaluation	310
9.4.2.1	De la validation rationnelle	311
9.4.2.2	Des transformations induites	311
	Conceptions portées par la méthode	311
	Compétences requises	312
9.4.2.3	De la forme	312

Ce dernier chapitre vise à prendre un peu de recul sur le parcours de mes recherches au cours de ce doctorat et de réfléchir aux difficultés auxquelles j'ai été confrontées. Cet exercice, très enrichissant pour moi, me paraît intéressant à partager car les problématiques que j'y aborde ne sont généralement pas spécifiques à mon sujet et pourront peut-être faire écho à d'autres expériences similaires. En effet, elles sont liées aux enjeux des recherches interdisciplinaires et, de manière plus générale, à l'organisation sociale de la recherche. On pourrait interpréter à tort cette partie, de par mon goût parfois trop prononcé pour l'abstraction et la théorisation, comme une mauvaise analyse de sociologie des sciences. Je n'ai pas la prétention d'effectuer une telle entreprise, car je n'ai ni les compétences, ni le bagage théorique nécessaires pour la mener à bien. Il faut plutôt voir ce chapitre comme un retour d'expérience qui vise avant tout à me donner les moyens de conscientiser et formaliser certaines problématiques en espérant que cela puisse m'aider à les surmonter et à améliorer mes pratiques de recherche dans le futur.

Un tournant important de mon doctorat fut de me rendre compte du rapport particulier qu'entretient chaque personne avec les données analysées¹. Cela m'a fait prendre conscience des problèmes que posait le fait d'analyser des données que je n'avais pas moi-même collectées². Cette séparation entre collectage et analyse informatisée des données devenait de fait une situation de travail interdisciplinaire dont j'ai pu entrevoir certaines limites³. Enfin mes interrogations quant à la légitimité de mes recherches m'ont amené à réfléchir plus généralement aux problèmes posés par l'utilisation dans une discipline (l'*ethnomusicologie*) de méthodes développées par une autre discipline (la *systématique biologique*)⁴.

9.1 La relation aux données

9.1.1 Quand commence l'archive ?

Un jeu de données homogène. Les données qu'il m'a été donné d'analyser au cours de cette thèse⁵ ont été entièrement collectées et transcrites par Sylvie Le Bomin, ma directrice de thèse. Compte tenu à la fois de l'ampleur de la zone géographique étudiée et du détail de l'analyse (souvent plus d'une dizaine de pièces pour une même population), c'est une qualité rare pour un jeu de données que d'avoir été entièrement constitué par une seule et même personne, selon une perspective et une méthodologie cohérentes et unifiées. Ce jeu de données, présenté comme issu d'un corpus de première main homogène du point de vue de la collecte et de la transcription, devait donc constituer des bases stables pour mes recherches et me permettre d'avancer vite dans un contexte où les thèses doivent être réalisées dans un temps relativement court, y compris en ethnomusicologie.

1. ↑ Voir [partie 9.1](#) pages 297 à 300.

2. ↑ Voir [partie 9.2](#) pages 301 à 304.

3. ↑ Voir [partie 9.3](#) pages 304 à 306.

4. ↑ Voir [partie 9.4](#) pages 306 à 313.

5. ↑ Et en particulier dans le chapitre 6 pages 187 à 245.

Des données de seconde main. Ce n'est que progressivement, à travers les problèmes que j'ai pu rencontrer au cours de son analyse, que je me suis rendu compte que ce corpus n'était *pas* de *première* main. En effet, dès lors que ce n'est plus Sylvie Le Bomin (la collectrice) qui prend en charge son analyse, il devient un corpus de *seconde main* et soulève des questions semblables à tout jeu de données collecté par d'autres et, en particulier, les problématiques posées par ce qui peut être considéré comme un « terrain d'archive »⁶ [249, pp. 86–87].

9.1.2 De nombreux implicites

La première et peut-être la plus importante de ces problématiques est sans doute tout l'implicite et l'ineffable de la collecte.

9.1.2.1 Implicite car évident

Pour une partie il s'agit d'éléments de contexte qui vont de soi pour la collectrice et qui à ce titre ne sont pas mentionnés dans les données ou dans la description des méthodes. Par exemple Sylvie m'a expliqué, au commencement de ma thèse et alors que je découvrais les méthodes de l'ethnomusicologie, que les enregistrements dont provenaient les données étaient réalisés par la méthode du *re-recording*. Développée dans les années 1960 afin d'analyser les polyphonies et les polyrythmies d'Afrique centrale, cette méthode est décrite en détail par Arom [13, pp. 190–213]. J'ai donc pensé tout naturellement que c'était cette technique qui avait été utilisée par Sylvie Le Bomin pour enregistrer les musiques qui m'intéressaient. Ce n'est que plus tard, en fait aux premières écoutes des enregistrements analytiques, que j'ai réalisé que la méthode était différente. Interrogé à ce sujet, Sylvie m'a expliqué qu'il s'agissait de la méthode d'*enregistrement continu des parties séparées* (EnConPaS), imaginée par Sylvie Le Bomin et Susanne Fűrniß dans les années 2000. L'utilisation de cette méthode, permettant de gagner en efficacité sur le terrain, n'a été formellement décrite dans aucune publication⁷. Elle fait partie du bagage implicite des ethnomusicologues de l'équipe *systématique et catégorisation culturelle* (SCC), qui ne s'apprend vraiment qu'au contact du terrain. C'est d'ailleurs à l'occasion de mes propres collectages que j'ai pu vraiment me rendre compte du déroulement réel d'une séance d'EnConPaS.

Cela m'a non seulement permis de mieux comprendre les enregistrements auxquels j'avais accès, mais aussi de me faire une meilleure idée de leurs limites éventuelles. Je peux maintenant mieux évaluer quels sont les éléments dans lesquels j'ai une plus grande confiance. Par exemple j'ai pu constater que la technique d'EnConPaS déstabilise légèrement les interprètes lors des premiers enregistrements⁸, ce qui n'est plus le cas après quelques pièces enregistrées. Il existe ainsi de nombreux éléments implicites (évidents pour la collectrice), dont il est nécessaire de

6. ↑ Ces enjeux sont décrits dans sa thèse par Émeline Lechaux [249, pp. 86–87].

7. ↑ Émeline Lechaux [249, p. 82] décrit succinctement cette méthode dans sa thèse de doctorat. J'en propose un schéma explicatif dans le chapitre 6, page 223.

8. ↑ Pour éviter ce biais, une première pièce (non enregistrée) est jouée en simulant un EnConPaS, afin de familiariser les interprètes avec cette technique. Néanmoins la déstabilisation perdure parfois encore au cours des premiers enregistrements.

tenir compte (du moins d'avoir conscience) au cours de l'analyse, par exemple la dénomination endogène des pièces, l'ordre d'intervention des interprètes, ou encore le nommage des fichiers informatiques.

9.1.2.2 Collecter pour soi

Un certain nombre de ces implicites vient de ce que les données ont été collectées par Sylvie dans une perspective où c'est elle-même qui les analyserait. Dès lors un certain nombre d'éléments contextuels ne sont pas nécessaires à expliciter puisqu'ils ne lui sont pas utiles et peuvent être retrouvés indirectement. Ainsi Sylvie est-elle capable de reconnaître, de par le style musical mais aussi par l'ambiance sonore ou par les voix identifiées, le contexte de la plupart de ses enregistrements.

9.1.2.3 Implicite car ineffable

Une partie des informations implicites ne peuvent cependant être explicitées car il est très difficile, voire impossible de les retransmettre. Il s'agit de toutes les rencontres informelles sur le terrain, le vécu des relations humaines (à la croisée d'une multitude de rapports sociaux entre l'ethnomusicologue et les interprètes), les intuitions, les réactions spontanées ou les sentiments et émotions. Autant d'éléments qui, soit que l'on n'en ait pas conscience, soit qu'il soit difficile de les partager, ne sont pas accessibles à l'analyste de données collectées par d'autres. Bien que la méthode scientifique cherche souvent à minimiser leur importance par des procédures neutralisant leurs effets, ils ont pourtant un rôle déterminant dans la compréhension et l'acquisition de connaissances, ne serait-ce que par leur rôle heuristique, qui guide la démarche de recherche. Ils ne peuvent en tout cas pas être négligés. Je possède moi-même ma propre part d'implicite, issue de mon expérience et notamment de ce « terrain » particulier qu'est ma relation avec ma directrice et ses données. C'est d'ailleurs sans doute dans le décalage entre l'implicite de Sylvie et mon propre implicite que réside la différence de perception de données en apparence identiques. Ainsi Sylvie et moi n'auront pas la même confiance dans une même observation puisqu'elle aura été vécue dans un cas comme une expérience de vie pendant plusieurs mois dans des régions très distinctes, dans l'autre comme un déchiffrement de notes manuscrites et l'analyse de données informatiques.

9.1.3 Travailler avec l'implicite

9.1.3.1 De l'ignorance à la méfiance

Cette faible connaissance⁹ des conditions de collecte me conduisent à avoir une certaine méfiance envers les données. En effet, il n'est pas toujours possible d'avoir la même confiance dans l'ensemble des données. Par exemple, lorsque nous enregistrons chez les Eshira de Libreville,

9. ↑ Ou plutôt ma faible expérience, c'est à dire le fait de n'avoir pas moi-même *vécu* la collecte de ces données.

nous avons pu constater que les interprètes n'avaient qu'une faible connaissance des pièces réalisées et ces enregistrements pouvaient donc difficilement être considérés comme représentatifs de leur village d'origine. En analysant des données de seconde main, je n'ai pas accès à ce type d'information implicite qui permet d'évaluer la fiabilité relative de différentes données. Ainsi, ne connaissant pas finement les limites de mes données, les éléments solides et ceux plus instables, je préfère exercer une réserve *a priori* sur l'ensemble.

Cette attitude et sans doute salutaire scientifiquement parlant puisqu'elle cherche à éviter une surinterprétation des données ou des conclusions issues d'une mauvaise compréhension. Mais elle conduit aussi à diminuer le savoir que je peux tirer de ces données. Ainsi j'ai choisi d'écarter de mon analyse les informations pour lesquelles je n'ai pas retrouvé l'enregistrement correspondant ou encore toutes celles provenant d'enregistrements conventionnels. Sans doute l'expérience qu'a Sylvie lui aurait permis de choisir avec plus de discernement entre ces données et donner un meilleur aperçu des musiques du Gabon. Cependant mon propre rapport à ces mêmes données et le besoin d'explicitier rationnellement et de manière cohérente le choix des données m'amènent à écarter ces informations.

9.1.3.2 Directrice et informatrice

Prenant conscience de ces nombreux implicites qui me manquaient pour comprendre mes données (et ainsi pour avoir confiance en elles), j'ai questionné Sylvie à de nombreuses reprises afin d'explicitier tous ces éléments invisibles. J'ai enfin pu reconsidérer mes données non plus comme des données de première main, mais comme des données de seconde main pour lesquelles j'avais la chance d'avoir une informatrice de choix (la collectrice) prête à me fournir toutes les informations dont j'aurais pu avoir besoin. Dès lors ma directrice devenait aussi mon informatrice et nos séances peuvent être vues tantôt comme des réunions de travail, tantôt comme une enquête de terrain. Les documents qu'elle a mis à ma disposition deviennent des sources qu'il s'agit de décrire et d'indexer afin de pouvoir m'y référer. J'ai essayé tant que possible de caractériser chaque source potentielle d'information à ma disposition (par exemple un carnet de terrain de Sylvie, l'annonce précédant un enregistrement analytique, mais aussi un courriel ou mes notes issues d'une réunion de travail) afin de pouvoir sourcer chacune de mes données à l'aide de ses identifiants. Cette prise de conscience a cependant été relativement tardive et ce travail de sourçage n'est pas complètement terminé. Je ne peux donc pas présenter ces données ici, chose malheureuse puisqu'elles seraient utiles non seulement à Sylvie ou moi-même mais aussi aux ethnomusicologues désirant travailler sur ces données dans le futur.

9.2 Le rôle d'analyste

9.2.1 Analyser des données préparées par d'autres

Une relation interdisciplinaire. L'émergence de cette question des « données de seconde main » peut d'une certaine manière trouver son origine dans la construction de mon projet de thèse. Sylvie ayant rassemblé un important volume de données, il m'a été confié la tâche de les analyser, ma formation aux méthodes d'analyse de données issues de la biologie devant me fournir les outils pour le faire. La relation avec ma directrice est alors devenue, en plus d'une relation d'encadrement, une relation de travail interdisciplinaire, entre une ethnomusicologue et un « analyste de données ». En effet, malgré mes deux ans de formation en musicologie classique, mes connaissances en ethnomusicologie étaient alors très faibles et ne furent acquises qu'au cours de mes recherches. Cette relation est interdisciplinaire en ce sens qu'elle implique des scientifiques de disciplines différentes mobilisant des méthodes et des paradigmes différents. Cette relation particulière a bien entendu des conséquences sur la pratique de la recherche et je propose de réfléchir ci-dessous à quelques-unes d'entre elles.

Une interdisciplinarité asymétrique. Cependant c'est une interdisciplinarité particulière puisque les disciplines ne sont pas mobilisées de manières symétriques : d'abord Sylvie, ethnomusicologue aguerrie, a collecté et transcrit un ensemble de pièces puis en a constitué une matrice de données. J'étais ensuite supposé analyser ces données avec les méthodes tirées de la biologie (et que j'étais sensé maîtriser). Avec le recul cette approche me semble être problématique et pose la question de la place de la collecte dans la recherche, du statut de la transcription et plus généralement de la possibilité même de « l'analyste professionnel ». Les données analysées dans cette thèse sont issues d'enregistrements effectués par Sylvie entre 2005 et 2009¹⁰. Je n'ai donc pas participé à la réflexion, au choix et à la pratique du collectage des musiques que j'ai analysées.

Pourtant ces étapes font partie intégrante de la conduite de la recherche et doivent permettre d'éclairer au mieux le sujet traité. Les collectages effectués par Sylvie ne s'inscrivaient pas dans la problématique à la lumière de laquelle ils sont aujourd'hui analysés : la perspective principale d'alors était la caractérisation de la **systématique musicale** des populations visitées dont les musiques étaient alors pas ou peu connues. Ce n'est que récemment que la question de l'inférence historique par comparaison de ces musiques a vraiment émergée comme thématique de recherche. On voit que la perspective dans laquelle sont analysées les données est amenée à évoluer, modifiant l'adéquation de ces données avec la problématique posée. Ainsi Sylvie, guidée par la poursuite de l'analyse comparative de ses musiques, cherche depuis 2007 à noter la présence ou l'absence des différents répertoires ou des instruments de musique qu'elle a pu rencontrer au Gabon dans chacune des populations qu'elle visite, afin de combler les lacunes de sa matrice de données

10. ↑ J'ai moi-même participé à une mission de collectage de données d'un mois, mais les données qui en sont issues n'ont pas encore été analysées.

(puisque les premiers collectages n'avaient pas été effectués dans ce but).

9.2.2 Un analyste déconnecté du terrain

9.2.2.1 Scinder l'étape d'analyse ?

La transcription est analyse. La séparation interdisciplinaire entre collecte et transcription d'un côté et analyse statistique de l'autre vient se confronter au fait que l'étape de transcription, loin d'être une simple étape technique, fait partie intégrante de l'analyse. Séparer en deux étapes radicalement disjointes la transcription et l'analyse statistique des données qui en sont issues, et même les séparer en deux personnes distinctes, peut aussi être problématique.

L'analyse doit être globalement cohérente. Ainsi la codification des rythmes élaborés par Sylvie l'avait été dans la perspective d'une caractérisation d'une culture musicale. Mes propres recherches s'étant dirigées vers une description comparative des musiques de l'ensemble du Gabon, mes choix d'analyse lors de l'étape de transcription auraient pu être différents. Par exemple, Sylvie a choisi de coder les **cellules rythmiques** de manière différentes selon que la métrique de la pièce soit de **subdivision binaire** ou **ternaire**. Dès lors, il est impossible de comparer l'organisation rythmique de deux pièces ayant une pulsation de subdivision l'une ternaire et l'autre binaire. En effet, les cellules étant codées de manière différentes selon ces deux organisations métriques, on ne pourra pas identifier de cellule commune, en raison même du choix de cette représentation.

Pourtant un cadre métrique différent n'exclut pas toute parenté de rythme. Susanne Fürniss [151] a par exemple montré que la formule polyrythmique du répertoire du *bèkà* des pygmées Baka du Cameroun a été empruntée au Kwélé Djako, mais en l'intégrant dans le cadre métrique des Baka, comportant invariablement 4 pulsations. Il est donc possible que deux **formules rythmiques** soient historiquement apparentées, même si chacune d'elle est associée à un cadre métrique différent. En codant différemment les cellules rythmiques selon que la pulsation soit binaire ou ternaire, il est impossible de détecter une telle parenté. Aussi, dans une perspective comparative il aurait été intéressant de pouvoir décrire l'organisation rythmique de manière relativement indépendante de l'organisation métrique¹¹.

En outre, comme on l'a vu¹², la subdivision de la pulsation montre une nette organisation géographique : la subdivision de la pulsation de la plupart des pièces enregistrées dans la partie est du Gabon est ternaire, alors qu'elle est binaire dans celles provenant de l'ouest du pays. Comme les cellules rythmiques sont codées selon la subdivision de la pulsation, cette organisation géographique se retrouve aussi dans la distribution des cellules. Ainsi la cellule 23 ($\underline{\text{♩. ♩. ♩}}$), utilisée pour coder les cellules rythmiques des pièces dont la pulsation a une subdivision ternaire, est trouvée uniquement dans la partie est du pays. Inversement la cellule 46 ($\underline{\text{♩. ♩}}$), utilisée lorsque la pulsation a une subdivision binaire, est trouvée presque exclusivement dans la partie ouest. Ceci montre que la manière dont sont codées les cellules rythmiques, en fonction de la métrique,

11. ↑ Voir [partie 8.2.2](#) pages 288 à 292.

12. ↑ Voir [chapitre 6](#), pages 191 à 245.

impacte l'analyse comparative qui en est faite. On voit donc que la deuxième partie de l'analyse (l'analyse comparative) peut amener à questionner la manière dont est faite la première partie (le codage des cellules).

9.2.2.2 Séparer collecte et analyse ?

Spécialisation de l'analyste. L'augmentation des données informatiques issues des travaux de recherche¹³ d'une part et l'augmentation du nombre et de la puissance des méthodes pour les analyser d'autre part conduit aujourd'hui à la spécialisation de certains et certaines scientifiques en « analystes professionnels » dont le terrain se limite à leur ordinateur et qui se chargent d'analyser grâce aux outils informatiques des données produites par d'autres. Le phénomène me semble particulièrement présent dans l'étude comparative et quantitative des données culturelles, qui constituent le cœur de cette thèse, facilité par la séparation entre des disciplines rompues aux méthodes du terrain et d'autres à celles de l'analyse informatisée. C'est par exemple mon cas, biologiste de formation, ou encore celui de Godfried Toussaint, peut-être l'un des premiers à introduire des *méthode issue de l'inférence phylogénétique* pour étudier la musique¹⁴, qui est chercheur en informatique [420].

Dans un contexte où la recherche scientifique est contrainte par la difficulté d'obtention des financements d'une part et les critères scientométriques poussant à multiplier les publications d'autre part, une thèse se concentrant sur l'analyse et négligeant le travail de terrain a l'intérêt d'être à la fois peu coûteuse et relativement rapide puisqu'elle se base sur des données déjà existantes.

L'analyse liée au terrain. Si l'intérêt stratégique de telles publications est compréhensible, on peut cependant questionner son impact du point de vue scientifique. En effet la pratique même de l'enquête ethnomusicologique aussi bien que les différentes étapes d'analyse et de construction des données (y compris la transcription) modèlent et guident les questions de recherche. J'ai pu constater à quel point la courte mission que j'ai effectuée au Gabon a changé la perception de mon sujet d'étude. Intervenue tardivement (à la fin de ma deuxième année de doctorat) elle a transformé des questionnements abstraits dont je ne trouvais le sens qu'en me référant constamment à ma directrice, en des questionnements et intuitions concrètes : Pourquoi deux enregistrements sont-ils considérés équivalents par deux cultures alors que tout, à mon oreille, les différencie ? Comment relier le discours, très précis, sur l'origine et les échanges musicaux tenus par les interprètes avec leur production musicale et l'histoire de ces populations ? Si l'on peut considérer que « l'analyste professionnel » dont j'ai en quelque sorte tenu le rôle pendant ma thèse, a son propre terrain et ses propres problématiques dans les données informatiques que constitue son corpus, cette séparation entre production et analyse des données informatiques me paraît scientifiquement dommageable et il nous revient, à mon avis, de revendiquer les moyens

13. ↑ Ou du moins utilisables dans une perspective de recherche.

14. ↑ Voir [partie 2.1.5.2](#) page 71.

matériels de conduire nos recherches tout en évitant cet écueil.

9.3 Limites de l'interdisciplinarité

9.3.1 Les « points aveugles » de la coopération interdisciplinaire

Des zones délaissées. On voit ici se dessiner certaines limites de l'interdisciplinarité. Lorsque plusieurs personnes, au bagage méthodologique et épistémologique différent, s'associent pour aborder un même sujet, l'unité et la cohérence globale ne peut pas être aussi forte sans doute que lorsqu'une seule personne maîtrise l'ensemble de l'analyse.

Reprenons l'exemple de l'utilisation des *méthodes issues de l'inférence phylogénétique* pour analyser des données musicales. La méthode *cladistique* est basée sur le principe de *congruence* : on fait l'hypothèse que la plupart des *caractères* analysés ont la même histoire. Même si cette hypothèse est sans doute fautive pour certains d'entre eux, le choix d'un grand nombre de caractères permet de juger, grâce au principe de *parcimonie*, que l'arbre obtenu reflètera cette histoire principale. Dans un cadre interdisciplinaire où le jeu de données est construit par l'ethnomusicologue et la méthode conçue par un ou une spécialiste de l'*inférence phylogénétique*, cette notion centrale de congruence risque pourtant d'être un « point aveugle » de la coopération scientifique. Pour l'ethnomusicologue, la notion de congruence peut-être vue comme intrinsèque à la méthode et il revient donc à l'analyste de s'assurer de la pertinence de cette méthode compte tenu des données fournies. Pour l'analyste, en revanche, le respect du critère de congruence est à la charge de l'ethnomusicologue puisqu'il s'agit de s'assurer que les différents caractères choisis soient en accord.

Occultées par les différences disciplinaires. Cette zone « aveugle » pourrait même être occultée un peu plus par les différences des terminologies et des conceptions disciplinaires. Ainsi la notion de *caractère* en biologie semble équivalente à celle de *paramètre* en ethnomusicologie, de même que la notion liée d'*état* de caractère semble équivalente à celle de *valeur* de paramètre. Pourtant le caractère en biologie, défini dans le cadre de l'*inférence historique*, doit être présumé partagé pour des raisons historiques :

« **Caractère** n.m. Ensemble d'attributs sur lesquels on peut émettre au moins une hypothèse d'homologie (homologie primaire).

[...]

Homologie (caractère homologue) n.f. [...] 2. similarités chez une ou plusieurs espèces, d'organes ou de parties d'organes, lorsqu'on peut présumer que cette correspondance provient de l'héritage d'un ancêtre commun. »

Lecointre et Le Guyader (2001) [251, pp. 539–540]

Le **paramètre** de l'ethnomusicologue en revanche appartient aux outils de la **catégorisation musicale** et ne s'inscrit pas nécessairement dans une optique historique. Ainsi la notion de paramètre ne fait pas appel à une notion équivalente de celle de l'**homologie** et le partage des valeurs de paramètre n'est pas nécessairement imputable à une cause historique. En effet, lorsque l'on définit un caractère en biologie, cela revient à faire l'hypothèse *a priori* que la ressemblance n'est pas due à une convergence, ce qui n'est pas le cas en systématique musicale. Ainsi la similarité apparente de ces notions, en facilitant le dialogue interdisciplinaire, cache des différences plus profondes et peut occulter le problème de la « zone aveugle » de la **congruence**. Si la notion de *paramètre* est équivalente à celle de *caractère* il n'y a pas de raison de douter de l'hypothèse de congruence des données. En effet, on peut penser que l'hypothèse de congruence découle de ce que les caractères sont supposés homologues : semblables de part leur histoire commune, l'utilisation d'un grand nombre de caractères devrait permettre de retrouver cette histoire¹⁵.

9.3.2 Le besoin d'unifier

De la coopération à l'intégration. La coopération interdisciplinaire peut donc engendrer des zones de faiblesse, moins robustes méthodologiquement et épistémologiquement¹⁶. On comprend le besoin qu'une personne puisse avoir une vue d'ensemble de l'analyse et puisse juger de la cohérence globale de celle-ci. C'est ce qui m'a conduit à délaissé le seul rôle d' « analyste » pour me former aux méthodes et conceptions aussi bien de l'**ethnomusicologie** que de l'**inférence phylogénétique**. J'ai aussi eu le besoin de comprendre l'origine et les modifications qu'avaient subies les données que j'avais sous la main. Remonter à partir d'un code dans une case de la matrice de données à la **valeur** de **paramètre** identifiée par Sylvie lors de ses transcriptions et comprendre comment elle avait été obtenue. C'est ainsi que j'en suis venu à me demander si l'hypothèse d'homologie des paramètres était raisonnable et affectait les analyses présentées dans le chapitre 5¹⁷.

Maitriser conjointement deux disciplines très différentes demande cependant beaucoup de temps¹⁸. Il faut en outre réussir à concilier les cadres épistémologiques, mais aussi lexicaux, des deux disciplines, alors même qu'on ne se sent maitriser aucune d'entre elles.

Des ressources pour le travail interdisciplinaire. Le cadre du département HNS du MMHN m'a été sur ce point très profitable, car il m'a permis de présenter mes travaux à des

15. ↑ On voit aussi que le lien entre hypothèses d'**homologie** et de **congruence** repose sur le fait qu'il existe en biologie, dans la grande majorité des cas, une manière largement dominante de transmettre la ressemblance (la filiation). En ce qui concerne la musique il en existe plusieurs (souvent schématisées en **transmission horizontale** et **transmission verticale**). Cette différence de conception a sans doute aussi un rôle dans le problème du point aveugle de la convergence, mais je ne le détaillerai pas plus ici, l'exemple présenté me paraissant suffisamment parlant.

16. ↑ On peut relier ce problème à ce que Alix Levain appelle plus généralement la « perte des conditions de la robustesse » (communication personnelle).

17. ↑ Pages pages 155 à 185.

18. ↑ En outre cette formation doit se faire en autonomie puisqu'il n'existe pas d'institution dédiée. Le risque est alors de se concentrer aux seuls éléments directement utiles, ce qui ne permet pas nécessairement d'appréhender l'épistémologie sous-jacente de la discipline.

publics des deux disciplines (parfois au cours d'un même exposé), facilitant la distinction entre ce qui relevait des mes erreurs de jeune chercheur d'une part et de conflits interdisciplinaires d'autre part. L'habitude d'interdisciplinarité du laboratoire offre par ailleurs des outils concrets à commencer par le séminaire *Interdisciplinarité* de Frank Alvarez-Pereyre où j'ai pu être confronté à de nombreux exemples mêlant des disciplines variées.

Le groupe de jeunes chercheurs et chercheuses *Atelier terrain et théorie en anthropologie* (AT-TA) a aussi été essentiel par l'espace de parole et d'échanges qu'il offre et parce qu'il permet de mettre en lien les problématiques purement scientifiques avec l'organisation sociale de nos recherches au sein du laboratoire. En outre de nombreuses personnes de ce groupe étaient elles aussi confrontées à une problématique similaire à la mienne : travailler en situation d'interdisciplinarité (c'est à dire avec des scientifiques de plusieurs disciplines différentes) mais aussi avoir la tâche de synthétiser cette interdisciplinarité en une problématique unie, cohérente d'un point de vue méthodologique et épistémologique.

9.3.3 L'individualisation de la discipline

Mais créer un cadre cohérent d'un point de vue des méthodes, des objectifs et des conceptions épistémologiques, n'est-ce pas recréer une discipline ? Une discipline que l'on est seul-e cependant à pouvoir appréhender dans sa globalité, une sorte de « singu-disciplinarité » si l'on veut.

On imagine les problèmes de cette individualisation de la discipline, qui tout en joignant deux disciplines en arrive à leur rupture. La science est fondamentalement un processus collectif, reposant sur la critique des pairs. Si l'on est seul à appréhender d'une manière globale l'ensemble de notre problématique, comment avoir un retour critique global et pas seulement en provenance de l'une ou l'autre des disciplines ? Comment savoir s'il n'y a pas un « point aveugle »¹⁹ que l'on n'a pas considéré, éventuellement car s'y confronter semblait trop ardu ? Si l'interdisciplinarité est sans nul doute source de richesses dans l'exploration scientifique, elle pose de nombreux problèmes pratiques et les méthodes tout comme les institutions sociales²⁰ permettant de les contourner restent encore largement à construire.

9.4 La transposition de méthodes

Ce travail m'a aussi conduit à me questionner sur ma démarche consistant à utiliser des méthodes issues de la **systematique biologique** pour étudier les musiques du Gabon. Ceci m'a amené à réfléchir plus généralement sur les enjeux entourant la réutilisation d'un outil, développé dans le cadre d'une certaine discipline, pour analyser un objet dont l'étude relève traditionnellement d'une autre discipline. C'est ce que j'appellerai ici la *transposition méthodologique interdisciplinaire*, ou plus simplement la *transposition de méthodes*. À travers l'exemple de l'utilisation des

19. ↑ Voir partie 9.3.1 page 304.

20. ↑ J'ai déjà cité certaines de celles qui m'ont aidé dans le paragraphe *Des ressources pour le travail interdisciplinaire* page 305.

méthodes issues de l'inférence phylogénétique, je m'intéresserai d'abord au problème de la légitimité et de la pertinence de la transposition de méthodes²¹, avant d'étudier les enjeux de son évaluation par la discipline où elle est introduite²².

9.4.1 Légitimité de la transposition

Il peut sembler enrichissant de s'inspirer des méthodes proposées par d'autres disciplines pour répondre à une problématique proche. C'est dans cette perspective qu'a commencé mon doctorat, qui devait consister à utiliser des méthodes venues de la *systematique biologique* pour étudier les musiques du Gabon. Mais cela m'a rapidement amené à me poser la question de la *légitimité* d'une telle transposition. En effet, l'ethnomusicologie avait-elle réellement besoin de méthodes venues de la biologie ? N'était-ce pas plutôt moi qui, de par ma formation de biologiste mais ignorant tout des méthodes et de l'histoire de l'ethnomusicologie, décrétait ce besoin pour répondre à une envie qui était la mienne ?

J'en suis venu à me poser la question de l'intérêt scientifique de la démarche : s'il est bien entendu nécessaire que je m'épanouisse dans mes recherches, il est avant tout fondamental qu'elles aient une pertinence scientifique et qu'elles puissent apporter un éclairage nouveau sur mon sujet. Poser la question de la légitimité de la transposition, c'est se demander de manière rationnelle si la méthode peut permettre de mieux comprendre ou, du moins, de comprendre différemment, l'objet de l'étude. La légitimité d'une telle transposition ne va pas de soi et peut être controversée. Ainsi dans la seconde moitié du XX^{ème} siècle, l'idée d'appliquer des modèles structuralistes de la linguistique à l'ethnomusicologie a fait débat et ce fut déjà l'occasion de réfléchir à la légitimité de cette transposition. Dans une réflexion détaillée sur le sujet, Feld [134, p. 201] distingue d'une part des *raisons* d'effectuer cette transposition, c'est-à-dire des arguments heuristiques incitant à explorer cette voie, et d'autre part des *justifications* de cette transpositions, c'est-à-dire des preuves ou avancées significatives montrant les apports effectifs du modèle pour la discipline qui l'accueille. C'est en suivant ces deux types d'arguments, *raisons* et *justifications*, que je propose de réfléchir à la légitimité de la transposition de méthodes.

9.4.1.1 Des raisons

L'analogie a une valeur heuristique. Si l'on attend d'une méthode ou d'une théorie nouvelle une meilleure compréhension, ou du moins différente, il est impossible de savoir avant l'analyse son apport effectif à la discipline. On peut bien sûr avoir des raisons d'écarter une méthode, aussi bien que des raisons de l'essayer, sans pouvoir en affirmer définitivement l'(in)utilité. On voit par ailleurs que la transposition doit d'abord être conçue, l'idée de la transposition méthodologique devant bien entendu précéder sa réalisation. Parmi les raisons possibles capables de susciter la transposition de méthodes, c'est sans doute à l'analogie que l'on a le plus souvent recours [134, pp. 202-208].

21. ↑ Voir [partie 9.4.1](#) pages 307 à 310.

22. ↑ Voir [partie 9.4.2](#) pages 310 à 313.

Ainsi, comme on a pu le voir dans le chapitre 2²³, l'utilisation de *méthodes issues de l'inférence phylogénétique* sur des données culturelles a souvent été justifiée par des raisons analogiques. Par exemple Macé et coll. [276, pp. 256–257] expliquent en détail les ressemblances à la fois entre les objets d'études et entre les méthodes²⁴ de la *critique textuelle* et de la *systématique biologique*. Selon les auteur·e·s, cette observation rend « légitime » [276, p. 256] l'application des méthodes issues de l'inférence phylogénétique à la critique textuelle.

La ressemblance n'est cependant pas un fait matériel, mais bien une construction sans doute venue de notre tendance à repérer des schémas de régularité, aussi l'analogie est-elle subjective. Pour Feld [134, p. 202], « [m]usic is somewhat like and almost the same as many things, depending on how good your imagination is [...] »²⁵ et l'on peut sans aucun doute généraliser ce constat à la plupart des analogies. Dès lors toutes les analogies ont-elles la même valeur heuristique ? Est-il possible d'évaluer à l'avance et rationnellement ses apports éventuels ? Comme l'explique Robert-Demontrond [350, p. 35] :

« Toute comparaison entre objets est vraie sur un nombre indéfini de points, et fautive sur un nombre également indéfini de points. [...] On ne peut définir le degré de pertinence d'une comparaison puisqu'il est toujours un nombre indéfini de dimensions sur lesquelles peut porter le processus de comparaison. »

Robert-Demontrond (2005) [350, p. 35]

Toute analogie est donc légitime, dès lors qu'elle est assumée et présentée en tant qu'approche heuristique plutôt que preuve. C'est sa fécondité, sinon en terme de résultats du moins en termes d'images et de possibilités évoquées, qui suffit à la justifier [350, p. 43].

Sur les conditions de naissances de l'analogie. Il convient alors de s'interroger sur les conditions de son émergence : si toute comparaison peut plus ou moins conduire à une analogie, seules certaines sont pensées comme telles et un plus petit nombre encore sont utilisées dans le travail scientifique. L'analogie doit tout d'abord être perçue, ce qui nécessite pour le ou la scientifique d'être capable de mettre en relations des propriétés homologues d'objets ou de disciplines distinctes, donc d'avoir des connaissances variées et une habitude d'abstraction des concepts permettant d'envisager leur recontextualisation. L'émergence individuelle de l'analogie est bien entendu influencée par les conditions sociales, l'enseignement pluridisciplinaire et l'habitude prise au cours de la formation de transposer des méthodes ou des modèles entre les différents domaines favorisant le nombre et la richesse des comparaisons potentielles.

Le cheminement m'ayant conduit au travail présenté dans cette thèse en est un exemple. Si la possibilité d'étudier l'histoire de la musique à partir des outils et méthodes issues de l'inférence phylogénétique m'a été donnée, c'est sans doute d'abord grâce aux passerelles existantes au

23. ↑ Voir par exemple la partie 2.1.2.3 page 58.

24. ↑ Même si Macé et coll. [276] mettent sur un même plan l'analogie des objets d'études et l'analogie des méthodes, on peut cependant penser que ces deux types d'analogie pourraient avoir un statut heuristique différent.

25. ↑ « [l]a musique est en quelque sorte comme et presque identique à beaucoup de choses, selon la richesse de votre imagination [...] »

cours de ma formation et qui m'ont permis de compléter des études en biologie évolutive par des études de musicologie. C'est aussi grâce à l'existence du département *Homme, Nature, Société*²⁶ qui a fourni, grâce à sa longue tradition interdisciplinaire, un cadre adéquat pour la conduite de cette thèse. C'est enfin parce que, plus généralement, les perspectives interdisciplinaires jouissent d'une aura positive dans le milieu de la recherche française, au moins dans les discours. On voit ainsi comment les structures sociales peuvent favoriser l'apparition d'une analogie, mais aussi les conditions de son développement en une véritable thématique de recherche.

L'analogie peut ensuite être formalisée et approfondie, permettant éventuellement d'améliorer la compréhension de l'objet étudié. Cela implique donc pour la personne à son origine de conférer à cette analogie une valeur heuristique valable scientifiquement, c'est-à-dire qui mérite d'être étudiée et utilisée. La discrimination entre l'analogie à valeur scientifique de celle qui n'en a pas passe donc par un processus d'évaluation et éventuellement d'autocensure, qui dépend bien entendu de l'opinion raisonnée qu'aura la personne sur cette comparaison, mais aussi de sa perception de la recevabilité de celle-ci par la discipline. L'analogie est finalement évaluée par la discipline²⁷. Si son intérêt est reconnu par une partie au moins de la communauté scientifique, elle pourra éventuellement donner lieu à d'autres appropriations et développements au sein de la discipline.

9.4.1.2 Des justifications

Si comme on l'a vu la validité de la transposition ne peut être déterminée *a priori*, il s'agit maintenant de voir comment elle peut l'être *a posteriori*. Le positivisme logique aurait sans doute répondu que l'outil transposé peut-être considéré valide pour la discipline s'il permet de construire une théorie ou un modèle de l'objet meilleur que les précédents, mais on sait bien la difficulté de comparer des théories dont les paradigmes et les conceptions sont généralement incompatibles. Aussi, s'il ne peut permettre de mieux comprendre l'objet étudié, l'outil doit au moins produire des connaissances permettant de voir cet objet sous une vision nouvelle, l'expliquant différemment. En somme, l'outil doit apporter des résultats qui soient pertinents pour la discipline concernée. Cette évaluation rationnelle de la pertinence des résultats permet donc en retour de porter un jugement sur l'outil transposé et donc sur la légitimité de la transposition.

L'obtention de résultats jugés pertinents n'est cependant pas la seule raison scientifiquement valable à la reconnaissance de la validité d'un outil. La transposition et l'utilisation de l'outil, même s'ils ne donnent pas de résultat satisfaisant (c'est-à-dire assez éclairant ou suffisamment novateur), peut apporter une ouverture lui conférant une valeur heuristique accrue. En effet, la manipulation même de l'outil au contact concret des données conduit à modifier la perception des problèmes de la discipline. Ceci guide à son tour vers des perspectives et des idées nouvelles²⁸. Ainsi les premiers essais d'utilisation des méthodes cladistiques à l'échelle des patrimoines musicaux nous ont conduit à nous questionner sur la pertinence de choisir cette unité d'analyse. C'est

26. ↑ Et plus particulièrement de l'UMR 7206.

27. ↑ J'y reviendrai dans la partie 9.4.2 page 310.

28. ↑ Voir la partie 9.4.1.1 page 307.

cette réflexion qui nous a amené à envisager d'étudier les relations historiques entre des *pièces* musicales et qui a guidé la mission de terrain que Sylvie Le Bomin et moi-même avons réalisée²⁹. Impossible de dire dès maintenant si cette piste portera ses fruits, mais on voit qu'une première démarche heuristique amène à ouvrir l'horizon des possibles. C'est donc déjà un résultat en soi, en ce sens qu'il permet de questionner autrement nos connaissances des musiques du Gabon.

Mais cet argument, du type « ce qui n'a rien donné peut donner » peut être retourné : ce n'est pas parce qu'une méthode a produit des résultats pertinents que sa réutilisation aura le même effet. Un outil est toujours utilisé dans un contexte particulier, sur des données particulières. Aussi, les résultats apportés ne peuvent jamais « prouver » son utilité générique. Pourtant, dans toutes les disciplines scientifiques, les outils gagnent en popularité et se répandent en fonction des résultats qu'ils ont permis d'obtenir. Comme l'explique Bourdieu [55, p. 15], il n'existe pas de science « pure » dont les résultats sont indépendants de la société dans laquelle elle se pratique. Au contraire la science constitue un *champ* social, c'est-à-dire un microcosme social en partie autonome et en partie dépendant du reste de la société [55, p. 14]. Le *champ* scientifique a la particularité de récompenser (entre autres) la recherche de l'objectivation du réel [55, p. 27]. Dans cette perspective, on comprend que les résultats de la transposition ne confèrent pas à l'outil une validité scientifique « pure », mais une légitimité dans le champs scientifique où il sont introduits basée sur la supposition qu'une méthode productive dans un cas particulier a plus de chances de l'être de nouveau dans un autre cas qu'une autre méthode qui n'a rien donné ou qui n'a pas été essayée. Si cette hypothèse paraît raisonnable, on comprend que le problème de la légitimité de la transposition d'outils a bien sûr une composante rationnelle, mais qui ne peut se comprendre et s'expliquer que dans le contexte du champ social dans lequel elle s'inscrit. Aussi la transposition de la méthode sera jugée par les scientifiques du champ disciplinaire concerné, aussi bien théoriquement que d'après les résultats obtenus. C'est donc aux enjeux de cette étape d'évaluation et éventuellement de reconnaissance de la transposition méthodologique que j'essaierai maintenant de réfléchir.

9.4.2 Évaluation

Alors que la transposition interdisciplinaire ne peut être évaluée objectivement *a priori* ou *a posteriori*, le problème initialement posé de la légitimité épistémologique de l'importation d'outils ne peut être séparé de celui de sa légitimation par le champ scientifique concerné. L'enjeu est alors de comprendre comment et pourquoi la transposition va être évaluée par la communauté scientifique et j'ai, d'ailleurs, déjà donné un aperçu de certains aspects individuels du problème (comme l'autocensure). En fait, même si l'évaluation disciplinaire intègre un jugement rationnel des résultats apportés par l'outil, la reconnaissance ou non de la validité de la transposition dépend largement des changements de paradigme introduits et de la manière dont l'import est effectué.

29. ↑ Ces données n'ont pas encore été exploitées.

9.4.2.1 De la validation rationnelle

L'évaluation de la transposition est avant tout sociale, mais il faut nuancer le subjectivisme présenté jusqu'à maintenant : l'organisation sociale de la science favorise par de nombreuses manières la rationalité, aussi l'évaluation au sein du champ disciplinaire n'exclut pas l'évaluation rationnelle. Le ou la scientifique, de par sa formation au raisonnement logique, évalue de manière rationnelle tout d'abord la méthodologie employée et la logique sous-tendant la transposition de méthode, ensuite l'application effective de la méthode à la nouvelle discipline et, enfin, les résultats obtenus par la méthode. En outre l'ensemble de la démarche doit relever d'une cohérence interne et d'une adéquation avec le sujet étudié.

De plus, même si la raison est elle-même subjective, la structure sociale et institutionnelle que constitue le champ scientifique favorise dans une certaine mesure l'objectivité. Tout d'abord, en instituant des normes de rationalité partagées au sein de ce champ, qui fournissent des références communes pour l'évaluation réciproque des travaux scientifiques [55, p. 27] (par exemple la valeur p acceptable pour considérer un test statistique significatif). Ensuite en donnant des indices de légitimité et de valeur aux travaux, notamment via la vérification par plusieurs scientifiques et la stimulation du débat contradictoire. Même si la raison et la logique varient entre les sociétés et au cours du temps, l'organisation sociale de la science permet d'obtenir une « objectivité relative », valable dans une culture donnée et s'affranchissant dans une certaine mesure de la diversité individuelle des raisonnements³⁰.

9.4.2.2 Des transformations induites

Conceptions portées par la méthode. La seule rationalité individuelle et institutionnelle ne suffit cependant pas à expliquer comment est évaluée la transposition méthodologique. L'outil a en effet été construit dans un contexte disciplinaire particulier, au contact de données particulières et porte par conséquent la trace de ce contexte, mais aussi de ses retombées scientifiques. Il n'est donc pas neutre, mais véhicule des concepts propres à la discipline d'origine, au niveau *ontologique* et au niveau *symbolique*. La méthode du *neighbor joining*, par exemple, est une méthode développée en *systématique biologique* consistant à construire une *matrice de distances* puis à regrouper deux à deux les éléments de cette matrice dont la distance est la plus faible jusqu'à avoir regroupé tous les éléments.

Dans son principe même cette méthode porte l'idée de hiérarchie *dichotomique* : 1) elle regroupe successivement les éléments, en organisant ces regroupements sur différents plans (donc de manière hiérarchique) et 2) elle regroupe à chaque fois exactement deux éléments, produisant un arbre dont chaque branche mère se sépare en exactement deux branches filles (dichotomie). Cette idée de hiérarchie dichotomique est intrinsèquement liée à la procédure de *neighbor joining* et lui est donc associée au niveau *ontologique*.

Cette méthode a par ailleurs été introduite dans le but d'effectuer des *inférences phylogéné-*

30. ↑ Gingras [159, pp. 54–56].

tiques en biologie et la plupart de ses utilisations successives ont été faites dans ce cadre. Ceci fait que l'on associe cette méthode à une perspective historique. Cette méthode pourra bien sûr être utilisée dans un autre cadre, mais à moins d'insister sur le fait qu'elle n'implique *pas* cette dimension évolutive, on prend le risque que les autres scientifiques lui associent de fait cette signification. Ainsi l'utilisation de cet outil lui confère une perspective évolutive qui ne vient pas de la méthodologie elle-même, mais de la manière dont elle a été socialisée dans le champ disciplinaire. Cette association avec la notion d'évolution n'est donc pas intrinsèque à la méthode, mais relève du niveau *symbolique*.

Ces concepts, transportés par l'outil, affectent donc plus ou moins profondément l'objet auquel il est appliqué et, plus généralement, les conceptions de la discipline dans laquelle il est introduit. Or ces changements doivent être jugés pertinents par les scientifiques de la discipline, ce qui n'est pas forcément évident en fonction de l'importance de ces transformations induites et des paradigmes régnants dans le domaine d'arrivée.

Compétences requises. Par ailleurs, tout outil nécessite des compétences pour son utilisation, mais aussi pour comprendre son usage par autrui. Or, plus que la simple reconnaissance de la validité de la transposition, sorte de tolérance d'un objet exogène, c'est l'adoption plus large de l'outil qui pourra réellement modifier les conceptions de la discipline où il est introduit. Aussi un outil nécessitant des concepts ou des compétences peu répandues dans la discipline d'arrivée sera difficilement compris et encore moins utilisé, amoindrissant ainsi son impact.

L'utilisation de la méthode d'*inférence bayésienne* présentée dans le chapitre 5 nécessite par exemple des connaissances en probabilités afin de comprendre la méthode, mais aussi des compétences en informatique afin de formater les données et utiliser le logiciel ainsi que des compréhensions des principes sous-jacents à l'*inférence phylogénétique*. Bien que chacun de ces concepts plus ou moins étrangers à l'ethnomusicologie ne soit pas exagérément compliqué, la nécessité de leur acquisition simultanée puis de leur intégration pour l'utilisation de cette méthode rend difficile sa compréhension et son appropriation. En outre si les pairs d'une discipline n'ont pas les moyens de comprendre ces méthodes, les résultats obtenus ne pourront bénéficier que de peu de retours critiques, pourtant essentiels au progrès scientifique collectif et donc à la légitimation de la méthode. On voit donc que les transformations conceptuelles et pratiques induites par l'outil influent sur son évaluation et son éventuelle adoption³¹.

9.4.2.3 De la forme

Enfin les conditions dans lesquelles est effectuée la transposition de l'outil sont déterminantes. Tout d'abord par la manière dont il est présenté, au-delà de l'aspect purement scientifique de

31. ↑ En conséquence, lorsque plusieurs outils d'analyse différents applicables à un même objet requièrent des compétences importantes et surtout dont l'apprentissage est généralement disjoint, cela peut entraîner des séparations disciplinaires en fonction de ces outils. On pourrait voir de tels signes dans le développement d'une communauté scientifique s'intéressant à l'évolution culturelle, dont le bagage scientifique se rattache peut-être davantage à l'analyse quantitative et statistique qu'aux méthodes de l'anthropologie.

l'explication. L'introduction de concepts exogènes peut sembler remettre en cause la capacité de la discipline à répondre elle-même aux questions qu'elle se pose et suggère la nécessité d'avoir recours à l'aide d'un autre domaine. Comme on l'a vu, les raisons de la transposition d'outil sont bien éloignées de l'ingérence³², mais viennent plutôt des apports heuristiques et créatifs de cette forme d'interdisciplinarité. Il convient cependant de ne pas laisser planer cette hypothèse d'ingérence. Aussi l'apport heuristique de la transposition doit être explicité et l'humilité de la démarche me paraît indispensable.

Peut-être plus fréquente dans ces cas de travail interdisciplinaire, la réflexion épistémologique pourrait elle aussi participer à un jugement plus favorable. On peut ainsi penser que les nombreux problèmes méthodologiques et épistémologiques que je me suis posés et que j'ai présentés dans ce chapitre viennent de ce besoin d'explicitier les enjeux de la transposition, afin qu'ils ne soient pas mal interprétés. Par ailleurs, compte tenu de l'importance des changements conceptuels et pratiques nécessaires à l'adoption de l'outil, la pédagogie déployée lors de l'introduction de la nouvelle méthode ne peut être négligée. J'ai pour ma part essayé³³ d'expliquer au mieux différentes méthodes statistiques d'inférence historique³⁴ dans l'idée de favoriser les conditions de telles transpositions.

Le contexte social plus général dans lequel a lieu la transposition a aussi un rôle important. Les relations scientifiques et idéologiques entre les disciplines d'origine et d'arrivée peuvent influencer l'évaluation. On peut, par exemple, penser que l'introduction d'une méthode provenant de la biologie dans les sciences sociales pourra être gênée par l'association forte entre ces disciplines de sciences dites « naturelles » et le réductionnisme traditionnellement associé à la naturalisation. Inversement, l'aura positive de sciences d'ailleurs qualifiées d'« exactes »³⁵ (et en particulier les mathématiques) rend attractif l'importation de concepts depuis ces disciplines [350, p. 213]. De la même manière, la capacité de la personne introduisant l'outil à manipuler les concepts et le langage de la discipline d'arrivée (en quelque sorte son « intégration » dans la discipline), et plus généralement sa réputation au sein de celle-ci, influence l'évaluation de la transposition. Je n'entrerai pas plus dans le détail de l'influence de la forme et du contexte de la transposition sur son évaluation, mais d'autres facteurs comme l'imaginaire social mobilisé, ou encore la médiatisation dont bénéficie la transposition, entrent bien entendu en jeu.

32. ↑ Ou en tout cas devraient l'être.

33. ↑ Même si le résultat est certainement moins accessible que je ne l'aurais voulu, sans doute pour avoir choisi d'expliquer un nombre trop grand de méthodes, plutôt que de mieux en expliquer quelques unes.

34. ↑ Voir [partie 1.3](#) pages 24 à 47.

35. ↑ Sans doute mieux reconnues socialement car plus susceptibles de donner lieu à des applications commercialisables et donc plus financées.

Conclusion générale

Quand j'ai commencé ce doctorat, passionné des mécanismes de l'évolution biologique, j'arrivais convaincu que les concepts et méthodes développées par cette discipline pouvaient être utiles pour étudier la musique. Rencontrant par chance Sylvie qui me proposait d'analyser un jeu de données issu de ses enregistrements de musiques du Gabon, d'une cohérence et d'une finesse rares en ethnomusicologie, dans le cadre privilégié d'un laboratoire où se côtoyaient tous les jours biologistes et ethnomusicologues (et bien d'autres), tout cela s'annonçait être facile. Rien pourtant ne devait l'être et seuls le doute et les remises en questions trouvaient aisément leur chemin dans l'exploration de mon sujet de recherche. Car cette thèse aura été avant tout un chemin d'apprentissage de l'humilité et de la patience.

Découvrant avec vertige les nombreux précédents consistant à utiliser des méthodes issues de la biologie pour étudier la culture, je me suis rendu compte aussi de la riche diversité de méthodes qui avaient été développées par chaque discipline. Parmi celles-ci, l'émergence dans trois disciplines distinctes de méthodes curieusement similaires avaient déjà étonné de nombreuses personnes avant moi. Ainsi la linguistique, la biologie et la critique textuelle, constatant que les similarités observées dans leurs objets d'études (les langues, les êtres vivants et les manuscrits anciens) étaient bien souvent dues à une histoire commune, proposèrent toutes les trois des approches semblables pour essayer de déduire l'histoire de ces objets à partir de la comparaison des structures actuelles. C'est sans doute cette découverte qui m'amena à formaliser plus clairement la question qui guidera tout mon doctorat : peut-on retrouver l'histoire des musiques en comparant différentes musiques pratiquées aujourd'hui ? Et si oui, comment ?

Relier des questionnements convergents. Mais loin d'être identiques, j'ai aussi rapidement constaté la diversité de ces méthodes d'*inférence historique*, ce qui fut l'occasion de remettre en question ma démarche et ma posture de chercheur. Pourquoi utiliser nécessairement des méthodes venues de la biologie ? Sont-elles vraiment plus appropriées que les méthodes venues d'autres disciplines et dont j'ignore tout ? J'ai donc voulu comprendre et décrire un certain nombre de ces méthodes qui cherchaient à retrouver l'histoire en comparant des structures (biologiques, linguistiques ou textuelles) observées aujourd'hui, que j'ai appelé les *méthodes d'inférence historique par comparaison des structures formelles*³⁶. Il m'était bien entendu impossible d'en donner un inven-

36. ↑ Voir [chapitre 1](#) pages 15 à 47.

taire exhaustif, mais ces descriptions m'ont sans aucun doute ouvert l'esprit à d'autres pratiques et d'autres problématiques, souvent bien plus proches des préoccupations de l'ethnomusicologie que ne le sont celles venant de la biologie. J'espère aussi que ces descriptions favoriseront l'échange et la réflexion entre des disciplines historiques dont les objets sont peut-être très divers, mais dont les questionnements se rejoignent et devraient pouvoir s'enrichir mutuellement.

Documenter la transposition de méthodes issues de la biologie. J'ai cependant constaté que cette remise en question, qui me faisait soupçonner dans ma démarche une sorte d'intrusion ingérante et dérangeante de la biologie dans l'ethnomusicologie, semblait à l'exact opposé d'une tendance forte dans l'étude du changement culturel. En effet depuis plusieurs décennies les méthodes venant de la systématique biologique sont de plus en plus employées pour étudier des données culturelles. À la fois incertain de ma légitimité et curieux de ces expériences, je me suis alors attaché à recenser et décrire au mieux ces explorations³⁷. J'ai vu la diversité des tentatives qu'elles ont suscitées, j'ai vu aussi les imprécisions sur lesquelles elles se fondaient bien souvent, j'ai vu encore les échanges qu'elles ont nourris et les nouvelles pistes qu'elles ouvraient. J'ai pensé aussi que la démarche que j'avais entamée valait d'être poursuivie à condition qu'elle reste modeste et de garder à l'esprit qu'elle n'était qu'une exploration dont le principal intérêt résidait sans doute plus dans les critiques qu'elle pourrait susciter que dans d'hypothétiques découvertes. Peut-être le récit de ces expériences, qui m'ont guidées à de nombreux égards, permettra-t-il à d'autres de trouver une inspiration venue de disciplines qui leurs sont étrangères. Peut-être aussi que l'aperçu donné de ces travaux, explorant l'histoire de manuscrits, de statuettes, de mots, ou encore d'instruments de musiques, contribuera à poursuivre la fédération en cours de disciplines scientifiques, non plus concentrées sur leur objet mais rassemblées dans leur quête de compréhension de l'histoire.

Utiliser et adapter les outils ethnomusicologiques. Soutenu par ces précédents sur lesquels je pouvais m'appuyer et quelque peu réconforté sur la légitimité de mes recherches, je manquais néanmoins des outils conceptuels et pratiques nécessaires pour décrire et comparer les musiques. Je me suis donc plongé dans les méthodes développées par l'ethnomusicologie structuraliste, permettant de décrire d'une manière pertinente les musiques³⁸. Je comprenais aussi peu à peu en quoi les méthodes qu'elle proposait étaient essentielles dès lors qu'il est question de comparer des musiques. En outre cette école structuraliste s'est beaucoup construite autour de l'étude des musiques d'Afrique centrale et je découvrais en même temps leurs caractéristiques et leurs spécificités. J'étais aussi intrigué par le concept de *modèle*, central dans cette approche structuraliste, qui me semblait ouvrir des pistes intéressantes du point de vue de la cognition musicale et indirectement de l'approche historique qui m'occupait. Mais je n'arrivais pas à saisir complètement la notion car ses contours semblaient toujours m'échapper. Ce n'est que par la comparaison des différents concepts dérivés de celui de *modèle* utilisés par plusieurs ethnomusi-

37. ↑ Voir chapitre 2 pages 49 à 103.

38. ↑ Voir chapitre 3 pages 105 à 128.

colgues qui j'ai pu commencer à en saisir l'histoire et les enjeux. Tout ceci me donnait finalement les outils pour comprendre les données qui m'étaient confiées et inspirer mes analyses.

Rappeler et s'inspirer des recherches passées. J'ai découvert à l'occasion de cette plongée dans l'ethnomusicologie les relations tumultueuses avec la comparaison³⁹. D'abord érigée comme but ultime de la *musicologie comparée* puis décriée comme prématurée, hasardeuse voire impossible, la comparaison a connu une popularité fluctuante dans les théories ethnomusicologiques. Elle a cependant été d'une présence et d'une importance constante tout au long de son histoire. Mais les fins aussi bien que les moyens de la comparaison musicale ont toujours été multiples. Celle qui m'intéressait, la comparaison à visée historique, n'était pas en reste en traversant cette histoire agitée : encensée aux débuts de la discipline et placée aux côtés de la linguistique et de l'anatomie comparées, on ne peut donner tort à ses critiques qui lui reprochaient de n'avoir pas seulement pris la peine de comprendre et d'étudier les musiques dont on cherchait à dresser le podium. Plus d'un siècle plus tard, relisant ce cours d'humilité qu'est le témoignage de la science passée, je compris que les apports de l'ethnomusicologie structuraliste et nos connaissances plus approfondies sur les musiques d'Afrique centrale rendaient sans doute l'époque plus propice à de telles comparaisons, mais qu'elles ne se feraient pas sans difficulté et qu'il ne me fallait pas être trop impatient d'en voir les apports.

Expérimenter et tâtonner. Mes premières recherches, menées alors que j'explorais ces différentes thématiques et que j'aiguisais encore mes réflexions sur le sujet, pâtirent sans doute de cette immaturité. Enthousiasmé par les travaux de Le Bomin et coll. [240] qui analysaient par une méthode *cladistique* les patrimoines musicaux d'une soixantaine de populations du Gabon, poussé par l'exigence de la publication scientifique au cours du doctorat, je jetais mon dévolu sur une méthode d'*inférence phylogéographique*⁴⁰. Celle-ci devait me permettre de retrouver les localisations ancestrales des patrimoines musicaux que Sylvie Le Bomin et ses collègues avaient étudiés grâce à leur arbre qui proposait (croyais-je) une hypothèse concernant leurs relations historiques et les positions actuelles de ces patrimoines. Je découvrais, au fur et à mesure que les échanges avec mes collègues affinaient mes réflexions, ma mauvaise compréhension des résultats de l'analyse sur laquelle je m'étais basé, ainsi que le flou qui entourait mes pensées concernant les présupposés, les hypothèses ou même la finalité de l'étude.

Décrire et cartographier les musiques du Gabon. Était-ce à ce moment que je découvrais l'injonction de Mantle Hood à ne pas trop anticiper la comparaison ? Peut-être et peut-être aussi est-ce cela qui m'a poussé à penser que cette première exploration phylogéographique était prématurée et trop ambitieuse. Les hypothèses sur lesquelles elle se basait n'étaient pas vérifiées faute d'avoir pris le temps d'analyser et comprendre plus en profondeur les données sur lesquelles je travaillais. En outre décrire ces données pouvait être d'un réel intérêt pour l'ethnomusicologie

39. ↑ Voir [chapitre 4](#) pages 129 à 143.

40. ↑ Voir [chapitre 5](#) pages 155 à 185.

du Gabon qui n'avait encore jamais tenté de dégager des tendances générales à partir des musiques de ce pays.

Sylvie Le Bomin et moi-même nous sommes donc penché-e-s sur l'exploration plus fine des pièces transcrites, en particulier en ce qui concerne les organisations métrique, rythmique et scalaire⁴¹. Ce fut l'occasion de mettre en évidence la diversité des musiques du Gabon et de dégager certaines généralités qui, bien que toujours contredites par telle ou telle pièce, n'en demeuraient pas moins significatives. Ainsi l'organisation **métrique** des pièces analysées révélait-elle une nette organisation géographique, en particulier en ce qui concerne la subdivision de la pulsation, majoritairement binaire à l'ouest et ternaire à l'est du pays. Certaines constatations contrastaient aussi avec ce que l'on connaissait des musiques d'Afrique centrale, et si nous en avons parfois l'intuition, elles n'avaient jamais été formellement établies. Ainsi les **échelles** musicales montraient-elles une étonnante diversité aussi bien à l'échelle du pays qu'au sein de chaque population, et une seule valeur de paramètre musical suffit rarement à caractériser un répertoire au sein d'une population. Cette étude a surtout été l'occasion d'explorer les stimulantes perspectives ouvertes par les travaux de Le Bomin et coll. [240] suggérant que différents types de paramètres musicaux se transmettent plus ou moins verticalement. Creusant cette piste, nous avons pu voir que le partage de **formules rythmiques** s'observe plus souvent entre populations proches, suggérant que les formules rythmiques pourraient être moins sujettes que d'autres paramètres musicaux à la convergence.

Décortiquer et affiner les présupposés. Si cette exploration détaillée de notre jeu de données ouvrait la perspective de tentatives plus éclairantes d'inférence historique, celles-ci ne pouvaient se passer d'un fond théorique mieux stabilisé. J'ai cherché à contribuer à ma mesure à cette réflexion en explorant les liens qu'il est possible d'établir entre des données contemporaines et des histoires passées⁴². Mais cette perspective étant très large, je me suis concentré sur ce que peut nous apprendre la comparaison des musiques : l'histoire des musiques elle-même ou aussi celle des populations ? En défrichant cette question complexe à la lumière de l'expérience d'autres disciplines, il m'a semblé que différents types de données, gènes, langues, musiques ou autres, n'ont pas nécessairement cheminé ensemble le long de l'histoire, et ce même s'il s'agit de différents types de données musicales. Ceci suggère que l'histoire des populations n'est pas directement déductible de la comparaison des musiques, et que dans ce type d'études la question de la congruence des données comparées est primordiale.

Je me suis aussi demandé comment pouvait être choisi le type de données le plus à même de nous renseigner sur l'histoire. Ceci m'a conduit à proposer différents types de critères permettant d'évaluer ce que j'appelle le *potentiel historiographique*. Il m'a semblé que les indices les plus intéressants étaient ceux déduits de notre compréhension plus générale des musiques, car ils permettraient d'utiliser la musique comme un marqueur historique indépendant d'autres marqueurs (par exemple linguistiques ou génétiques), et de tisser avec eux un faisceau de preuves

41. ↑ Voir [chapitre 6](#) pages 187 à 245.

42. ↑ Voir [chapitre 7](#) pages 247 à 274.

plus robuste.

Inspirer les recherches à venir. Ce cheminement théorique révélait ainsi le peu de connaissances dont nous disposons pour justifier d'utiliser la comparaisons des musiques pour inférer l'histoire. Alors que ce doctorat s'achevait, le moment venait de proposer quelques pistes de recherches qui me semblaient mériter d'être explorées⁴³. J'ai donc voulu insister sur la nécessité de mieux comprendre la manière dont les musiques se transmettent et se transforment car ces éclairages viendront guider nos recherches de l'histoire. Et puisqu'en guise de phare pour nous orienter nous n'avons qu'une faible lueur, je suggère aussi d'approfondir l'étude des formules rythmiques dont la similarité semble, mieux peut-être que d'autres paramètres musicaux, refléter l'histoire.

Susciter l'échange. J'ai été accompagné tout au long de ce travail par la question de l'interdisciplinarité, qui m'ouvrait des perspectives toujours passionnantes, mais dont les pièges ont tant marqué ce doctorat que j'ai tenu à partager, comme un exorcisme, quelques unes de mes réflexions concernant mes pratiques de recherche⁴⁴. Car au-delà de la seule connaissance des musiques du Gabon et de leur histoire, la recension parallèle de problématiques communes à différentes disciplines, leur mise en perspective et l'exploration des problèmes qui en surgissent constituent je l'espère un apport utile. Je souhaite en tout cas qu'il contribue à l'approfondissement mutuel de ces savoirs.

43. ↑ Voir [chapitre 8](#) pages 279 à 292.

44. ↑ Voir [chapitre 9](#) pages 295 à 313.



Muséum National d'Histoire Naturelle

École Doctorale Sciences de la Nature et de l'Homme – ED 227

Année 2017

N° attribué par la bibliothèque

THÈSE

Pour obtenir le grade de

DOCTEUR DU MUSÉUM NATIONAL D'HISTOIRE NATURELLE

Spécialité : Évolution culturelle

Présentée et soutenue publiquement par

Jérémy Gardent

le 30 juin 2017

Mesurer les musiques pour parler du passé La comparaison des musiques du Gabon comme source d'informations historiques

Volume III : Glossaire — Annexes

Sous la direction de : Évelyne HEYER et Sylvie LE BOMIN

JURY :

Dana Rappoport,	Directrice de Recherche, CNRS, Paris	Rapporteuse
Stephan Shennan,	<i>Professor</i> , University College, Londres	Rapporteur
Olivier Tourny,	Chargé de Recherche, CNRS, Paris	Examineur
Olivier Morin,	<i>Research group leader</i> , Max Planck Institute, Jena	Examineur
Évelyne Heyer,	Professeure, MNHN, Paris	Directrice de Thèse
Sylvie Le Bomin,	Maître de Conférence, MNHN, Paris	Co-directrice de Thèse



Travail sous licence Creative Commons
Attribution – Partage dans les Mêmes Conditions
4.0 International

Table des matières

Volume I : Introduction	7
Remerciements	7
Préambule	9
Introduction	13
1 L'inférence historique par comparaison des structures formelles	15
1.1 De quoi s'agit-il?	17
1.2 Historique disciplinaire	18
1.3 Description de méthodes d'inférence historique	24
1.4 Conclusion du chapitre	47
2 Les méthodes issues de l'inférence phylogénétique appliquées aux données culturelles	49
2.1 État de l'art disciplinaire	53
2.2 Considérations générales sur ces transpositions interdisciplinaires	75
2.3 Tableau comparatif de quelques études	81
2.4 Conclusion du chapitre	103
3 Systématique musicale	105
3.1 Outils théoriques et méthodologiques	107
3.2 Des outils au service de la comparaison interculturelle	114
3.3 Réflexions autour de la notion de <i>modèle</i>	118
3.4 Conclusion du chapitre	128
4 Musique, comparatisme et histoire	129
4.1 Quelle place pour la comparaison?	130
4.2 La comparaison dans la pratique de l'ethnomusicologie	132
4.3 La comparaison musicale pour l'inférence historique	139
4.4 Conclusion du chapitre	143
Conclusion de la première partie	145
Volume II : Résultats — Discussion	153
Résultats	153
5 Essai d'inférence phylogéographique	155
5.1 Contexte et objectifs	157
5.2 Méthodes	158
5.3 Résultats	165
5.4 Discussion	180
5.5 Perspectives	184

6	Un aperçu de la diversité des musiques du Gabon	187
6.1	Introduction	188
6.2	<i>Diversity and distribution of musics in Gabon</i>	191
6.3	Conclusion et perspectives	245
7	Quelles données, quelle(s) histoire(s) ?	247
7.1	Positionnement épistémologique	249
7.2	Croisements d'histoires	251
7.3	Les indicateurs de choix des marqueurs historiques	268
7.4	Conclusion du chapitre	274
	Conclusion de la deuxième partie	275
	Discussion	277
8	Pistes de recherches	279
8.1	Comprendre la transmission et la transformation des musiques	280
8.2	Approfondir l'étude des formules rythmiques	287
9	Réflexions sur la pratique de la recherche	295
9.1	La relation aux données	297
9.2	Le rôle d'analyste	301
9.3	Limites de l'interdisciplinarité	304
9.4	La transposition de méthodes	306
	Conclusion générale	315
	Volume III : Glossaire — Annexes	327
	Glossaire argumenté	327
	Annexes	370
A	Conventions	373
A.1	Citations et traductions	373
A.2	Sources et références bibliographiques	374
B	Suppléments	377
B.1	Chapitre 3 : Citations autour de la notion de <i>modèle</i>	378
B.2	Chapitre 5 : Inférence phylogéographique	387
B.3	Test de Mantel	421
C	Bibliographie	425
C.1	Publications consultées	425
C.2	Autres publications citées	445
D	Listes des figures et des tableaux	457
D.1	Liste des figures	458
D.2	Liste des tableaux	460
E	Table exhaustive des matières	461

Volume III
Glossaire — Annexes

Glossaire argumenté

Le sujet de cette thèse recoupe des disciplines parfois très différentes ayant chacune leur vocabulaire spécifique, notamment en ce qui concerne les aspects méthodologiques. Aussi une même explication peut s'avérer triviale et fastidieuse ou au contraire indispensable selon que l'on soit familier de la discipline concernée ou non. Le glossaire détaillé présenté ici vise à expliquer les termes et notions qui sont à la fois trop généraux pour pouvoir être systématiquement exposés dans le corps du texte et trop spécifiques pour pouvoir être compris par une personne étrangère à la discipline. Ce glossaire argumenté me permet aussi de discuter des différentes acceptations de termes dont l'usage est parfois ambigu et par là de clarifier le sens précis dans lequel je les utilise ici.

ADNmt [p. 253, 255, 259, 263, 333]

ADN mitochondrial (ou *mtDNA* en anglais). Il s'agit de l'ADN contenu dans les mitochondries, organites présents dans les cellules eucaryotes (y compris celles des humains) qui ne se transmettent que de mère en fille. L'ADN mitochondrial humain est utilisé en génétique des populations, tout comme le **chromosome Y** pour inférer l'histoire des migrations humaines.

analogie [p. 35, 342, 343]

Dans le vocabulaire de la **cladistique**, une analogie est une ressemblance (structurelle ou fonctionnelle) entre deux caractéristiques observables, qui sont dites *analogues*. Par exemple l'aile des insectes, des oiseaux et des chauves-souris remplissent la même fonction et peuvent à ce titre être qualifiées d'analogues. De même, les pattes des oiseaux et celles de mammifères sont analogues. Dans une perspective **phylogénétique**, on distingue deux types d'analogie : l'**homologie** est imputable à une ascendance commune alors que l'**homoplasie** ne l'est pas. La constatation d'une analogie est la première étape conduisant à formuler une hypothèse d'homologie.

apomorphie [p. 23, 334, 339, 357, 362]

En **cladistique**, une apomorphie est l'**état dérivé** d'un **caractère**, c'est-à-dire que chez l'ancêtre (théorique) de tous les groupes comparés, le caractère en question n'est pas dans

cet état. Cet état de caractère a donc évolué plus tardivement au cours de l'évolution des taxons considérés. L'apomorphie s'oppose à la **plésiomorphie**, qui est l'état ancestral d'un caractère.

Ainsi, sur la figure 1.4 page 37, la présence de fleurs (état « présence » pour le caractère « Fleurs ») est une apomorphie car la présence est l'état dérivé (définie grâce au groupe externe, le pin blanc de Provence) du caractère C1 « Fleurs ».

Selon la méthode cladistique, seul le partage d'une apomorphie (appelée *synapomorphie*) par plusieurs groupes (mais pas tous) est utile et pertinent pour inférer les relations de filiation entre espèces, alors que le partage d'une plésiomorphie (une *symplesiomorphie*) n'est pas informatif. Il y a différentes manières d'identifier l'état dérivé d'un caractère, l'une d'entre elles étant l'utilisation de **groupes externes**.

arborescent [p. 18, 62, 66, 80, 87, 103, 158, 180, 184, 260, 263, 328, 336, 360, 361]

Qui a la forme d'un arbre. On trouve aussi les termes de *dendroïde*, *dendriforme* ou encore (mais plus rarement) d'*arboriforme*. J'utilise pour ma part uniquement ce terme d'*arborescent*.

arboriforme [p. 328]

Qui a la forme d'un arbre. J'utilise pour ma part son synonyme *arborescent*.

arbre [p. 67, 68, 71, 72, 79–81, 86, 87, 275, 286, 329, 336, 337, 348, 351, 352, 357, 360, 361, 367, 369]

Dans la théorie mathématique des graphes, un arbre est un ensemble de nœuds tous reliés par des branches (ou arêtes) de manière non cyclique. Il existe donc un et un seul chemin de branches permettant de relier deux nœuds donnés. On dit que l'arbre est *raciné*⁴⁵ si on a choisi une branche de l'arbre où placer une *racine* déterminant la base de l'arbre et orientant sa lecture. Les nœuds reliés à un seul autre nœud sont appelés *nœuds internes* ou *feuilles*. Les nœuds reliés à plus d'un autre nœud sont appelés *nœuds externes*. Un arbre est dit *dichotomique* si chaque branche ne se sépare jamais en plus de deux autres branches, c'est-à-dire que chaque nœud n'est jamais relié à plus de trois autres nœuds. Tous les nœuds internes sont donc des *bifurcations*. Un nœud interne relié à plus de trois autres nœuds est appelé une *multifurcation*.

Ce concept est souvent utilisé en *systématique biologique*⁴⁶ pour représenter des relations de parentés entre **taxons** : on parle alors d'*arbre phylogénétique*. Un arbre construit par une méthode **cladistique** est appelé un *cladogramme*. Lorsque des réticulations sont permises dans la structure de l'arbre⁴⁷ on parle de *réseau*.

45. ↑ On dit *orienté* dans la théorie des graphes.

46. ↑ Les conceptions portées par la métaphore de l'*arbre* sont très diverses en biologie, voir à ce sujet le travail de Fislér [A140].

47. ↑ C'est-à-dire que le graphe peut-être cyclique.

arbre de consensus [p. 64, 180, 184]

En **systematique biologique**, un arbre de consensus est un **arbre** représentant le consensus d'un ensemble d'arbres différents. Il est utilisé lorsqu'une méthode d'**inférence phylogénétique** produit plusieurs arbres pouvant représenter la même **phylogénie**, sans que l'on puisse déterminer lequel d'entre eux devrait être choisi. On pourra trouver dans le livre de Darlu et Tassy [107, pp. 120–123] des explications plus détaillées concernant les méthodes d'obtention des différents types d'arbre de consensus (strict et majoritaire).

arbre phylogénétique [p. 32, 34, 39, 76, 79, 328, 334, 343–345, 356, 369]

Pour Lecointre et Le Guyader [251, p. 202], un **arbre** peut être considéré comme phylogénétique si « [l]e concept de descendance avec modification a été formulé dans la méthode qui a présidé à sa construction » et « s'il permet une découverte *a posteriori* des homologies et des homoplasies ».

Pour ma part, j'utilise l'adjectif **phylogénétique** au sens large comme se référant à une phylogénie. Ainsi un arbre phylogénétique est un arbre proposant une inférence concernant une phylogénie. C'est donc une **historiographie** de cette phylogénie présentée sous la forme d'un arbre. Selon cet entendement, ce qui préside est le but dans lequel il est réalisé, plutôt que la méthode ayant permis de l'obtenir (et qui peut éventuellement être jugée mauvaise pour parvenir à ce but).

arbritude [p. 81]

Degré selon lequel un objet donné (par exemple un jeu de données ou un **réseau**) peut être représenté sous la forme d'un **arbre**. Il s'agit d'un néologisme que je propose comme traduction du terme anglais *treeness*. J'aurais pu aussi proposer le terme de *dendritude*, mais je privilégie le terme issu de la racine « arbre » car sa signification est plus transparente pour les francophones.

artéfact [p. 66, 70, 86]

En archéologie, un artéfact est un objet fabriqué par l'espèce humaine, par exemple une pointe de flèche ou une poterie.

bayésien [p. 64, 77, 158]

Voir **inférence bayésienne**.

binaire (caractère) [p. 32, 330, 339, 360]

Dans l'analyse **cladistique**, un caractère est dit binaire lorsqu'il ne peut prendre que deux **états**. Ces deux états sont souvent l'absence et la présence de ce caractère, et sont alors généralement codés respectivement 0 et 1. Par exemple, le caractère « Utilisation de la technique vocale du **jodel** » (défini par exemple pour une pièce musicale) a deux états : « oui » et « non ». Certains caractères peuvent n'avoir que deux états observés qui ne sont pas l'absence ou la présence, mais sont néanmoins dits binaires. C'est par exemple le cas

du caractère « Type de **subdivision de la pulsation** », tel que nous l'avons défini ici pour une pièce, qui ne prend dans notre jeu de données que les valeurs « binaire » et « ternaire ».

binaire (subdivision de la pulsation) [p. 291, 293, 302, 350]

Voir *modalité de subdivision de la pulsation*.

biologie de l'évolution [p. 18, 26, 250, 340, 346, 349]

La biologie de l'évolution est le domaine de la biologie qui s'intéresse à comprendre les mécanismes de l'évolution du vivant, ainsi qu'à **inférer l'histoire** de cette évolution. On trouve aussi parfois le terme de *biologie évolutive*.

bootstrap [p. 60, 66, 67, 70, 180, 367]

Dans les **méthodes issues de l'inférence phylogénétique**, l'analyse de *bootstrap* est une méthode permettant de tester la **robustesse** de l'inférence. Pour cela, une nouvelle **matrice de caractères** est constituée en tirant au hasard des caractères dans la matrice ayant servi à l'analyse. On tire le même nombre de caractères qu'il y avait au départ, mais en autorisant les remises, c'est-à-dire que certains caractères seront présents plusieurs fois et d'autres aucune. On utilise ensuite cette matrice pour effectuer la même analyse que celle dont on cherche à évaluer la robustesse. On répète cette procédure (tirage avec remise des caractères puis inférence) un grand nombre de fois et on compare les arbres obtenus. Pour chaque nœud obtenu dans l'arbre dont on teste la robustesse, on compte le pourcentage d'arbres obtenus dans l'analyse de bootstrap qui comporte ce nœud. Cette *valeur de bootstrap* est un indicateur de la robustesse du nœud : plus ce pourcentage est élevé, plus le nœud est considéré comme robuste.

caractère [p. 32, 34, 35, 54, 62, 85, 87, 158, 283, 287, 304, 327, 330, 331, 335, 336, 339, 342, 343, 346, 352–354, 357, 360, 362, 367, 369]

Dans le vocabulaire de la **cladistique**, « [o]n appelle caractère⁴⁸ tout attribut observable d'un organisme » [107, p. 23]. Dans une perspective comparatiste, on parlera surtout de caractères pour désigner des caractéristiques qui sont observables dans l'ensemble des **taxons** comparés. Un caractère peut être observé sous plusieurs formes, que l'on appelle **états**. Ainsi le caractère « Type de sol » sur lequel pousse une plante a comme **état** « basique » chez le pin blanc de Provence et « acide » chez le crocus d'automne⁴⁹.

On parle de *caractère binaire* lorsqu'un caractère ne peut prendre que deux états différents et de *caractères à états multiples* lorsqu'il peut prendre plus de deux états.

La notion de caractère est équivalente à celle utilisée en **ethnomusicologie** de *paramètre*⁵⁰. Je choisis donc d'utiliser le terme de *paramètre* lorsqu'il s'agit de données de nature musicale et celui de *caractère* lorsque j'évoque des données de nature biologique.

48. ↑ On trouve aussi parfois le terme *trait*, plutôt utilisé pour des caractères phénotypiques.

49. ↑ Voir figure 1.4 page 37.

50. ↑ Elles se ressemblent mais ne sont pas strictement équivalentes, comme je le discute dans la partie 9.3.1 page 304.

caractère continu [p. 70, 158, 331]

Une variable continue est une variable pouvant prendre un nombre infini de valeurs de manière continue. Un caractère continu est donc un **caractère** dont la **valeur** est une variable continue. Par exemple, la latitude d'une localité peut prendre toutes les valeurs comprises entre 8,7 et 14,5 : c'est un caractère continu⁵¹. Il s'oppose en cela à un *caractère discret* qui ne peut prendre qu'un nombre fini de valeurs.

caractère discret [p. 331]

Une variable discrète est une variable qui ne peut prendre qu'un nombre fini de valeurs. Un caractère discret est donc un **caractère** dont la **valeur** est une variable discrète. Par exemple le nombre de degrés d'une échelle musicale peut être de 5, 6 ou 7, mais pas de 5,8 : c'est donc un caractère discret. Il s'oppose en cela à un *caractère continu* qui peut prendre un nombre infini de valeurs.

catégorie [p. 114, 333, 360, 365]

En **systématique musicale**, une catégorie est un ensemble de **pièces** musicales possédant une propriété qui leur est commune :

« Par catégorie musicale, on entend ici un ensemble vernaculaire doté d'un trait musical – c'est-à-dire concernant uniquement la substance sonore : échelle, forme, périodicité, dispositif vocal et/ou instrumental, etc. – qui permet de distinguer cet ensemble de tous les autres, donc de toute autre catégorie. Ce trait se manifeste dans toutes les pièces constitutives de cette catégorie. »

Arom et Martin (2015) [24, p. 235]

catégorisation [p. 332]

Activité consistant à **catégoriser**, c'est-à-dire à construire un **système catégoriel**.

catégorisation musicale [p. 107, 111, 117, 128, 305, 350, 364, 365]

Le terme de catégorisation musicale peut prendre deux sens différents, relevés par Susanne Fürniss [154, p. 27] :

« Le travail de catégorisation musicale consiste à rendre compte, d'une part, de la façon dont différentes cultures conçoivent et ordonnent leur monde musical. C'est la mise au jour des critères musicaux qui fondent l'organisation vernaculaire d'un patrimoine. D'autre part, il s'attache à la détermination du degré de correspondance entre les critères culturels en action et ceux issus de l'analyse musicologique. »

Fürniss-Yacoubi (2007) [154, p. 27]

Dans un premier sens, la catégorisation musicale s'intéresse à la manière dont sont organisés les éléments d'ordre musical d'une culture, du point de vue endogène. Le terme peut

51. ↑ Ces valeurs sont les longitudes extrêmes du Gabon, respectivement à l'ouest et à l'est.

désigner, selon les contextes, soit le système d'organisation endogène lui-même⁵², soit (le plus souvent) l'activité scientifique s'attachant à mettre au jour cette organisation, comme dans la citation ci-dessus.

Dans un deuxième sens, la catégorisation musicale a trait au travail d'organisation de manière logique et cohérente des éléments d'une culture musicale en un ensemble de catégories par les scientifiques. Elle peut alors désigner aussi bien l'activité scientifique d'organisation que le résultat produit :

« La catégorisation peut être envisagée selon deux acceptions, suivant que l'on considère l'acte de catégoriser en lui-même ou bien le résultat qui en découle. Dans le premier cas, on s'intéresse à l'opération qui consiste à classer les éléments relevant d'un même domaine, qu'il soit technique, linguistique, artistique ou de toute autre nature; l'accent est alors mis sur les processus qui aboutissent à la classification. Dans le second, c'est le résultat qui tient une place centrale, résultat à partir duquel on peut opérer des comparaisons, que ce soit au sein d'un même domaine comme entre différents domaines d'une même culture ou, pour un domaine donné, entre différentes cultures. »

Arom et coll. (2008) [21, p. 279]

Ce travail de catégorisation scientifique ne saurait cependant s'affranchir de la catégorisation endogène. Comme le relève Fürniss-Yacoubi [154, p. 27] dans la citation ci-dessus, un des objectifs est justement de confronter catégorisations endogène et scientifique.

Du point de vue méthodologique, la catégorisation musicale s'appuie sur l'analyse de la *systematique musicale* de la culture considérée. Elle est par ailleurs fondamentalement comparative au niveau intraculturel et permet la comparaison au niveau interculturel⁵³ :

« La catégorisation musicale repose sur une méthode comparative intraculturelle, c'est-à-dire qu'elle met en paradigme les répertoires qui composent le patrimoine musical à un moment donné, *hic et nunc*. Cette méthode permet de décrire chaque répertoire et de déterminer à la fois ses spécificités formelles et ce que, d'un point de vue musical, il partage avec la majorité des autres répertoires. »

Fürniss (2014) [151, p. 12]

Je choisis ici de séparer les différents sens contenus dans le terme de *catégorisation musicale*. J'entends le terme de *catégorisation* comme l'activité consistant à *catégoriser*, c'est-à-dire à créer un ensemble de catégories. Cet ensemble résultant de la catégorisation est appelé un *systeme catégoriel*. Étant donné que je n'aborde pas la question de savoir comment

52. ↑ Par exemple dans le titre de la deuxième partie de la thèse d'Hugo Ferran : « Systematique et catégorisation de l'univers musical maale » [138, p. 161].

53. ↑ Voir la citation de Arom et coll. [21, p. 279] ci-dessus, ainsi que la partie 3.2 pages 114 à 118.

une crée un système catégoriel⁵⁴, le terme de *catégorisation musicale* sous-entend ici une activité *scientifique* de catégorisation, c'est-à-dire la création d'un système catégoriel à visée scientifique. Comme le système catégoriel lui-même, en tant qu'objet cognitif, est trop difficile à décrire, il s'agit le plus souvent de modéliser ce système :

« Notre démarche vise à *rendre compte de la manière dont les tenants d'une culture classent leur patrimoine musical*. [...] [N]otre méthode [...] permet de proposer une *modélisation* de la catégorisation vernaculaire. »

Arom et coll. (2008) [21, p. 279]

Activité scientifique	systématisation	catégorisation
Production (résultat de l'activité)	systématique	système catégoriel
- scientifique	systématique (scientifique)	système catégoriel (scientifique)
- endogène	systématique endogène	système catégoriel endogène

TAB. 9.1 : **Terminologies concernant la systématisation et la catégorisation musicale.** Ce tableau explicite les choix terminologiques utilisés dans cette thèse concernant la systématisation et la catégorisation musicale.

J'appelle donc ce travail la *modélisation du système catégoriel vernaculaire*. Cela passe bien entendu le plus souvent par une activité de catégorisation (scientifique) puisque le modèle du système catégoriel vernaculaire est lui-même un système catégoriel (mais scientifique, puisque produit à des fins et par des méthodes scientifiques).

catégoriser [p. 331, 332]

Construire un *système catégoriel*, c'est-à-dire un ensemble cohérent de *catégories*.

cellule rythmique [p. 288, 289, 291, 293, 302, 333, 353, 362]

En *ethnomusicologie*, une cellule rythmique est une unité rythmique élémentaire, distinctive et culturellement pertinente. Ainsi le plus souvent une *formule rythmique* peut être divisée en plusieurs *cellules rythmiques*. Les cellules rythmiques identifiées par Sylvie Le Bomin au cours de ses analyses des musiques du Gabon sont représentées dans le tableau du chapitre 6, page 226.

chaîne d'états [p. 162, 166, 338]

Dans le cadre de l'*inférence bayésienne*, la chaîne d'états est l'ensemble des *états* successifs parcourus au cours de l'analyse. Cette notion est expliquée et illustrée plus en détail dans la partie 5.2.3.1 page 162 et en particulier dans la figure 5.2 page 163.

chromosome Y [p. 253, 327]

Le chromosome Y est, chez l'être humain, un chromosome qui ne se transmet que de père en fils. Pour cette raison il est utilisé en génétique des populations, tout comme l'*ADNmt*, pour inférer l'histoire des migrations humaines.

54. ↑ L'ethnomusicologie s'attache le plus souvent à décrire le système catégoriel, plutôt qu'à identifier les processus qui l'ont créé, généralement jugés trop complexes ou inaccessibles.

CI [p. 181, 182, 185]

Voir *indice de cohérence*.

clade [p. 73, 160, 181, 343, 391]

Dans le vocabulaire de l'analyse **cladistique**, un clade est un groupe constitué d'un nœud du **cladogramme** et de tous les nœuds qui en descendent (et eux seuls). Par exemple dans l'arbre 2 de la figure 1.4 page 37, le brome, le crocus et leur plus proche ancêtre forment un clade⁵⁵.

cladisme [p. 334]

Le terme de cladisme désigne un courant de la **systématique biologique** dérivé des travaux de Willi Hennig [A189]⁵⁶. Le terme, qui n'a pas été inventé par Hennig mais par les auteur-e-s anglo-américain-e-s ayant repris ses travaux par la suite [107, p. 31], est issu du mot *clade* (du grec $\chi\lambda\alpha\delta\omicron\sigma$ signifiant *branche*) [251, p. 539].

cladistique [p. 21, 31, 34, 54, 55, 59, 61, 62, 64, 66–70, 74, 75, 84, 87, 158, 180, 260, 263, 283, 304, 317, 327–330, 334, 335, 339, 342–344, 347, 348, 352, 354, 357, 362, 367]

L'adjectif *cladistique* se réfère au **cladisme**. Par nominalisation⁵⁷, la cladistique est aussi un synonyme du cladisme [251, p. 539], mais je n'utiliserai pour ma part le terme qu'en tant qu'adjectif. Le terme de *méthode cladistique* est utilisé pour désigner un ensemble de méthodes d'**inférence phylogénétique** issues des travaux de Willi Hennig [A189]⁵⁸. Pour Lecointre et Le Guyader [251, p. 539] il s'agit d'une « [m]éthode d'analyse des caractères qui vise à mettre en évidence la séquence évolutive de leurs transformations, c'est-à-dire déterminer leur état **plésiomorphe** (primitif) et leur(s) état(s) **apomorphe(s)** (dérivés) ». Il n'existe bien entendu pas « une » méthode cladistique. Je désigne pour ma part, lorsque j'utilise le terme de *méthode cladistique*, celle décrite par Darlu et Tassy [107] et présentée succinctement dans la partie 1.3.6 pages 34 à 36.

cladogramme [p. 40, 68, 328, 334]

En **systématique biologique**, un cladogramme est un *arbre* obtenu par une méthode **cladistique**.

clique [p. 32]

La méthode d'**inférence phylogénétique** par **compatibilité** utilise le terme de clique par analogie avec l'utilisation qui en est faite en théorie des graphes. Dans cette discipline, une clique d'un graphe est un ensemble de sommets qui sont tous reliés les uns aux autres. En biologie, une clique est un ensemble de caractères pour lesquels on peut construire un **arbre phylogénétique** ne nécessitant aucune hypothèse d'**homoplasie** [107, p. 145]. Ces caractères

55. ↑ Ce pourrait être par exemple le clade des espèces monocotylédones.

56. ↑ Citée par Lecointre et Le Guyader [251, p. 539].

57. ↑ La nominalisation consiste à utiliser comme nom un mot d'une autre classe grammaticale, en l'occurrence un adjectif.

58. ↑ Citée par Lecointre et Le Guyader [251, p. 539].

sont dits **mutuellement compatibles** et sont donc en quelque sorte tous « reliés » les uns aux autres par cette compatibilité.

cognat [p. 27, 86, 346]

En **linguistique historique**, des cognats (appelés aussi *mots apparentés*) sont des mots qui ont une origine évolutive commune. Il s'agit en général de mots de différentes langues, que l'identification de ces cognats permet de rapprocher.

« A word (or morpheme) that is related to a word (morpheme) in sister languages by reason of these words (morphemes) having been inherited by the related languages from a common word (morpheme) of the **proto-language** from which they descend. »⁵⁹ Campbell (1999) [75, p.33]

co-incidence [p. 57, 61, 336]

En **critique textuelle**, la co-incidence désigne l'apparition de la même **variante** dans deux manuscrits de manière indépendante. On trouve aussi les termes de *parallélisme*, de *variation coïncidente* ou encore de *variation accidentelle* [366, p. 67]. On peut relier ce concept à celui de **convergence** utilisé en **systematique biologique**.

congruence [p. 286, 304, 305]

De manière générale, le terme de *congruence* désigne le fait que plusieurs éléments soient cohérents, en accord les uns avec les autres. En **cladistique**, il désigne plus précisément le fait que les arbres inférés à partir de l'analyse de différents **caractères** sont cohérents entre eux et présentent la même **topologie** [251, p. 539]. De manière plus générale, j'utilise la notion de congruence pour désigner le fait que différentes données (voire différents types de données) apportent les mêmes informations historiques et qu'il est donc adéquat de les utiliser conjointement pour inférer l'histoire.

consistency index [p. 343]

Voir *indice de cohérence*.

contamination [p. 57, 337, 367]

En **critique textuelle**, on parle de contamination lorsque les **variantes** trouvées dans un manuscrit sont issues d'au moins deux manuscrits différents. On trouve aussi les termes de *combinaison*⁶⁰, de *bâtardise textuelle*⁶¹ ou encore de *changement relationnel*⁶² [366, p. 67]. On peut relier ce concept à celui de *transfert horizontal* utilisé en **systematique biologique**, ou encore à celui d'*emprunt* utilisé en **linguistique historique**.

59. ↑ « Un mot (ou un morphème) qui est semblable à un mot (morphème) d'une langue sœur parce que ces mots (morphèmes) ont été hérités par les langues parentes à partir d'un même mot (morphème) de la **proto-langue** de laquelle elles descendent. »

60. ↑ *Conflation*, en anglais.

61. ↑ *Text bastardy*, en anglais.

62. ↑ *Change in relationships*, en anglais.

continuous recording of individual parts [p. 336]

Traduction en anglais du nom de la méthode d'enregistrement continu des parties séparées (abrégée EnConPaS, ou *CRIP* en anglais).

convergence [p. 37, 57, 80, 287, 335, 343]

En *systématique biologique*, la convergence est l'apparition d'un même état d'un caractère dans deux taxons différents [251, p. 539]. Par exemple dans l'arbre 1 de la figure 1.4 page 37, le bois secondaire disparaît⁶³ deux fois de manière indépendante chez le brome et chez le crocus. On peut rapprocher cette notion de celle de *co-incidence* utilisée en critique textuelle. La convergence, tout comme le *transfert horizontal*, est une forme d'*homoplasie*.

CRIP [p. 336]

Sigle pour *continuous recording of individual parts*. Voir l'entrée *enregistrement continu des parties séparées* pour la description de cette méthode.

critique textuelle [p. 18, 29, 47, 53, 70, 103, 308, 335–338, 346, 362, 367–369]

Suivant la proposition de Salemans [366, p. 11], la critique textuelle est définie ici comme la discipline cherchant à retrouver le contenu d'un texte originel perdu, à partir de copies plus récentes de ce texte.

Plusieurs méthodes ont été proposées pour parvenir à cette fin, parmi lesquelles la *méthode de l'erreur commune*⁶⁴.

J'utilise le terme de *critique des textes* pour désigner les personnes pratiquant cette discipline.

cycle [p. 336, 354]

Voir *période*.

cyclique [p. 336, 341, 350, 354, 355, 362]

En *ethnomusicologie*, une *pièce* musicale est dite *cyclique* (ou *périodique*) lorsqu'un même matériau mélodico-rythmique est reproduit à intervalles réguliers [16, p. 407]. Chacune de ces répétitions est appelée une *période* (ou *cycle*).

dendriforme [p. 328, 336]

Synonyme d'*arborescent*.

dendroïde [p. 328]

Qui a la forme d'un *arbre*. On trouve aussi son synonyme *dendriforme*, moins cohérent puisqu'il mêle le radical grec δένδρον (arbre) et le suffixe latin *-forma* (forme). L'utilisation du suffixe issu du grec εἶδος (forme, aspect) est donc plus logique. Je privilégie pour ma part le terme d'*arborescent*, dont la signification est plus évidente pour les francophones et dont l'emploi est plus fréquent.

63. ↑ Passage du caractère « Bois secondaire » de l'état « 0 » (présence) à l'état « 1 » (absence).

64. ↑ *Common error method*, en anglais.

diachronie [p. 133, 266, 267, 282, 363]

La diachronie s'intéresse à la comparaison au cours du temps : on compare donc deux objets considérés équivalents ou liés, à deux moments différents. Ce terme, issu de la linguistique et proposé par Ferdinand de Saussure, s'oppose à celui de *synchronie*.

dichotomique [p. 56, 72, 311, 328, 351]

Un *arbre* est dit dichotomique si, à chaque ramification, la branche mère est séparée en exactement deux branches filles. Autrement dit, chaque *nœud interne* est relié à exactement trois autres nœuds. Un tel arbre ne possède donc que des bifurcations et aucune *multifurcation*. En *systematique biologique*, un arbre dichotomique est aussi dit *parfaitement résolu* car les multifurcations représentent en général une indétermination sur les relations de parentés.

échelle musicale [p. 73, 138, 181, 271, 318, 353, 357, 362, 365, 460]

Dans le domaine de la *systematique musicale*, Fernando [137] définit ainsi l'échelle musicale :

« L'échelle correspond à une segmentation du continuum sonore en unités discrètes qui s'opposent les unes aux autres. Elle s'inscrit à l'intérieur d'un intervalle cadre – généralement l'octave – au-delà duquel elle se reproduit identique à elle-même. L'échelle contient un potentiel intervallique et renvoie à un inventaire fini, à un réservoir clos de tous les degrés – délimitant la nature et le nombre d'intervalles – à partir desquels sont élaborés les différents modes. »

Fernando (2007) [137, p. 946]

emprunt [p. 63, 85, 284, 335, 367]

En *linguistique historique* un emprunt est l'intégration dans une langue d'un élément linguistique provenant d'une autre langue [77, pp. 25–26]. On peut rapprocher cette notion de celle de *transfert horizontal* utilisée en *systematique biologique* et de *contamination* utilisée en *critique textuelle*.

EnConPaS [p. 298, 336, 338]

Abréviation pour la technique d'*enregistrement continu des parties séparées*.

enregistrement analytique [p. 113, 115, 288, 292, 338, 358]

En *ethnomusicologie*, un enregistrement analytique est un enregistrement hors contexte effectué à des fins d'analyse musicale.

La méthode d'*enregistrement en playback* décrite par Arom [13, pp. 190–203] est par exemple une méthode d'enregistrement analytique essentielle pour comprendre les polyphonies et polyrythmies d'Afrique centrale. Il en existe d'autres, comme la méthode d'*enregistrement continu des parties séparées* imaginée par Sylvie Le Bomin et Susanne Fürniss⁶⁵.

65. ↑ Lechaux [249, p. 82].

enregistrement continu des parties séparées [p. 113, 298, 336, 337, 359]

L'enregistrement continu des parties séparées (ou **EnConPaS**)⁶⁶ est une technique d'**enregistrement analytique** imaginée par Sylvie Le Bomin et Susanne Fürniss⁶⁷. Cette méthode, dérivée de la méthode de *re-recording* de Simha Arom [13, pp. 190–213], permet d'enregistrer chacune des parties d'une **pièce** musicale séparément en un seul enregistrement audio. Elle facilite donc la transcription et l'analyse musicale de la pièce tout en permettant une grande efficacité des enregistrements sur le terrain. Cette méthode est décrite plus en détail dans la partie *Supporting information* du chapitre 6 et illustrée dans la figure page 223.

Susanne Fürniss et Sylvie Le Bomin se réfèrent à cette méthode par le terme d'*enregistrement analytique*⁶⁸, ne permettant pas de la distinguer lexicalement de la méthode de *re-recording* qui est aussi une méthode d'**enregistrement analytique**. J'ai donc choisi d'utiliser ce terme d'*enregistrement continu des parties séparées*, espérant ainsi clarifier la description des méthodes d'enregistrement. Le terme d'*enregistrement continu* se réfère au fait que les enregistrements analytiques de chacune des **parties** sont réalisés les uns à la suite des autres, sans interruption de l'enregistrement.

enregistrement conventionnel [p. 112, 300]

En **ethnomusicologie**, un enregistrement conventionnel est un enregistrement qui n'a pas spécialement été réalisé suivant une méthodologie facilitant l'analyse musicale ultérieure. Il s'oppose donc à l'*enregistrement analytique*.

enregistrement en *playback* [p. 337, 358]

Voir *re-recording*.

erreur [p. 369]

En **critique textuelle**, une erreur est une modification introduite dans un texte lors de sa copie. On utilise aussi le terme de *variante*, dont la connotation est moins négative et dénuée d'*a priori*.

état (inférence bayésienne) [p. 162, 333, 345]

Dans le cadre de l'**inférence bayésienne**, le terme d'*état* désigne un ensemble des valeurs des paramètres que l'on cherche à optimiser. Autrement dit, l'état est l'ensemble de valeurs prises par les paramètres inférés à une **itération** donnée. L'ensemble des états successifs parcourus au cours de l'analyse constitue la **chaîne d'états**. Cette notion est expliquée et illustrée plus en détail dans la partie 5.2.3.1 page 162.

La signification du terme *état* dans le cadre de l'inférence bayésienne ne doit pas être confondue avec celle d'*état de caractère* utilisée en **systématique biologique**.

66. ↑ *Continuous recording of individual parts (CRIP)*, en anglais.

67. ↑ Lechaux [249, p. 82].

68. ↑ Lechaux [249, p. 82].

état (de caractère) [p. 32, 34, 35, 54, 87, 158, 287, 304, 329, 330, 336, 338, 339, 343, 346, 354, 367, 369]

Dans le vocabulaire de la *cladistique*, un état est l'une des formes que peut prendre un *caractère*. Ainsi le caractère « Type de sol » sur lequel pousse une plante a comme *état* « basique » chez le pin blanc de Provence et « acide » chez le crocus d'automne⁶⁹.

Cette notion peut être rapprochée de celle de *valeur de paramètre* en *ethnomusicologie*.

état ancestral [p. 35, 287, 342, 357, 362]

En *cladistique*, l'état ancestral est l'*état* de *caractère* qui était présent chez l'ancêtre commun à tous les *taxons* étudiés. Cette notion est expliquée dans l'entrée de son synonyme *plésiomorphie*.

état dérivé [p. 35, 287, 327, 342, 362]

En *cladistique*, un état dérivé est un *état* de *caractère* qui n'était pas présent chez l'ancêtre (théorique) de tous les groupes comparés. Cette notion est expliquée dans l'entrée de son synonyme *apomorphie*.

états multiples (caractère) [p. 330]

En *systématique biologique*, un *caractère* est dit à *états multiples* lorsqu'il peut prendre plus de deux *états*. Cette notion s'oppose à celle de *caractère binaire*. Dans l'exemple de la figure 1.4 page 37, le caractère C3 « Nombre de cotylédons dans l'embryon » est un caractère à états multiples car il peut prendre trois états : « plus de deux cotylédons », « deux cotylédons » ou « un seul cotylédon ».

ethnogenèse et phylogenèse [p. 66, 355]

En archéologie, les concepts de *phylogenèse* et d'*ethnogenèse* visent à expliquer l'apparition des nouveaux assemblages culturels observés. Ils sont respectivement basés sur les processus de *ramification*⁷⁰ et de *mélange*⁷¹.

« In phylogenesis a new cultural assemblage is the result of descent with modification from an ancestral assemblage, whereas in ethnogenesis a new cultural assemblage arises through the blending of elements of two or more contemporaneous assemblages. »⁷² Collard et Shennan (2000) [93, p. 89]

ethnomusicologie [p. 130, 297, 305, 330, 333, 336–339, 341, 350–354, 358, 363, 364, 366, 369, 370]

69. ↑ Voir figure 1.4 page 37.

70. ↑ *branching*, en anglais.

71. ↑ *blending*, en anglais.

72. ↑ « Lorsqu'il y a phylogenèse un nouvel assemblage culturel est issu de la descendance avec modification à partir d'un assemblage ancestral, alors que lorsqu'il y a ethnogenèse un nouvel assemblage culturel émerge par mélange d'éléments de deux ou plusieurs assemblages contemporains entre eux. »

L'ethnomusicologie est une discipline s'intéressant traditionnellement, en tant que descendante de la **musicologie comparée**, aux musiques non occidentales et aux musiques occidentales non savantes. Les contours de cette discipline sont cependant très variables. L'ethnomusicologie est parfois caractérisée par l'importance de l'approche anthropologique, ou encore par la spécificité de ses méthodes. Je n'entrerai donc pas dans une discussion plus détaillée de ce terme.

ethnonyme [p. 73, 159, 190]

En anthropologie, l'ethnonyme est le nom donné à une ethnie. Je privilégie ici l'utilisation de l'endonyme, c'est-à-dire le nom utilisé par cette ethnie pour se désigner elle-même.

évolution [p. 20, 22, 282, 340, 346, 356]

En **biologie de l'évolution**, l'*évolution* désigne, dans son sens le plus général, une modification au cours du temps. En pratique cette transformation est souvent pensée de manière populationnelle : il s'agit alors de la modification des caractéristiques d'une population au cours des générations successives.

En anthropologie culturelle, le concept d'*évolution* est largement associé à l'idée que toutes les cultures se développent en passant par les mêmes stades successifs, formant le courant *évolutionniste* dominant la fin du XIX^{ème} siècle en anthropologie. Le terme porte donc encore, pour nombre d'anthropologues, des connotations progressistes, ethnocentristes et finalistes.

On voit cependant renaître aujourd'hui un courant s'intéressant à l'*évolution culturelle* où le terme d'*évolution* se rapproche plus de son sens biologique que de celui qui a cours en anthropologie culturelle. Au sein de ce courant, le terme est parfois ambigu et peut désigner aussi bien la modification au cours du temps d'un élément culturel que l'apparition de cet élément culturel au sein des populations humaines en lien avec l'évolution biologique humaine, comme illustré dans l'entrée *évolution musicale*.

Lorsque j'emploie le terme, c'est toujours dans le sens biologique actuel, c'est-à-dire pour évoquer une modification des caractéristiques d'une **population statistique** au cours des générations successives.

évolution artificielle [p. 87, 368]

On parle d'*évolution artificielle* lorsqu'un processus d'**évolution** est simulé, notamment à des fins scientifiques. La simulation peut-être numérique si elle est calculée par ordinateur à partir d'un modèle informatique⁷³ ou analogique si elle est réalisée par des expériences de **transmission en chaîne**.

évolution musicale [p. 280, 340]

Le terme d'*évolution musicale*, ou *évolution de la musique*, peut être entendu de deux manières :

73. ↑ Voir par exemple le travail de Atkinson et coll. [28, pp. 209–213].

- La modification des caractéristiques d'un ensemble de pratiques musicales au cours des générations successives.
- L'apparition chez l'homme des pratiques musicales et les modifications biologiques (évolution biologique) qui l'ont permise, telles que l'évolution du système nerveux permettant la musicalité ou l'évolution morphologique du système phonatoire⁷⁴.

Je préfère pour ma part parler de *l'évolution biologique des aptitudes musicales* en ce qui concerne le deuxième aspect, réservant le terme d'*évolution musicale* au premier.

figure rythmique [p. 341]

Voir *formule rythmique*.

forme vocale [p. 73, 114, 353, 361]

En *ethnomusicologie* la forme vocale est un *paramètre* caractérisant, dans une pièce comportant une opposition entre un ou une soliste et un chœur, la relation entre ces deux parties vocales [13, pp. 58–59]. Si la forme vocale est *antiphonale*, chaque phrase énoncée par le ou la soliste est reprise par le chœur. Si la forme vocale est *responsoriale*, la partie du chœur (appelée *répons*) est différente de la partie soliste.

formule rythmique [p. 71–73, 108, 113, 142, 181, 245, 270–272, 280, 287–289, 293, 302, 318, 333, 341, 353, 357, 360, 362]

Dans le cas de *pièces* musicales *cycliques*, le terme de *formule rythmique* désigne « l'organisation rythmique de la *période*, à savoir le mode de répartition, à l'intérieur de ce cadre, des événements rythmiques » [19, p. 290].

Lorsqu'une pièce comporte plusieurs *parties* constitutives, on peut caractériser la formule rythmique de chacune d'entre elles. Dans le cas des musiques du Gabon, une ou quelques unes des parties instrumentales sont plus épurées et plus répétitives que les autres. Ces parties, généralement réalisées par des tambours (et parfois par des idiophones) ont un rôle de repère rythmique pour le reste des parties et constituent en cela la *référence rythmique*. Aussi j'utilise parfois le terme de *formule rythmique* sans préciser la partie concernée : il s'agit alors de la formule rythmique de la partie constituant la référence rythmique.

On trouve aussi parfois le terme de *figure rythmique* [19, pp. 290–291]. Pour ma part j'utilise le terme de *formule rythmique* en français et le terme *rhythmic figure* en anglais.

glottochronologie [p. 24]

La glottochronologie est une méthode de *linguistique historique* proposée en 1952 par Swadesh [A404]⁷⁵ pour calculer les dates de séparation entre plusieurs langues ou groupes de langues. Elle suppose que la vitesse de modification des éléments lexicaux d'une langue

74. ↑ La discipline s'intéressant à ce dernier aspect est parfois appelée *musicologie évolutionnaire*. Voir par exemple l'ouvrage de Wallin et coll. (2000) [437]

75. ↑ Citée par Atkinson et Gray [27, p. 520].

est constante au cours du temps et permet donc d'estimer ces dates de divergence. Cette méthode évoque le principe de l'horloge moléculaire utilisé en biologie pour dater les nœuds internes des **phylogénies** construites à partir de caractères moléculaires.

groupe externe [p. 35, 328, 357]

Dans l'analyse **cladistique**, un groupe externe⁷⁶ est un **taxon** choisi pour **polariser** les **caractères** analysés. On choisit des groupes que l'on sait plus éloignés de tous les taxons étudiés qu'ils ne le sont entre eux. L'**état ancestral** de chaque caractère est alors supposé être l'état observé de ce caractère chez ces groupes externes. Dans l'exemple de la figure 1.4 page 37, le groupe externe choisi est le pin blanc de Provence⁷⁷. On observe que le caractère C5 « Bois secondaire » est présent chez le pin : on suppose donc que cet état est l'état ancestral et par conséquent que l'absence de bois secondaire est un **état dérivé**.

histoire [p. 268, 342, 349, 356]

Tucker [424, p. 275] propose la séparation suivante entre *histoire* et *historiographie* :

« *Historiography* means here the scientific study of the past, its methodologies and practices. *Historiography* also means the representation of history (the past) that historians infer. *History* means the past. »⁷⁸

Tucker (2012) [424, p. 275]

Je reprendrai ici cette suggestion de séparer le passé tel qu'il a existé et qui demeure inaccessible (l'*histoire*), de la représentation du passé construite à partir de l'étude scientifique (l'*historiographie*). Je distinguerai cependant les deux sens de ce dernier terme en parlant de **sciences historiographiques** concernant les disciplines scientifiques se chargeant d'étudier le passé. En plus de préciser le sens de mon propos, cette distinction me permet d'introduire un pluriel mettant l'accent sur la diversité des disciplines, de leurs méthodes, de leur ancrage social et de leur histoire.

historiographie [p. 268, 329, 342, 356]

L'historiographie est entendue ici comme la représentation du passé inférée par les **sciences historiographiques**. Elle se distingue donc de l'*histoire* qui désigne le passé tel qu'il s'est déroulé et sera à jamais inaccessible : elle ne peut être que supposée à partir des indices actuels. L'historiographie est donc un construit social, résultat de l'**inférence historique**.

homologie [p. 23, 35, 268, 305, 327, 343]

Dans le vocabulaire de la **cladistique**, une homologie est une similarité due à une ascendance commune. C'est donc une forme particulière d'**analogie**. Elle s'oppose à l'**homoplasie**. Les

76. ↑ Ou *extra-groupe*, ou encore *outgroup* en anglais.

77. ↑ Cet exemple ne montre qu'un seul groupe externe, mais il est nécessaire de faire appel à plusieurs groupes externes pour lever toute ambiguïté [107, pp. 46–49].

78. ↑ « *Historiographie* signifie ici l'étude scientifique du passé, ses méthodes et pratiques. *Historiographie* signifie aussi la représentation de l'histoire (le passé) que les historiens infèrent. *Histoire* signifie le passé. »

pattes des oiseaux et celles des mammifères sont issues de l'ancêtre commun à tous les amniotes et sont donc des caractères *homologues*.

On ne peut généralement pas savoir *a priori* si deux caractères sont homologues. On commence donc par supposer, à partir de la constatation de l'analogie entre deux caractères et des connaissances sur les objets étudiés, l'homologie entre les deux caractères : c'est l'*homologie primaire*. Il est cependant très rare que tous les caractères supposés homologues puissent être expliqués par un même arbre phylogénétique et certaines hypothèses d'homologie formulées *a priori* seront donc invalidées par l'analyse cladistique. Si l'analyse cladistique confirme l'homologie entre ces caractères, on parle alors d'*homologie secondaire*.

homoplasie [p. 22, 23, 32, 79, 327, 334, 336, 342, 348, 352, 354]

Dans le vocabulaire de la cladistique, une homoplasie est une similarité qui n'est pas due à une ascendance commune. C'est donc une forme particulière d'analogie. Elle s'oppose à l'homologie.

Les ailes des insectes, des oiseaux et des chauves-souris sont analogues, mais ne sont pas issues d'un ancêtre commun. Chacune de ces structures est apparue indépendamment au cours de l'évolution : elles sont donc homoplasiques.

L'homoplasie peut être expliquée par un processus de convergence (comme dans l'exemple ci-dessus) mais aussi par un événement de transmission horizontale ou par la réversion d'un état de caractère [251, p. 540].

horizontale (transmission) [p. 361]

Voir *transmission horizontale*.

IC [p. 70, 80]

Voir *indice de cohérence*.

indice de Bremer [p. 70, 81]

Dans une analyse cladistique, l'indice de Bremer [A64]⁷⁹ est le nombre de pas minimum à ajouter à l'arbre le plus parcimonieux pour qu'un clade donné disparaisse [280, p. 168].

On trouve parfois le nom d'*indice de chute du clade*^{80,79} ou d'*indice de support de branche*^{81,79}.

indice de cohérence [p. 60, 67, 73, 79, 262, 266, 272, 286, 334, 335, 343, 344]

L'indice de cohérence (IC)⁸² est un indice statistique proposé par Kluge et Farris [A226]⁸³ et utilisé lors d'une analyse cladistique pour estimer le degré d'adéquation entre la matrice de caractères et l'arbre phylogénétique produit par l'analyse [107, pp. 117–118].

79. ↑ Citée par Marwick [280, p. 168].

80. ↑ *Decay index value*, en anglais.

81. ↑ *branch support value*, en anglais.

82. ↑ *consistency index*, en anglais.

83. ↑ Citée par Darlu et Tassy [107, p. 117].

Il est calculé en divisant le nombre minimal de changements d'états de caractères nécessaires pour expliquer l'ensemble des données de la matrice de caractères (R) par le nombre de changements d'états nécessaires sur l'arbre considéré (L) : $IC = \frac{R}{L}$.

Sur l'exemple de la figure 1.4 page 37, il faut au minimum 5 changements d'états pour expliquer l'ensemble de la matrice : 0 pour le caractère C2, 1 pour les caractères C1, C4 et C5 et 2 pour le caractère C3. Sur l'arbre 2 (le plus parcimonieux), on a besoin de 6 changements d'états (marqués par des barres horizontales). L'indice de cohérence de l'arbre 2 vaut donc : $IC = \frac{R}{L} = \frac{5}{6} = 0,83$

En théorie la valeur du CI est comprise entre 0 (cohérence minimale) et 1 (cohérence maximale), mais en pratique cet indice ne peut guère descendre en dessous de 0,38 [A10]⁸⁴. Un autre problème de cet indice est qu'il est augmenté artificiellement par la présence de caractères non informatifs [107, p. 118] (par exemple le caractère C2), ce qui peut être évité en supprimant ces caractères du calcul. Enfin la valeur de cette indice dépend du nombre de taxons comparés [A8]⁸⁵. On lui préfère ainsi souvent l'indice de rétention qui corrige ces défauts.

indice de rétention [p. 60, 67, 73, 79, 266, 272, 344, 345, 361]

L'indice de rétention (IR)⁸⁶ est un indice statistique proposé par Farris [A133]⁸⁷. Comme l'indice de cohérence (IC), il est utilisé lors d'une analyse cladistique pour estimer le degré d'adéquation entre la matrice de caractères et l'arbre phylogénétique produit par l'analyse [107, pp. 117–118].

inférence bayésienne [p. 22, 64, 70, 74, 91, 157, 275, 312, 329, 333, 338, 345]

Les méthodes d'inférence bayésienne⁸⁸ sont des méthodes probabilistes d'inférence historique qui cherchent à trouver, parmi un ensemble d'hypothèses, celle dont la probabilité est la plus grande en tenant compte des données observées. En langage probabiliste, on dit que l'on cherche l'hypothèse dont la probabilité *sachant les données* est la plus grande.

inférence historique [p. 17, 26, 47, 83, 85, 103, 128, 145, 266, 271, 280, 304, 315, 330, 342, 344, 345, 347–349, 358, 359, 368]

Inférer signifie « tirer une conclusion d'une proposition ou d'un fait » [439]. La notion d'inférence ne précise pas quel mécanisme est utilisé pour tirer la conclusion (induction, déduction ou abduction), mais insiste d'une part sur le fait que la conclusion est *construite* à partir des prémisses et que ce processus est *rationnel*.

L'inférence historique est donc entendue ici comme la formulation d'hypothèses sur le passé à partir de l'analyse rationnelle des données disponibles. Ici, il s'agira le plus souvent de données contemporaines et synchroniques entre elles.

84. ↑ Citée par Nunn et coll. [321, p. 3812].

85. ↑ Citée par Darlu et Tassy [107, p. 118].

86. ↑ *retention index (RI)*, en anglais.

87. ↑ Citée par Darlu et Tassy [107, p. 120].

88. ↑ Voir aussi la partie 1.3.7.2 page 39.

J'utilise ce terme plutôt que celui de *reconstruction historique* dont le sens est peut-être plus restreint et que je trouve plus ambigu. À ce sujet, il est intéressant de noter que l'ouvrage anglophone de référence concernant les méthodes *phylogénétiques* soit intitulé *Inferring phylogenies* [136], alors que l'ouvrage francophone qui est la référence en la matière a pour titre *La reconstruction phylogénétique* [107].

inférence phylogénétique [p. 34, 304, 305, 311, 312, 329, 334, 348, 349, 356]

Une inférence phylogénétique est une *inférence historique* concernant la *phylogénie* d'un ensemble d'entités étudiées. C'est donc une forme d'*inférence historique* dont le résultat peut être exprimé sous la forme d'un *arbre phylogénétique*, d'un *réseau* phylogénétique, ou de toute autre manière.

inférence phylogéographique [p. 26, 40, 275, 317]

J'entends ici le terme d'*inférence phylogéographique* comme l'*inférence* de la distribution géographique passée à partir des connaissances établies concernant un certain nombre d'objets apparentés. Un exemple d'application d'une méthode d'inférence phylogéographique est détaillé dans le chapitre 5⁸⁹. Cette signification se distingue de l'acceptation usuelle du terme de *phylogéographie*, plus générale⁹⁰.

IR [p. 70, 80]

Voir *indice de rétention*.

itération [p. 162, 166, 338]

Dans le cadre de l'*inférence bayésienne*, le terme d'*itération* désigne chacune des étapes élémentaires de l'analyse. À chaque itération, les valeurs des paramètres dont on recherche la valeur optimale sont légèrement modifiées, constituant un nouvel *état*. Cette notion est expliquée et illustrée plus en détail dans la partie 5.2.3.1 page 162.

jodel [p. 252, 329, 366]

D'après Walter Wiora⁹¹ :

« Im wissenschaftlichen Sprachgebrauch bedeutet Jodeln textloses Singen mit fortwährendem Registerwechsel. Entgegen der Norm des Registerausgleichs springt hier die Stimme immer wieder vom “Brust”- in das “Kopf”-Register (bei Männern in Fistel oder Falsett) zurück, indem sie von dem einen zum andern unvermittelt umschlägt oder umbricht. Sie hält sich dabei an musikalische Tonstufen [...]. Gewöhnlich jodelt man [...] auf Jodelsilben. »⁹² Wiora (1958) [A445, p. 73]⁹¹

89. ↑ Pages 155–185.

90. ↑ Voir par exemple Crisci et coll. [101, p. 13].

91. ↑ Citée par Fürniß [152, p. 168].

92. ↑ « Dans la terminologie scientifique, “jodel” désigne un chant sans paroles, avec alternance permanente de registres. Contrairement à la norme d'égalisation des registres, la voix saute constamment du registre de “poitrine” à celui de “tête” (chez les hommes, voix de sifflet ou de fausset), allant de l'un à l'autre sans transition. Elle suit des degrés musicaux [...]. Normalement, on [...] jodle sur des [...] syllabes propres au jodel. »⁹³

93. La traduction est de Fürniß [152, p. 168].

On trouve parfois aussi l'orthographe *yodel*.

lexicostatistique [p. 24, 26, 83]

La lexicostatistique est une méthode de **linguistique historique** proposée dans les années 1950 par Swadesh [A404, A405]⁹⁴ consistant à inférer les relations entre langues à partir du calcul de pourcentage de **cognats** partagés entre les langues. Cette méthode peut faire penser aux **méthodes de distances** utilisées en **biologie de l'évolution**.

lieu variant [p. 29, 53, 55, 86, 369]

En **critique textuelle**, un lieu variant est un « passage du texte, pour lequel tous les manuscrits ne sont pas d'accord ; en tout lieu variant, les différents manuscrits présentent au moins deux variantes » [318, p. LXI]⁹⁵.

linguistique historique [p. 18, 47, 103, 335, 337, 341, 346, 347, 367, 368]

Lyle Campbell définit ainsi la linguistique historique :

« **historical linguistics** (sometimes called **historical and comparative linguistics**) The study of **language change**, of how and why languages change. »⁹⁶
Campbell et Mixco (2007) [77, p. 77]

C'est donc une discipline qui s'intéresse à l'histoire des langues (en particulier à **l'évolution** des langues) et aux processus qui structurent cette histoire.

matrice de caractères [p. 35, 70, 73, 79, 330, 343, 344, 354, 360]

En **systématique biologique**, une matrice de caractères est une matrice indiquant, pour chaque **taxon** étudié, l'**état** de chacun des **caractères** analysés. Un exemple de matrice de caractères est donné dans la figure 1.4 page 37.

matrice de distances [p. 24, 54, 71, 72, 77, 80, 90, 91, 253, 254, 311, 347, 348, 357, 360, 366, 369]

Une matrice de distances est une matrice donnant, pour chaque paire d'objets comparés, une mesure de distance entre ces objets. La distance mesure le degré de différence de ces deux objets, selon le critère choisi. Ainsi deux objets complètement dissemblables auront une distance de 1 alors que deux objets identiques auront une distance de 0.

Les matrices de distances sont fréquemment utilisées pour comparer un ensemble d'objets du même type. On compare tous les objets deux à deux, on mesure la distance entre ceux-ci (grâce à une mesure choisie en fonction de l'objectif) et on place la valeur dans la colonne correspondant au premier objet et dans la ligne correspondant au second. On obtient ainsi une matrice *carrée* (c'est-à-dire qui contient autant de lignes que de colonnes) correspondant

94. ↑ Citées par Atkinson et Gray [27, p. 520].

95. ↑ Citée par Macé et coll. [276, p. 6].

96. ↑ « **linguistique historique** (parfois appelée **linguistique historique et comparative**) L'étude du changement linguistique, de comment et pourquoi les langues changent. »

au nombre d'objets comparés. La plupart des mesures de distance sont *commutatives* : la distance entre A et B est la même que celle entre B et A. Dans ce cas, la case $[A : B]$ de la matrice est identique à la case $[B : A]$. Une moitié de la matrice est donc redondante et peut être omise : on obtient alors une matrice *triangulaire*.

matrice de similarités [p. 53]

Une matrice de similarité est une matrice donnant, pour chaque paire d'objets comparés, une mesure de similarité entre ces objets. C'est conceptuellement la même chose qu'une **matrice de distances**, la similarité étant mathématiquement le complément de la distance (ces deux mesures prenant généralement des valeurs entre 0 et 1). Ainsi si deux objets A et B ont une similarité de 0, ils auront une distance (ou une dissimilarité) de 1. Si ils ont une distance de 0,2, ils auront une similarité de 0,8.

matrilinéaire [p. 73, 365]

Voir *système de filiation*.

maximum de parcimonie [p. 354]

Une méthode de maximum de parcimonie est une méthode qui cherche à maximiser la **parcimonie**, c'est-à-dire à minimiser le nombre d'hypothèses *ad hoc* nécessaires pour expliquer les observations. Le terme désigne parfois directement les méthodes **cladistiques**, qui sont les plus répandues des méthodes de maximum de parcimonie.

maximum de vraisemblance [p. 22, 370]

Les méthodes dites de maximum de vraisemblance^{97,98} sont des **méthodes probabilistes d'inférence historique** qui cherchent à trouver l'arbre dont la probabilité d'engendrer (sous un certain modèle d'évolution choisi *a priori*) les données observées est la plus grande. En langage probabiliste, elles cherchent donc à trouver l'arbre dont la vraisemblance *sachant les données* est la plus grande.

MCMC [p. 91, 158, 162, 349]

Sigle pour *Markov chain Monte Carlo*. Voir *méthode de Monte-Carlo par chaînes de Markov*.

méthode comparative (linguistique historique) [p. 22, 61, 76, 83]

La méthode comparative est une méthode de **linguistique historique** utilisée pour comparer un ensemble de langues entre elles, déterminer leurs relations de filiation et retrouver leurs formes ancestrales. Le principe de cette méthode est présenté dans la partie 1.3.3 pages 26 à 29.

méthode comparative phylogénétiquement informée [p. 64, 76, 78, 84, 260, 354, 355]

Les méthodes comparatives phylogénétiquement informées⁹⁹ (MCPIs) sont des méthodes

97. ↑ *Maximum likelihood*, en anglais.

98. ↑ Voir aussi la partie 1.3.7.1 page 39.

99. ↑ *Phylogenetic comparative methods (PCMs)*, en anglais.

d'**inférence historique** basées sur la comparaison. Cependant, contrairement aux méthodes d'**inférences phylogénétiques**, elles ne permettent pas elles-mêmes d'inférer des relations historiques. Elles se servent de relations déjà établies (généralement d'un **arbre**, issu de l'analyse de données biologiques ou culturelles) pour analyser la distribution ou la structuration d'autres données (souvent culturelles, dans les cas qui m'intéressent) en tenant compte des informations apportées par ces relations.

Ainsi Holden et Mace [A197] cherchent à savoir si le type de **système de filiation** coévolue avec le pastoralisme dans des sociétés africaines. La corrélation entre ces deux paramètres ne peut être testée directement car les différentes sociétés ne sont pas indépendantes : elles ont une histoire commune (c'est ce qu'on appelle le *problème de Galton*). Aussi la corrélation entre la patrilinéarité et le pastoralisme peut simplement venir de l'héritage commun de ces populations. Les auteur.e.s se servent donc d'un arbre des relations entre les langues de ces populations, inféré au cours d'une étude précédente [196], pour prendre en compte l'histoire commune des populations lors du test de corrélation.

méthode de compatibilité [p. 21, 63, 87, 334]

La méthode de compatibilité est une méthode d'**inférence phylogénétique** basée sur la comparaison de caractères biologiques et faisant appel au principe de **parcimonie** [107, p. 145]. Elle ressemble par ces aspects aux méthodes **cladistiques**, mais elle diffère en ce qu'elle refuse le recours à toute hypothèse d'**homoplasie** dans l'évaluation des différents arbres expliquant le mieux les états de caractères observés. La méthode de compatibilité est présentée plus en détail dans la partie 1.3.5 page 31.

méthode de distances [p. 34, 54, 69, 346, 355, 357, 369]

Les méthodes de distances (ou méthodes **phénéétiques**) sont des méthodes d'**inférence phylogénétique** fondées sur le calcul de la distance globale entre les **taxons** étudiés. Pour chaque couple de taxons, une mesure de distance est calculée, formant ainsi une **matrice de distances**. Cette matrice est ensuite utilisée pour inférer des relations historiques entre les taxons. Il existe de nombreux algorithmes permettant de construire un arbre de proximité à partir d'une matrice de distances, par exemple l'algorithme du **plus proche voisin**.

méthode de l'erreur commune [p. 336, 348]

Voir *méthode de Lachmann*.

méthode de Lachmann [p. 19, 25, 348, 362]

La méthode de Lachmann (aussi appelée **méthode de l'erreur commune** [366, p. 11]) est une méthode élaborée en critique textuelle au cours du XX^{ème} siècle. Bien que cette méthode soit inspirée des travaux de Lachmann [A231]¹⁰⁰, elle n'a pas été élaborée par celui-ci, mais principalement par Maas [A272]¹⁰¹. Bordalejo en propose la définition suivante :

100. ↑ Citée par Timpanaro [A415, pp. 85–88], elle-même citée par Robins [353, p. 91].

101. ↑ Citée par Salemans [366, pp. 18–19].

« Independently of who formulated these ideas, the method which uses common errors in manuscripts to determine if they are or are not related and that assumes that the process of understanding the textual tradition must be carried out before one can decide which variants are to enter the emended text is usually called the Lachmann method [...]. »¹⁰² Bordalejo (2003) [49, p. 41]

Cette méthode est peut-être la première des méthodes **stemmatiques** et donnera naissance à plusieurs variantes, comme expliqué plus en détail dans la partie 1.2.1 pages 18 à 20.

méthode de Monte-Carlo par chaînes de Markov [p. 40, 91, 162, 347]

La méthode de Monte-Carlo par chaînes de Markov (**MCMC**) permettent de rechercher les valeurs d'un ensemble de paramètres maximisant une certaine fonction. Cette méthode est décrite dans la partie 5.2.3 page 162.

méthode issue de l'inférence phylogénétique [p. 47, 52, 58, 59, 62, 64–66, 69, 71–76, 81–85, 87, 88, 92, 103, 142, 273, 283, 286, 303, 304, 307, 308, 330, 352, 356]

J'entends par ce terme toute méthode ayant été développée à des fins d'**inférence phylogénétique**, c'est-à-dire afin de faire une **inférence** sur l'**histoire** d'un groupe d'organismes vivants et donc développée dans le cadre de la **systematique biologique**. Cela ne présage pas pour autant du but dans lequel est utilisée la méthode ou des données sur lesquelles elle est appliquée.

méthode probabiliste [p. 22, 63, 91, 158, 344, 347, 358, 370]

En **biologie de l'évolution**, les méthodes probabilistes sont des méthodes d'**inférence historique** basées sur une conception probabiliste de l'évolution, selon laquelle les modifications observées au cours de l'évolution ne sont ni déterminées, ni complètement aléatoires, mais que leur apparition peut être modélisée par des lois de probabilités.

Le principe de ces méthodes est présenté succinctement dans la partie 1.3.7 page 36.

métricité [p. 108, 114, 138, 181, 288, 350, 357]

Le terme de *métricité* est parfois utilisé comme synonyme de *métrique*. J'utilise ici le terme de *métricité* dans un sens un peu plus restreint. Ainsi je considère que la métricité d'une **pièce** musicale est la jonction de deux **paramètres** : la **périodicité** et la **modalité de subdivision de la pulsation**. Ce paramètre permet donc de décrire l'organisation métrique d'une pièce musicale. Ainsi si « P » est la nombre de **pulsations** dans chaque **période** et « S »¹⁰³ est le nombre de **valeurs opérationnelles minimales** dans chaque pulsation, la métricité est notée « P/S ».

102. ↑ « Indépendamment de qui a formulé ces idées, la méthode qui se sert des erreurs communes à différents manuscrits pour déterminer s'ils sont ou ne sont pas apparentés et qui suppose que la procédure de compréhension de la tradition textuelle doit être effectuée avant de décider des variantes qui devront figurer dans le texte corrigé est habituellement appelée la méthode de Lachmann [...]. »

103. ↑ La lettre « S » est choisi en référence au terme *subdivision*.

Par exemple chez les Ngom du village de Moualo, la métricité des pièces du répertoire *Makouma* est de huit pulsations de subdivision binaire et notée « 8/2 ».

métrique [p. 73, 140, 189, 271, 275, 292, 318, 349, 353, 358, 361, 460]

D'après Arom et coll. [22, p. 934] :

« La métrique, au sens où je l'entends, [est] *la détermination, par rapport à une valeur constante prise comme étalon, des intervalles de temps* obtenus soit par multiplication, soit par division, mais – et ce point est essentiel – *d'où est absente toute idée de hiérarchisation.* »
Arom (2007) [17, p. 934]

La métrique constitue donc un cadre temporel de référence alors que le rythme s'intéresse à la manière dont les durées sont réparties au sein de ce cadre :

« la métrique concerne *l'étalonnage du temps en quantités – ou valeurs – égales ; le rythme, les modalités selon lesquelles des durées différenciées sont regroupées.* »
Arom (2007) [17, p. 934]

Dans le cas de musiques **cycliques**, comme celles que j'étudie dans cette thèse, la métrique peut être caractérisée principalement par deux **paramètres** musicaux : la **périodicité** et la **modalité de subdivision de la pulsation**. Je regroupe ici ces deux paramètres sous le terme de **métricité**.

modalité de subdivision de la pulsation [p. 291, 330, 349, 350, 353, 358, 362, 366, 369]

La modalité de subdivision de la pulsation décrit la manière dont la **pulsation** d'une **pièce** musicale peut-être subdivisée. Si elle est divisée en un nombre paire de **valeurs opérationnelles minimales** (V.O.M.) (généralement 2 ou 4), la subdivision de la pulsation est dite **binaire**. Si elle est divisée en un nombre de V.O.M. multiple de 3 (généralement 3 ou 6), la subdivision de la pulsation est dite **ternaire**.

modèle [p. 108, 111, 128, 282, 378]

En **ethnomusicologie**, le modèle d'une **pièce** musicale est le résultat de l'effort scientifique pour représenter et condenser l'ensemble des caractéristiques pertinentes d'une pièce¹⁰⁴ :

« Par *modèle*, il faut entendre ici une *représentation sonore à la fois globale et simplifiée d'une entité musicale* ; le modèle en accuse la singularité *en ce qu'il condense, en une épure, l'ensemble de ses traits pertinents et eux seuls.* »
Arom (1991) [15, p. 70]

modélisation du système catégoriel vernaculaire [p. 333]

Activité consistant à créer un modèle scientifique du **système catégoriel** utilisé par une culture donnée. Il s'oppose en cela à la **catégorisation musicale** qui est (au sens où je l'entends dans ce travail) l'activité de construction d'un système catégoriel *scientifique*.

104. ↑ Voir la partie 3.1.3 pages 109 à 113 pour une présentation plus détaillée du *modèle* et la partie 3.3 pages 118 à 128 pour une discussion détaillée de ce concept et de ceux qui en sont issus.

morphème [p. 27]

En linguistique, un morphème est la plus petite unité porteuse de sens.

multifurcation [p. 56, 328, 337, 352]

Une multifurcation est un **nœud interne** d'un **arbre** connecté à quatre autres nœuds ou plus. En d'autres termes, c'est un point où une branche se sépare en au moins trois autres branches. Un arbre ne comportant aucune multifurcation est dit **dichotomique**.

musicologie comparative [p. 351]

Le terme de musicologie comparative est parfois trouvé comme synonyme de *musicologie comparée*, par exemple chez Arom et Alvarez-Pereyre [20, pp. 21–25]. J'ai pour ma part choisi d'utiliser ce second terme. En effet les scientifiques francophones de cette époque semblent utiliser plutôt le terme de *musicologie comparée*, à l'instar de Schaeffner [A372]¹⁰⁵. Ce terme me paraît par ailleurs plus cohérent avec les termes d'*anatomie comparée* ou de *linguistique comparée* alors préférentiellement utilisés par les francophones de ces autres disciplines, et dont la musicologie comparée (en particulier Hornbostel [201, p. 250]) s'est inspirée.

musicologie comparée [p. 130, 340, 351]

Comme l'ont montré (entre autres) Nettl [314, pp. 85–86] et Merriam [291], la discipline aujourd'hui appelée *ethnomusicologie* est la descendante directe (aussi bien du point de vue des théories, des méthodes, que des personnes la pratiquant) de la discipline appelée *musicologie comparée* (ou parfois *musicologie comparative*) avant 1950. S'il n'est pas certain que ce soit Guido Adler qui ait introduit pour la première fois ce terme, c'est en tout cas lui qui généralisa son usage en le définissant dans son ouvrage *Umfang, Methode und Ziel der Musikwissenschaft*¹⁰⁶ [A1]¹⁰⁷.

Au cours d'une intéressante analyse des définitions proposées pour les termes d'*ethnomusicologie* et de *musicologie comparée* et de leur évolution, Alan Merriam montre que les deux disciplines ne se distinguent guère ni par l'objet d'étude ni par les objectifs poursuivis [291], mettant en valeur l'ambiguïté entretenue par cette discipline avec l'analyse comparative, qu'elle a voulu évincer de son nom. De fait, la méthode comparative n'a jamais été au cœur de la musicologie comparée [291, p. 193], et celle-ci ne la distingue pas de la musicologie classique qui y a aussi recours [291, pp. 192–193]. Selon toute vraisemblance, ce changement de nom disciplinaire au cours des années 50 était même résolument orienté contre la dimension comparative de la discipline, les pratiques musicales ne pouvant être comprises, selon ces ethnomusicologues, que dans leur contexte ethnologique, toute comparaison étant alors, sinon dangereuse et proscrite, du moins prématurée [291, pp. 193–194].

Aussi j'utilise les termes d'*ethnomusicologues* et d'*ethnomusicologie* pour la discipline et les

105. ↑ Citée par Mayaud [281, p. 41].

106. ↑ La définition d'Adler est donnée dans la partie 4.1 page 130.

107. ↑ Citée par Nettl [314, p. 72].

scientifiques s’y dédiant, y compris avant les années 1950 où le terme n’était pas encore en usage.

mutuellement compatibles (caractères) [p. 32, 34, 335]

D’après Pierre Darlu et Pascal Tassy :

« Des caractères sont dits mutuellement compatibles quand il existe un arbre qui rende compte des changements d’états de ces caractères sans nécessiter d’hypothèses d’**homoplasie**. » Darlu et Tassy (1993) [107, p. 9]

mythème [p. 86]

Dans l’étude des traditions orales, en particulier des mythes, un mythème correspond à une unité de description d’un conte. d’Huy [121, p. 94] le définit ainsi : « irreducible and unchanging element of the myth that can be similar or different between two versions »¹⁰⁸.

Dans l’application des **méthodes issues de l’inférence phylogénétique** (en particulier des méthodes **cladistiques**) à l’étude des mythes, la notion de mythème est utilisée de manière équivalente à celle de **caractère**. J’étends ici la notion de mythème à la description d’autres traditions orales, par exemple pour l’analyse de contes comme dans l’étude de Tehrani [409].

neighbor joining [p. 54, 58, 59, 69, 70, 72, 90, 311, 360]

Voir *plus proche voisin*.

NeighborNet [p. 54, 67, 80, 254, 360]

Voir *réseau de voisins*.

nœud externe [p. 55, 328, 352]

Dans un **arbre** topologique, un nœud est dit *externe* lorsqu’il n’est connecté qu’à un seul autre nœud. Il s’oppose en cela à un **nœud interne**. On parle aussi de *feuille*.

nœud interne [p. 55, 56, 62, 184, 275, 328, 337, 351, 352]

Dans un **arbre** topologique, un nœud est dit *interne* lorsqu’il est connecté à au moins trois autres nœuds. Il s’oppose en cela à un **nœud externe**. Un nœud interne est appelé une *bifurcation* s’il est connecté à exactement trois autres nœuds et une *multifurcation* s’il est connecté à plus de trois autres nœuds.

paramètre (musical) [p. 86, 108, 112, 114, 116, 251, 275, 287, 304, 305, 330, 339, 341, 349, 350, 353, 357, 358, 360, 366, 369]

En **ethnomusicologie**, un paramètre musical est une caractéristique musicale observable et que l’on peut décrire dans un cas particulier en lui affectant une **valeur**. Par exemple le paramètre « **subdivision de la pulsation** » prend la valeur « ternaire » dans le **répertoire**

108. ↑ « Élément invariable et irréductible du mythe qui peut être similaire ou différent entre deux versions. »

Lisimbu chez les Awandji de Doumé et prend la valeur « binaire » dans le répertoire *Lisimbu* chez les Ngom de Bolapessa.

Cette notion peut être rapprochée de celle utilisée en biologie de *caractère*¹⁰⁹. Je choisis donc d'utiliser le terme de *paramètre* lorsqu'il s'agit de données de nature musicale et celui de *caractère* lorsque j'évoque des données de nature biologique. Suivant la proposition de Le Bomin et coll. [240], je distingue ici les *paramètres intrinsèques*, les *paramètres performatifs* et les *paramètres contextuel*.

paramètre contextuel [p. 181, 353]

En *ethnomusicologie*, les paramètres contextuels [240, p. 4] (ou *paramètres extramusicaux* [21, p. 286]) sont les *paramètres* musicaux relatifs aux éléments contextuels, sociaux et symboliques, associés à la performance musicale. Le *répertoire* dans lequel s'inscrit une pièce est ainsi un paramètre contextuel.

paramètre (musical) externe [p. 353]

Voir *paramètre performatif*.

paramètre extramusical [p. 353]

Voir *paramètre contextuel*.

paramètre (musical) interne [p. 72, 138, 140, 145, 181, 275, 353]

Voir *paramètre intrinsèque*.

paramètre (musical) intrinsèque [p. 73, 245, 286, 353]

En *ethnomusicologie*, un paramètre musical intrinsèque [240, p. 4] (ou *paramètre interne* [21, p. 300]) est un *paramètre* issu de l'analyse musicale du matériel sonore. Il se distingue ainsi du *paramètre performatif* (ou *externe*) et du *paramètre contextuel* (ou *extramusical*). Les organisations *métrique* (*périodicité*, *modalité de subdivision de la pulsation*), *rythmique* (*figures* et *cellules rythmiques*) et *scalaires* (*échelles musicale*) sont des paramètres intrinsèques.

paramètre performatif [p. 73, 181, 353]

En *ethnomusicologie*, les paramètres performatifs [240, p. 4] (ou *paramètre externe* [21, p. 300]) sont les *paramètres* relatifs aux techniques de la performance musicale. Ainsi les instruments de musique, les *procédés polyphoniques*, les *formes vocales* ou encore les *techniques vocales* sont des paramètres performatifs.

parcimonie [p. 21, 36, 40, 57, 87, 91, 180, 304, 347, 348]

Le principe de parcimonie, parfois appelé *rasoir d'Ockham*, préconise d'utiliser le plus petit nombre d'hypothèses *ad hoc*¹¹⁰ pour expliquer les données. Ce principe est au fondement

109. ↑ Elles se ressemblent mais ne sont pas strictement équivalentes, comme je le discute dans la partie 9.3.1 page 304.

110. ↑ C'est-à-dire des hypothèses formulées spécifiquement pour expliquer les données observées.

des méthodes de **maximum de parcimonie** et en particulier des méthodes **cladistiques**. Il s'agit alors de minimiser le nombre de changements d'états de **caractères** nécessaires pour expliquer les données de la **matrice de caractères**, en particulier le nombre d'**homoplasies**.

partie (constitutive) [p. 112, 121, 284, 338, 341, 358, 362]

En **systématique musicale**, le terme de *partie* désigne chacune des structures mélodico-rythmiques distinctes constituant une **pièce** musicale.

patrilinéaire [p. 73, 365]

Voir *système de filiation*.

patrimoine musical [p. 84, 107, 108, 113, 114, 181, 184, 275]

D'après Le Bomin et coll. [240], un patrimoine musical est « a set of musical features shared by a population »¹¹¹ [240, p. 2]. Plus précisément :

« A patrimony is defined as a set of everything that is necessary for performing music in a given population. As music is performed in specific contexts, patrimonies also encompass contextual (e.g. circumstances, functions) and linguistic (e.g. names of repertoires and instruments) data. »¹¹²

Le Bomin et coll. (2016) [240, p. 4]

PCM [p. 355]

Sigle pour *phylogenetic comparative method*. Voir *méthode comparative phylogénétiquement informée*.

période [p. 336, 341, 349, 354, 358, 362]

En **ethnomusicologie**, dans le cas d'une **pièce** musicale **cyclique**, une période (ou **cycle**) désigne chacune des répétitions d'un matériau mélodico-rythmique équivalent. La longueur de la période, caractérisée par le nombre de **pulsations** qu'elle contient, définit la **périodicité**.

périodicité [p. 349, 350, 353, 354, 358]

D'après Simha Arom :

« **période/périodicité** : de par leur caractère **cyclique** [...], la majorité des musiques d'Afrique subsaharienne s'appuient sur un principe de *périodicité*. L'unité de base de leur organisation temporelle est la *période*, boule de temps délimitée par la répétition – immédiate ou différée – d'événements rythmiques ou mélodico-rythmiques dont elle constitue l'armature temporelle. [...] Dans les musiques d'Afrique subsaharienne, au sein d'ensembles composés d'un effectif hétérogène, il y a fréquemment superposition de périodicités différentes, mais dont les ratios demeurent toujours simples. Leur point de jonction commun est appelé *macropériode*. »

Arom (2007) [16, p. 412]

111. ↑ « un ensemble de caractéristiques musicales partagées par une »

112. ↑ « Un patrimoine est défini comme l'ensemble de tout ce qui est nécessaire pour la réalisation musicale dans une population donnée. »

périodique [p. 336]

Voir *cyclique*.

permutation tail probability [p. 81, 358, 366]

Voir *test de longueur par permutation*.

phénétique [p. 21, 54, 66, 67, 69, 72, 74, 75, 87, 348, 366]

L'adjectif *phénétiq*ue est utilisé pour qualifier les *méthodes de distances* ou les classifications qui en sont issues. Le terme est aussi employé comme nom synonyme de *taxinomie numérique*. Pour ma part, je ne l'utiliserai que comme adjectif.

phylogénèse [p. 66, 355]

En *systématique biologique*, c'est un synonyme de *phylogénie*.

En archéologie, il s'agit d'un processus générant de nouveaux assemblages par ramification à partir d'un assemblage ancestral (voir *ethnogenèse*).

phylogenetic comparative method [p. 78, 354]

Terme anglais (abrégé *PCM*) pour désigner une *méthode comparative phylogénétiquement informée*.

phylogénétique [p. 327, 329, 345, 355, 356, 366]

Relatif à la *phylogénè*se, donc synonyme de *phylogénique*. Si le terme de *phylogénie* est un synonyme plus fréquent de *phylogénè*se, le terme de *phylogénét*ique est généralement préféré à celui de *phylogénique*. En anglais, le nom *phylogenetics* (proche de l'adjectif *phylogenetic*, équivalent à l'adjectif français *phylogénét*ique) désigne la discipline qui cherche à inférer les relations phylogénétiques entre espèces biologiques. En français c'est le terme de *systématique biologique* qui désigne cette discipline¹¹³, mais le terme *phylogénét*ique est parfois utilisé dans ce sens¹¹⁴, peut-être par emprunt à l'anglais. Pour ma part, je n'utilise ce terme que comme adjectif se rapportant à la *phylogénè*se, donc à la *phylogénie*.

phylogénie [p. 20, 83, 329, 342, 345, 355]

D'après Darlu et Tassy [107, p. 1], « [l]e terme “*phylogénie*” fut inventé par Ernst Haeckel en 1866 pour définir l'enchaînement des espèces animales et végétales au cours du temps. ». Il est formé à partir du grec ancien φῶλον (*phýlo*), signifiant *tribu*, *race*. Le mot est aujourd'hui polysémique et peut désigner « le cours historique de la descendance des êtres organisés » [107, p. 1], l'« arbre qui illustre la genèse des phyla » [251, p. 541] ou l'« arbre qui illustre “qui est le plus proche de qui” » [251, p. 541]. Le terme est parfois utilisé pour désigner la discipline cherchant à inférer les relations phylogénétiques entre espèces biologiques, équivalent du terme anglais *phylogenetics* (parfois aussi traduit par *la phylogénét*ique) et qui correspond à certaines acceptations du terme *systématique biologique*.

113. ↑ Du moins depuis qu'elle s'est éloignée du but classificatoire, voir l'entrée *systématique biologique* du glossaire.

114. ↑ On trouve aussi parfois le terme de *phylogénie* pour désigner cette discipline.

Pour Straffon [400, p. 2], « a phylogeny is a hypothetical reconstruction of those evolutionary relationships, built by identifying the distribution of characters among species and inferring their development. That is, a phylogeny constitutes a hypothesis of an evolutionary history. »¹¹⁵.

Une différence majeure entre ces distinctions réside dans la nature ontologique de cette phylogénie. Dans l'acceptation de Darlu et Tassy [107, p. 1], il s'agit d'une *histoire* (au sens défini dans ce glossaire), donc d'une série d'événements qui se sont effectivement déroulés auxquels nous n'aurons jamais accès et que nous ne pourrons qu'inférer. Dans celle de Straffon [400, p. 2] ou dans les sens 2 et 3 de Lecointre et Le Guyader [251, p. 541] présentés ci-dessus, il s'agit d'une *historiographie*, c'est-à-dire d'une représentation construite par les scientifiques de cette histoire.

Le terme de *phylogénie* est aussi utilisé pour parler de tout arbre construit à partir des *méthodes issues de l'inférence phylogénétique*. On peut citer pour exemple les articles de Barbrook et coll. [33] ou de Bordalejo [49], où le terme apparaît dans le titre. Au-delà de la question de la nature ontologique de la phylogénie (discutée ci-dessus et qui se pose aussi dans le cas de données culturelles) c'est ici la nature des données qui interroge. Issu du mot grec pour *race* ancrant le mot dans la systématique biologique, le terme de phylogénie pourrait indifféremment désigner la représentation de l'histoire de musiques, de langues ou de contes dès lors qu'elle a été produite par un algorithme issu de la biologie. Pourtant la question de savoir si une méthode propose une hypothèse pertinente sur une histoire (et même au-delà de savoir ce en quoi consiste la représentation de l'histoire) ne peut être défini et débattu qu'au sein de la discipline s'intéressant à cette histoire.

Pour ma part je conserverai la définition classique de Darlu et Tassy [107, p. 1] établissant la phylogénie comme « le cours historique de la descendance des êtres organisés ». Plus précisément et pour clarifier mes choix quant aux discussions ci-dessus, je considère la phylogénie comme une *histoire* (et non une *historiographie*) concernant les organismes vivants (et donc n'incluant pas des entités culturelles¹¹⁶). Je ne limite pas cette histoire à celle des relations de filiation entre ces êtres vivants même si, dans le cadre du sujet de cette thèse, ce sera souvent ce dont il s'agit. Cette acceptation du terme se répercutera naturellement sur le terme de *phylogénétique* entendu comme « relatif à la phylogénie » et sur les termes en découlant et en particulier l'*inférence phylogénétique* et l'*arbre phylogénétique*.

phylogénique [p. 355]

Relatif à la phylogénie, donc synonyme de *phylogénétique* quoique moins employé.

115. ↑ « une phylogénie est une reconstruction hypothétique de ces relations évolutives, construite par l'identification de la distribution des caractères parmi les espèces et en inférant leur développement. C'est-à-dire qu'une phylogénie est une hypothèse concernant l'histoire évolutive. »

116. ↑ Même si l'histoire de la culture est bien évidemment une partie de l'histoire des êtres humains qui sont des organismes vivants. Et même si les objets culturels sont parfois considérés se comporter comme des organismes vivants, notamment du point de vue de leur transformation (leur *évolution*) au cours du temps.

pièce [p. 107–109, 111–114, 117, 118, 120, 121, 128, 137, 138, 249, 275, 284, 310, 331, 336, 338, 341, 349, 350, 354, 358–360, 369, 370, 458]

En ethnomusicologie, une pièce musicale désigne un ensemble culturellement pertinent de réalisations sonores, caractérisées par un certain nombre de **paramètres** musicaux ou extra-musicaux (**métricité**, **formule rythmique**, **échelle** musicale, texte des paroles, instrumentation, nom endogène, contexte de réalisation, ...). Chacune des réalisations de cette pièce, production sonore particulière et unique, est appelée une **version** de cette pièce¹¹⁷.

« Par pièce, on entend ici la matérialisation, au sein d'une catégorie, d'un énoncé musical identifiable en tant que tel par les tenants de la tradition. »

Arom et coll. (2008) [21, p. 294]

playback recording [p. 358]

Voir *re-recording*.

plésiomorphie [p. 23, 328, 334, 339]

En **cladistique**, une plésiomorphie est l'**état ancestral** d'un caractère, c'est-à-dire que chez l'ancêtre (théorique) de tous les groupes comparés, le caractère en question est dans cet état, qualifié de *plésiomorphe*. Il s'oppose en cela à une **apomorphie** qui est l'état dérivé d'un caractère.

plus proche voisin [p. 90, 348, 352, 360, 369]

L'algorithme du plus proche voisin¹¹⁸, proposé par Saitou et Nei [A365]¹¹⁹ est une **méthode de distances** permettant de construire un **arbre** à partir d'une **matrice de distances**. Les **taxons** séparés par les distances les plus faibles se retrouvent regroupés dans l'arbre. La construction de l'arbre est très rapide, grâce à une approximation donnant d'assez bons résultats [136, pp. 166–168].

polarisation [p. 35, 342]

Dans l'analyse **cladistique**, la polarisation d'un **caractère** est l'identification de son **état ancestral**. Il y a plusieurs manières de polariser un caractère, la plus fréquente étant peut-être la comparaison avec des **groupes externes**.

population statistique [p. 340]

Une population statistique est un ensemble d'objets présentant une certaine pertinence du point de vue de l'analyse effectuée (justifiant le terme d'ensemble), mais dont les caractéristiques étudiées présentent une certaine variabilité. Par exemple l'espèce *Crocus nudiflorus*, constituée d'un ensemble d'individus (chaque plante de cette espèce) dont les caractéristiques biologiques sont variables, est une population statistique. De même, la pièce *Igogo*

117. ↑ Comme discuté dans le paragraphe *Instances et types* page 249, on peut donc considérer qu'une pièce est un *type* alors qu'une **version** de pièce est une *instance*.

118. ↑ *neighbor joining*, en anglais.

119. ↑ Citée par Felsenstein [136, pp. 166–168].

nuga des Galoa de Nombeduma est une population statistique, constituée de l'ensemble des réalisations de cette pièce (les versions) qui présentent une certaine variabilité, mais dont la pertinence de l'ensemble est attestée par son nom endogène.

potentiel historiographique [p. 251, 268]

Le *potentiel historiographique* d'un ensemble de données désigne les possibilités d'**inférence historique** offertes par celles-ci. Cette notion est introduite et détaillée dans le chapitre 7¹²⁰.

probabiliste [p. 74, 75, 87]

Voir *méthode probabiliste*.

procédé polyphonique [p. 73, 353]

En **ethnomusicologie** le procédé polyphonique¹²¹ est un **paramètre** permettant de caractériser, dans une pièce comportant plusieurs parties constitutives, les modalités d'organisation de ces parties [21, p. 281]. On parle d'*homorythmie* lorsque toutes les parties ont un découpage rythmique semblable [22, p. 1094], et de *contrepoint* lorsque l'on distingue plusieurs lignes mélodiques dont l'articulation rythmique est différente [22, p. 1095].

PTP [p. 67, 70, 366]

Sigle pour *permutation tail probability*. Voir *test de longueur par permutation*.

pulsation [p. 108, 293, 349, 350, 354, 360, 369]

Dans le cas de **pièces** musicales mesurées où la **métrique** est régulière, la pulsation désigne l'étalon de temps isochrone¹²² constituant le cadre temporel de la pièce, assurant la synchronisation des différentes **parties** :

« Dans la musique d'Afrique subsaharienne où la notion de “temps fort” n'a pas cours, la *pulsation* délimite des unités isochrones ayant toutes un statut équivalent. Chaque pièce est sous-tendue par une pulsation qui lui est inhérente ; elle en détermine le tempo et assure le synchronisme de son exécution. »

Arom (2007) [16, p. 413]

La **modalité de subdivision de la pulsation** est définie par la manière dont la pulsation peut-être divisée, en particulier par le nombre de **valeurs opérationnelles minimales** obtenues par une telle subdivision. La **périodicité** correspond au nombre de pulsations contenues dans une **période** de la pièce.

re-recording [p. 113, 298, 338, 357]

La technique de *re-recording*¹²³ est une méthode d'**enregistrement analytique** développée

120. ↑ Voir **partie 7.3.1** page 268.

121. ↑ On trouve aussi le terme de *procédé plurilinéaire* [13, pp. 84–101].

122. ↑ C'est-à-dire que toutes les pulsations ont la même durée.

123. ↑ Ou **enregistrement en *playback***, ou encore *playback recording*.

en 1967 par Arom [11]¹²⁴. Elle permet d'enregistrer chaque partie d'une *pièce* musicale sur une piste audio séparée, facilitant ainsi l'analyse musicale et en particulier la transcription.

Cette méthode est décrite succinctement dans la partie 3.1.3.3 page 112. La technique d'*enregistrement continu des parties séparées* est une amélioration de la méthode de *re-recording*.

reconstruction historique [p. 345]

Ce terme est largement utilisé par différentes disciplines historiques désignant en général une proposition (ou le processus ayant abouti à cette proposition) de description des caractéristiques passées d'un objet donné (par exemple une langue, une espèce, ...).

Le terme de *reconstruction*, ou *re-construction*, est intéressant puisqu'il met l'accent sur la *construction*, c'est-à-dire sur l'aspect construit de la connaissance historique : la description proposée est le produit de la construction de l'objet étudié par la science et les connaissances disponibles aujourd'hui. Cependant le choix du préfixe *re-* prête à confusion : il suggère le « fait de ramener en arrière » ou un « retour à un état antérieur » [348]. Si ce préfixe est entendu comme retour vers l'arrière (ce qui signifierait dans notre cas « vers le passé »), il paraît redondant avec l'adjectif *historique*. Si en revanche on le comprend comme un retour à l'état initial, cela suggère que le processus de construction pourrait permettre de retrouver les caractéristiques de l'objet *tel qu'il était*, oubliant que nos connaissances sur le passé ne peuvent qu'être des hypothèses dans lesquelles on peut avoir une confiance plus ou moins grande selon qu'elles soient confortées par un nombre plus ou moins grand d'observations indépendantes.

Pour ces raisons, j'utiliserai pour ma part exclusivement le terme d'*inférence historique*¹²⁵ qui est sans ambiguïté quant à la dimension hypothétique des résultats obtenus. Ce terme me paraît cependant un peu plus général que celui de *reconstruction historique* : l'inférence historique s'intéresse à toute hypothèse qui peut être formulée sur le passé à partir du présent, alors que la reconstruction historique se concentre sur les seules hypothèses concernant la description des caractéristiques passées. Par exemple, l'inférence s'intéresse aussi (et contrairement à la reconstruction) à formuler des hypothèses sur les *processus* évolutifs passés. Si l'on veut garder l'idée plus précise de description des caractéristiques passées, on trouve parfois le terme de *reconstruction hypothétique*, qui me paraît préférable : il conserve grâce au préfixe *re-* l'idée d'hypothèse sur le passé tout en insistant sur l'aspect hypothétique. Encore qu'il faille clairement expliciter ce qui est reconstruit (un ancêtre ? un processus évolutif ?), et selon les cas on pourrait ne plus voir la nuance avec la notion d'*inférence historique*.

Compte tenu de l'état préliminaire du travail d'inférence historique à partir de données musicales, il paraît pour le moins prématuré (sinon illusoire) de vouloir reconstruire des

124. ↑ La technique est décrite plus en détail dans l'ouvrage de Arom (1985) [13, pp. 190–213].

125. ↑ J'éviterai de manière générale le terme de *reconstruction* pour préférer celui d'*inférence*.

objets musicaux.

répertoire [p. 73, 107, 108, 113, 114, 117, 190, 270, 275, 350, 352, 353]

En **systematique musicale**, un répertoire est une ensemble de **pièces** musicales réalisées à l'occasion d'un contexte culturel de performance précis, identifié le plus souvent par un nom endogène. En Afrique centrale, un répertoire est souvent caractérisé par un **paramètre** musical précis¹²⁶.

Par exemple l'ensemble des pièces réalisées lors de la cérémonie d'initiation au *Lisimbu* chez les Bongo du village de Boumango constitue un répertoire appelé, par métonymie, le répertoire du *Lisimbu*. Il est caractérisé par sa métricité (huit **pulsations** de subdivision **ternaire**) ainsi que par les **formules rythmiques** des tambours.

réseau [p. 54, 80, 81, 86, 89, 286, 328, 329, 345, 360, 361, 367]

Dans la théorie mathématique des graphes, un réseau est un ensemble de nœuds (ou sommets) tous reliés entre eux par des arêtes. Il se distingue de l'**arbre** en ce qu'il peut être cyclique : on peut relier deux points donnés par plusieurs chemins différents. Les réseaux sont utilisés pour représenter des processus d'évolution qui permettent une **transmission horizontale** et qui ne sont donc pas strictement **arborescents**.

réseau de bipartitions [p. 71, 90, 91, 360]

Une bipartition est la séparation d'un ensemble en deux parties de manière à ce que tous les éléments de l'ensemble initial soient présents dans une et une seule de ces deux parties. Un réseau de bipartitions¹²⁷ est un **réseau** permettant de représenter un ensemble de bipartitions. Si toutes les bipartitions sont compatibles entre elles, le réseau de bipartitions produit est un **arbre**. Cependant toutes les bipartitions ne sont pas nécessairement compatibles et le réseau de bipartitions permet de visualiser les désaccords entre les différentes bipartitions.

Un réseau de bipartitions peut être obtenu à partir d'une **matrice** de caractères **binaires** (chaque **caractère** définissant alors une bipartition) mais aussi à partir d'une **matrice de distances**, notamment via la méthode du **réseau de voisins**.

réseau de voisins [p. 54, 67, 69, 74, 80, 254, 263, 352, 360]

L'algorithme du **réseau de voisins**¹²⁸, proposé par Bryant et Moulton [67]¹²⁹, permet de construire un **réseau de bipartitions** à partir d'une **matrice de distances**.

Comme la méthode du **plus proche voisin**¹³⁰, il utilise une méthode agglomérative regroupant progressivement les différents nœuds du réseau. À chaque étape, l'algorithme cherche les deux nœuds jugés les plus proches en minimisant un paramètre nommé \hat{Q} (étape de

126. ↑ Il constitue alors une **catégorie** musicale [21, p. 286].

127. ↑ *Splits network*, en anglais.

128. ↑ *NeighborNet*, en anglais.

129. ↑ Voir aussi Bryant et Moulton [68].

130. ↑ *neighbor joining*, en anglais.

sélection). Ces deux nœuds sont alors déclarés voisins et une nouvelle paire minimisant \hat{Q} est recherchée. S'il trouve un nœud qui a deux voisins, l'algorithme décide de transformer ces trois nœuds en deux nœuds¹³¹ (étape de *réduction*). Pour cela, la distance entre les deux nouveaux nœuds et chacun des autres nœuds est calculée par une formule basée sur les distances entre les trois anciens nœuds et les autres nœuds. Les nœuds sont ainsi progressivement regroupés jusqu'à ce qu'il ne reste plus que trois nœuds, constituant la topologie du réseau de bipartitions. Les longueurs des arêtes reliant les différents nœuds du réseaux sont finalement calculées (étape de *dimensionnement*) permettant d'obtenir le réseau final.

Cette méthode est fréquemment utilisée pour visualiser l'existence d'un signal **horizontal** dans un jeu de données et pour évaluer (via le calcul de la mesure δ) à quel point celui-ci se prête à une représentation **arborescente**.

responsorial [p. 114]

Voir *forme vocale*.

retention index [p. 344]

Voir *indice de rétention*.

réticulation [p. 81]

En **topologie**, une réticulation désigne une branche reliant plusieurs branches distinctes d'un **arbre**. Dès lors le graphe obtenu ne peut plus être considéré comme un arbre, et devient un **réseau**.

RI [p. 181, 182]

Voir *indice de rétention*.

robustesse [p. 59, 76, 180, 330]

Une méthode est dite *robuste* si les résultats qu'elle permet d'obtenir ne varient pas ou peu lorsque l'on modifie légèrement les données utilisées ou les paramètres de l'analyse.

rythme [p. 140, 189, 275]

En **systématique musicale**, le rythme désigne la manière dont des événements sonores, différenciés les uns des autres par une opposition de timbre, de durée ou d'accentuation, sont organisés. Elle se différencie ainsi de la **métrique**, qui constitue la trame temporelle sur laquelle se déploie le rythme :

« la métrique concerne *l'étalonnage du temps en quantités – ou valeurs – égales ; le rythme, les modalités selon lesquelles des durées différenciées sont regroupées.* »

Arom (2007) [17, p. 934]

131. ↑ Contrairement à l'algorithme du plus proche voisin qui condense à chaque étape deux nœuds en un seul.

Dans les musiques **cycliques**, le rythme formé par une **partie** pendant une **période** constitue une **formule rythmique**, généralement divisible en plusieurs éléments distinctifs et culturellement pertinents, les **cellules rythmiques**.

scalaire [p. 189, 275]

Relatif aux **échelles** musicales.

sciences historiographiques [p. 17, 249, 342]

Les sciences historiographiques sont l'ensemble des disciplines ayant pour objet l'étude scientifique du passé.

stemma [p. 18, 362]

En **critique textuelle** (et plus particulièrement en **stemmatologie**) un stemma est un arbre représentant les relations généalogiques entre différentes versions d'un même manuscrit.

stemmatique [p. 19, 20, 29, 53, 55, 59, 61, 76, 83, 349]

Le terme de stemmatique est utilisé comme nom synonyme de **stemmatologie**. Pour ma part, je ne l'utiliserai que comme adjectif se référant à la stemmatologie. Je l'utilise en général pour parler des méthodes stemmatiques, c'est-à-dire des méthodes développées par la stemmatologie, en particulier la **méthode de Lachmann** et les méthodes qui en sont dérivées.

stemmatologie [p. 19, 362]

J'utilise ici le terme de stemmatologie comme la branche de la **critique textuelle**, issue des travaux de Lachmann, qui cherche à inférer par une méthode objective les relations d'apparement entre différents manuscrits, c'est-à-dire à construire un **stemma**. Les méthodes stemmatiques sont celles utilisées en stemmatologie, et regroupent aussi bien la **méthode de Lachmann** que celles proposées par Paul Maas en 1958 [A272]¹³² ou par Ben Salemans en 1991 [366, pp. 22-35].

subdivision de la pulsation [p. 302, 330, 352]

Voir *modalité de subdivision de la pulsation*.

synapomorphie [p. 35, 57, 328]

Dans l'analyse **cladistique**, une synapomorphie est le partage de l'**état dérivé** d'un **caractère** (ou **apomorphie**) entre plusieurs groupes. Dans l'exemple de la figure 1.4 page 37, la présence de fleurs (état « présence » du caractère « Fleurs »), qui est l'état dérivé, est partagée par le brome dressé, le crocus d'automne et la bruyère callune : c'est donc une synapomorphie de ces trois espèces. Selon la méthode cladistique, seules les synapomorphies peuvent nous renseigner sur les relations de parentés entre les **taxons** (par opposition aux partages d'**états ancestraux**, ou **symplesiomorphies**).

132. ↑ Citée par Platnick et Don Cameron [339, p. 381].

synchronie [p. 133, 267, 283, 286, 337]

La *synchronie* s'intéresse à la comparaison à un même moment : on compare donc deux objets différents, mais décrits à un même moment. Ce terme, issu de la linguistique et proposé par Ferdinand de Saussure, s'oppose à celui de *diachronie*.

systématique biologique [p. 20, 22, 47, 57, 103, 145, 297, 306–308, 311, 328, 329, 334–339, 346, 349, 355, 364, 365, 367, 369]

La définition de la systématique biologique, en tant que discipline des sciences de la vie¹³³, est très variable selon les époques et les auteur·e·s [296, pp. 3–14] [396, pp. 10–13]. Elle est intimement liée à celle de *taxinomie*, et plus indirectement à celles de *classification* et de *système*. Les définitions peut-être les plus citées sont celles de George Simpson :

« Systematics is the scientific study of the kinds and diversity of organisms and of any and all relationships among them. »¹³⁴

« Zoological classification is the ordering of animals into groups (or sets) on the basis of their relationships, that is, of associations by contiguity, similarity, or both »¹³⁵

« Taxonomy is the theoretical study of classification, including its bases, principles, procedures, and rules. »¹³⁶ Simpson (1961) [A381, p. 7, p. 9, p. 11]

Pour Minelli [296, p. 3], la définition de *systématique* proposée par Simpson correspondrait plutôt à ce que l'on appelle aujourd'hui la *biologie comparée* et le sens du terme est aujourd'hui plus restreint. En effet, pour Guillaume Lecointre et Hervé Le Guyader la systématique est la « [s]cience des classifications biologiques et des relations évolutives entre organismes » alors que la taxinomie désigne la « [s]cience des lois de la classification des formes vivantes » [251, p. 542]. Kevin de Queiroz, suivant Graham Griffiths [176], propose cette séparation entre *classification* et *systematisation* :

« Classification is the ordering of entities (elements, individuals) into classes ; a class is a group defined by a property (attribute, character) or properties shared by its members. In contrast, systematization is the ordering of entities into systems ; a system is a more inclusive entity (whole) whose existence depends on some natural process through which its elements (component parts) are related. »¹³⁷ de Queiroz (1988) [342, p. 241]

133. ↑ Le terme de *systématique* est aussi utilisé en *ethnomusicologie*, voir l'entrée du glossaire concernant la *systématique musicale*.

134. ↑ « La systématique est l'étude scientifique des types et de la diversité des organismes et de toutes les relations entre eux. »

135. ↑ « La classification zoologique est l'organisation ordonnée des animaux dans des groupes (ou ensembles) sur la base de leurs relations, c'est-à-dire des associations par contiguïté, similarité ou les deux. »

136. ↑ « La taxinomie est l'étude théorique de la classification, en particulier ses bases, ses principes, ses procédures et ses règles. »

137. ↑ « La classification est l'organisation ordonnée d'entités (éléments, individus) dans des classes ; une classe est un groupe défini par une propriété (attribut, caractère) ou par des propriétés partagées par ses membres. Par

Lors de la définition de la systématique on voit que l'accent est mis sur le fait qu'elle repose sur un processus naturel¹³⁸, l'évolution et la filiation des êtres vivants, contrairement à la classification ou la taxinomie (termes parfois synonymes) qui peuvent reposer sur d'autres critères. Pour O'Hara [A324, p. 136]¹³⁹, la classification n'est plus depuis longtemps l'objectif de la systématique :

« The equation of systematics with classification is one of the ancient and petrified notions we have inherited from the pre-evolutionary period, and that we ought now to abandon. The problems of systematics now and in the future are not the problems of classification, they are the problems of historical reconstruction and historical representation : the reconstruction and representation of the evolutionary past. »¹⁴⁰ O'Hara (1992) [A324, p. 136]¹³⁹

Le mot est aussi utilisé pour désigner non plus la discipline, mais l'étude scientifique elle-même [251, p. 542], voire le résultat produit par les analyses de cette discipline.

systématique musicale [p. 107, 109, 111, 112, 114, 117, 128, 140, 250, 265, 301, 331, 332, 337, 354, 360, 361, 363, 365]

Comme l'expliquent Arom et Alvarez-Pereyre [20, p. 80] « [l]a systématique, c'est la grammaire d'une musique » :

« Par systématique, on entend les règles de fonctionnement d'une musique. Leur mise au jour doit prendre en compte aussi bien ses éléments constitutifs et les modalités de leur combinaison, que les structures formelles, de même que les procédés techniques (tels que monodie, hétérophonie, contrepoint) qui la caractérisent. » Arom et Alvarez-Pereyre (2007) [20, p. 31]

La systématique musicale désigne donc un ensemble de règles (souvent implicites) régissant la pratique musicale d'une culture. Par métonymie¹⁴¹, le terme est parfois utilisé pour désigner la discipline scientifique, appartenant à l'**ethnomusicologie**, qui cherche à comprendre

opposition, la systématization est l'organisation ordonnée d'entités dans des systèmes ; un système est une entité plus inclusive (le tout) dont l'existence dépend d'un processus naturel à travers lequel ses éléments (les parties) sont reliées. »

138. ↑ Mais le sens du mot *naturel* a pu varier selon les époques [296, pp. 11–12].

139. ↑ Citée par Minelli [296, p. 14].

140. ↑ « Le lien entre systématique et classification est une des vieilles notions pétrifiées héritées de la période pré-évolutionnaire et que nous devrions maintenant abandonner. Les problèmes de la systématique aujourd'hui comme demain ne sont pas ceux de la classification, ce sont ceux de la reconstruction historique et de la représentation historique : la reconstruction et la représentation du passé évolutif. »

141. ↑ Et peut-être par analogie avec d'autres usages du terme de *systématique*, comme la systématique biologique. On peut aussi y voir l'effet de l'usage conjoint du terme avec celui de *catégorisation musicale*. En effet la *systématique musicale* désigne plus volontiers un objet culturel alors que la *catégorisation musicale* se réfère le plus souvent à une activité scientifique. Aussi la formule *systématique et catégorisation musicales* peut correspondre soit à l'un (les systèmes de règles et d'organisation interne des éléments d'une culture musicale) soit à l'autre (l'activité scientifique visant à expliciter ces systèmes).

et expliciter ces règles¹⁴². Par analogie avec le terme de *catégorisation musicale*¹⁴³, je parlerai de *systématisation musicale* pour parler de l'activité consistant à créer un système de règles musicales.

systématisation musicale [p. 111, 365]

Activité consistant à élaborer une *systématique musicale*, c'est-à-dire un ensemble de règles régissant une pratique musicale. Je construis ce terme par analogie avec celui de *catégorisation musicale*. On pourra se référer à cette entrée pour plus de détails concernant la correspondance entre les différents termes que j'utilise, et en particulier le tableau 9.1 page 333.

En pratique ce terme désigne plus spécifiquement ici l'analyse et la formalisation scientifique d'une systématique musicale servant de modèle à une systématique musicale endogène.

système catégoriel [p. 107, 331–333, 350]

Ensemble cohérent de *catégories*, produit par l'activité de catégorisation. Cet ensemble est dit cohérent dans le sens où il forme un *système* : les différentes catégories ne sont pas indépendantes et organisent conjointement l'ensemble étudié (même si le système n'est pas nécessairement cohérent du point de vue de la logique formelle).

On pourra se référer à l'entrée *catégorisation musicale* pour plus de détails concernant la correspondance entre les différents termes que j'utilise et en particulier le tableau 9.1 page 333.

système de filiation [p. 64, 158, 160, 347, 348, 354]

En anthropologie, le système de filiation décrit la manière dont se transmet la parenté. Dans le cas d'une filiation *matrilinéaire*, les personnes sont rattachées à la lignée de leur mère. Dans le cas d'une filiation *patrilinéaire*, les personnes sont rattachées à la lignée de leur père.

système scalaire [p. 140]

Synonyme d'*échelle musicale* [15, p. 69] [20, p. 87].

taxinomie [p. 20, 363, 366]

La taxinomie (ou *taxonomie*) est une discipline des sciences de la vie. Sa pratique et sa définition étant aujourd'hui indissociables de la systématique biologique, la définition du terme *taxinomie* est traitée plus en détail dans l'entrée *systématique biologique*.

taxinomie numérique [p. 21, 53, 54, 355]

Le terme de *numerical taxonomy*, tel que défini à l'origine par Sokal et Sneath [A389, p. 48],

142. ↑ Par exemple dans cette phrase de Fürniss [151, p. 40] : « À l'instar de l'archéologie musicale et de l'iconographie, qui abordent l'évolution formelle des instruments de musique, la systématique et la catégorisation musicales sont des outils pour aborder l'évolution des langages musicaux. »

143. ↑ Voir cette entrée pour plus de détail concernant la correspondance entre les différents termes que j'utilise, et en particulier le tableau 9.1 page 333.

ne se restreint pas à la **phénétique** et met en avant l'utilisation de données quantitatives et de méthodes mathématiques et statistiques dans un but **taxinomique** :

« We mean by [numerical taxonomy] *the numerical evaluation of the affinity or similarity between taxonomic units and the ordering of these units into taxa on the basis of their affinities*. The term may include the drawing of phylogenetic inferences from the data by statistical or other mathematical methods to the extent to which this should prove possible. »¹⁴⁴

Sokal et Sneath (1963) [A389, p. 48]

La controverse entre méthodes phénétique et cladistique (Sokal et Sneath défendant la phénétique) a participé à assimiler *numerical taxonomy* et *phenetics*, les deux termes étant aujourd'hui souvent employés indifféremment pour désigner la méthode phénétique. Lemey et coll. [258, p. 23] préfèrent parler de méthodes basées sur des **matrices de distances**.

taxon [p. 26, 32, 34, 62, 77, 90, 268, 328, 330, 336, 339, 342, 344, 346, 348, 357, 362, 367]

D'après Simpson [A381]¹⁴⁵ « [u]n taxon est un groupe d'organismes reconnu en tant qu'unité formelle à chacun des niveaux de la classification ». Ainsi l'espèce *Crocus nudiflorus* aussi bien que le groupe des plantes à fleurs (les Angiospermes) sont des taxons.

technique vocale [p. 73, 353]

En **ethnomusicologie**, la technique vocale est un **paramètre** permettant de caractériser la manière d'utiliser le système phonatoire dans le cadre d'une production vocale. Le **jodel**, par exemple, est une technique vocale.

ternaire (subdivision de la pulsation) [p. 291, 293, 302, 350, 360]

Voir *modalité de subdivision de la pulsation*.

test de Kishino–Hasegawa [p. 261, 262]

Ce test statistique, proposé par Kishino et Hasegawa [A224] [216, p. 3882], permet de tester la similarité entre deux arbres A et B. Pour cela il construit un arbre à partir du jeu de données ayant servi à produire l'arbre A par une méthode cladistique, mais en faisant une recherche heuristique contrainte par l'arbre B. Cet arbre est ensuite comparé à l'arbre A (non contraint) : si ces deux arbres ne sont pas significativement différents, il y a une coévolution parfaite entre les données de l'arbre A et celles de l'arbre B.

test de longueur par permutation [p. 67, 81, 355, 358]

Ce test statistique¹⁴⁶, proposé par Archie [A8]¹⁴⁷, est utilisé pour évaluer la présence d'un signal **phylogénétique** dans un jeu de données.

144. ↑ « Nous entendons par taxonomie numérique *l'évaluation numérique de l'affinité ou de la similarité entre les unités taxonomiques*. Le terme peut inclure le tracé des inférences phylogénétiques à partir des données par les statistiques ou par d'autres méthodes mathématiques dans la mesure où cela serait possible. »

145. ↑ Citée par Darlu et Tassy [107, p. 26].

146. ↑ Appelé *permutation tail probability* (ou PTP) en anglais.

147. ↑ Citée par Tehrani et Collard [408, p. 449].

Les états de caractères décrits dans la matrice de données sont redistribués aléatoirement un grand nombre de fois¹⁴⁸ et une analyse cladistique est effectuée à partir de cet arbre. On regarde alors la distribution de la longueur des arbres obtenus. Si l'arbre issu du jeu de données originel est plus court que 95% des arbres obtenus après permutation, le jeu de données contient vraisemblablement un signal phylogénétique [408, p. 449].

test de Mantel [p. 79, 81, 254, 260, 367]

Le test de Mantel est un test statistique permettant d'évaluer l'existence d'une corrélation entre deux matrices de tailles identiques et représentant typiquement deux mesures de distances différentes pour un même ensemble d'objets. Il est par exemple fréquemment utilisé en génétique des populations pour tester la corrélation entre la distance géographique et la distance génétique au sein d'une même population.

Le test de Mantel est décrit plus en détail dans l'annexe B.3 page 421.

test partiel de Mantel [p. 254]

Le test partiel de Mantel est un dérivé du test de Mantel introduit par Smouse et coll. [385] en 1986. Il permet d'évaluer la capacité relative de deux distances X_1 et X_2 à expliquer une troisième distance observée Y . Elle permet donc de tester la significativité de la corrélation entre Y et X_1 (respectivement X_2) en retirant les effets de corrélation dus à X_2 (respectivement X_1).

topologie [p. 39, 65, 68, 76, 80, 180, 335, 361]

La topologie désigne la manière dont les nœuds et arêtes d'un arbre ou d'un réseau sont disposés les uns par rapport aux autres. Elle décrit donc uniquement leur organisation relative, et ne s'intéresse pas aux aspects quantitatifs comme les longueurs des arêtes.

tradition (critique textuelle) [p. 60, 61, 70, 82, 86]

En critique textuelle, une *tradition* désigne un ensemble de manuscrits apparentés ainsi que les relations entre ceux-ci. Si cette tradition a été créée artificiellement, notamment à des fins scientifiques, on parle de *tradition artificielle*.

trait [p. 330]

Quasi-synonyme du terme *caractère*, bien que plus souvent utilisé pour parler de caractères phénotypiques.

transfert horizontal [p. 57, 69, 79, 335–337]

En systématique biologique un transfert horizontal est l'acquisition par un taxon d'un état de caractère provenant d'un autre taxon. Cette acquisition se fait généralement par échange de matériel génétique et l'on parle le plus souvent de *transfert horizontal de gène* (THG).

On peut rapprocher cette notion de celle d'*emprunt* en linguistique historique et de *contamination* en critique textuelle. Voir aussi la notion de *transmission horizontale*.

148. ↑ Contrairement à la technique de *bootstrap*, c'est un tirage sans remise [408, p. 449] : chaque caractère est donc tiré une et une seule fois.

transmission en chaîne [p. 60, 87, 340]

Les expériences de transmission en chaîne permettent d'effectuer une *évolution artificielle* par simulation analogique. Des participant-e-s volontaires miment la transmission d'objets culturels : chaque personne reproduit l'objet lui-même reproduit par la personne précédente. Ces expériences permettent d'obtenir un modèle contrôlé de l'évolution culturelle qui peut être étudié de manière précise et sans lacune. De telles expériences ont par exemple été effectuées en *critique textuelle* [34, 361, 392] et en anthropologie de la culture matérielle [373]¹⁴⁹.

transmission horizontale [p. 76, 79, 80, 263, 305, 343, 360, 367]

En biologie, la transmission horizontale désigne la transmission de gènes ou de parasites entre espèces différentes et qui ne peuvent donc¹⁵⁰ échanger de matériel génétique. Le concept a été prolongé concernant l'étude de l'évolution culturelle. Il peut alors s'agir du processus de transmission d'éléments culturels entre lignées biologiques différentes. Les flux génétiques n'étant cependant pas limités au sein de l'espèce humaine, la transmission des gènes n'est pas à proprement parler verticale puisque différentes parties du génome peuvent porter différentes histoires.

Il peut aussi s'agir de la transmission horizontale d'éléments culturels par rapport à la tendance dominante de la transmission culturelle. En effet, tous les éléments culturels ne sont pas forcément transmis de la même manière et peuvent donc avoir des histoires différentes¹⁵¹.

De manière générale le concept de transmission horizontale ne peut être compris que par rapport à une transmission définie comme verticale pour les besoins de l'étude. La précision de la verticalité de référence est cependant omise la plupart du temps, rendant cette notion souvent floue. La transmission verticale sous-entendue peut ainsi se référer, par exemple, à la transmission de matériel génétique, à la transmission de parent à enfant ou encore à la transmission linguistique.

uniformitarisme [p. 23]

L'uniformitarisme est un postulat selon lequel les mécanismes de transformation intervenant dans le passé sont les mêmes que ceux s'exerçant aujourd'hui. Il a notamment été théorisé et appliqué en géologie par Charles Lyell, qui l'a aussi introduit en *linguistique historique* [77, p. 213]. Il a alors été utilisé par l'école néogrammairienne pour justifier l'utilisation des lois phonétiques à des fins d'*inférence historique*¹⁵².

149. ↑ Le laboratoire junior LaRÉMI avait commencé un projet, hélas avorté, d'une telle expérience concernant la transmission de la musique. Voir <http://laremi.net/activite/experiences/cartra/>.

150. ↑ Selon le concept biologique de l'espèce, tel que proposé par Mayr.

151. ↑ Cet aspect est expliqué en détail dans la partie 7.2 et en particulier dans les parties 7.2.2 page 260 et 7.2.3 page 262.

152. ↑ Voir partie 1.2.3 page 22.

UPGMA [p. 62, 69, 91]

Cet algorithme, proposé par Sokal et Michener [A388]¹⁵³, permet de construire un **arbre phylogénétique** à partir d'une **matrice de distances**.

Il cherche les deux taxons entre lesquels la distance donnée dans la matrice est la plus faible, puis les regroupe en un seul taxon dont la distance avec les taxons restants est calculée par la moyenne arithmétique. L'opération est alors recommencée jusqu'à ce que tous les taxons aient été regroupés et organisés en un **arbre**. C'est donc une **méthode de distances** agglomérative, tout comme la méthode du **plus proche voisin**.

UPGMA est le sigle pour le terme anglais *unweighted pair group method with arithmetic mean* (méthode de groupement non pondérée par paires à partir de la moyenne arithmétique).

V.O.M. [p. 293]

Sigle pour *valeur opérationnelle minimale*.

valeur (de paramètre) [p. 114, 275, 304, 305, 331, 339, 352]

En *ethnomusicologie*, la valeur d'un **paramètre** est la forme observée que prend ce paramètre dans un cas particulier. Par exemple le paramètre « subdivision de la pulsation » prend la valeur « ternaire » dans le répertoire *Lisimbu* chez les Awandji de Doumé et prend la valeur « binaire » dans le répertoire *Lisimbu* chez les Ngoms de Bolapessa.

Cette notion peut être rapprochée de celle d'*état de caractère* en *systématique biologique*.

valeur opérationnelle minimale [p. 292, 293, 349, 350, 358, 369]

D'après Simha Arom :

« **valeur opérationnelle minimale** : équivaut à la plus petite durée *pertinente* issue de la subdivision du mètre [...] en deux, trois, quatre, six – voire cinq – valeurs brèves, dont toutes les autres durées constituent des multiples. »

Arom (2007) [16, p. 416]

Ainsi la valeur opérationnelle minimale (V.O.M.) est la plus petite durée pertinente par laquelle peut-être divisée la **pulsation**¹⁵⁴. Le nombre de valeurs opérationnelles minimales dans chaque pulsation définit la **modalité de subdivision de la pulsation**.

variante [p. 29, 56, 57, 86, 335, 338]

En *critique textuelle*, une variante désigne chacun des textes trouvés en un même **lieu variant** dans les différents manuscrits comparés. On trouve aussi le terme d'*erreur*.

version (de pièce) [p. 357, 370]

En *ethnomusicologie*, une version de pièce désigne chacune des réalisations particulières d'une **pièce** musicale.

153. ↑ Citée par Felsenstein [136, p. 162].

154. ↑ Dans leur glossaire, Arom [16, p. 413] renvoient la notion de *pulsation* à celle de *mètre*.

version minimale [p. 111, 112, 120]

En **ethnomusicologie**, le terme de *version minimale* désigne une **version** de **pièce** particulière, volontairement réduite au strict minimum par les interprètes. Cette version minimale peut par exemple être réalisée à la demande de l'ethnomusicologue, ou encore dans un contexte d'apprentissage.

vraisemblance [p. 166]

Dans les **méthodes probabilistes**, la vraisemblance¹⁵⁵ désigne la probabilité d'obtenir, sous un modèle donné, un jeu de données de référence [116, p. 59] (le plus souvent un ensemble de données réellement observées).

Ainsi, dans la méthode présentée et appliquée dans le chapitre 5¹⁵⁶, la vraisemblance désigne la probabilité (d'après le modèle de déplacement des populations utilisé pour la simulation) que les populations soient à l'emplacement actuellement observé. Cette probabilité dépend des positions des ancêtres supposés des populations étudiées. Les méthodes de **maximum de vraisemblance** cherchent ainsi l'ensemble de ces positions ancestrales pour lesquelles la probabilité (d'après le modèle utilisé) d'observer les populations actuelles à l'emplacement réellement observé est maximale.

155. ↑ La vraisemblance est souvent notée *Lh*, pour l'anglais *likelihood*.

156. ↑ Pages 155–185.

Annexes

A Conventions	373
A.1 Citations et traductions	373
A.2 Sources et références bibliographiques	374
B Suppléments	377
B.1 Chapitre 3 : Citations autour de la notion de <i>modèle</i>	378
B.2 Chapitre 5 : Inférence phylogéographique	387
B.3 Test de Mantel	421
C Bibliographie	425
C.1 Publications consultées	425
C.2 Autres publications citées	445
D Listes des figures et des tableaux	457
D.1 Liste des figures	458
D.2 Liste des tableaux	460
E Table exhaustive des matières	461

Annexe A

Conventions

A.1 Citations et traductions

Sauf précision, les formatages présents dans les textes intégrés en citation sont ceux introduits par la ou les personnes ayant rédigé le texte. Ainsi dans la citation « Le tempo traduit le *mouvement intérieur* de la musique à laquelle il est associé. » [12, p. 10], c'est bien Arom qui souligne le terme de « mouvement intérieur » puisqu'il n'y a pas de précision à ce sujet. Les textes cités rédigés par l'auteur dans une langue autre que le français sont toujours donnés dans leur version originale dans le corps du texte, et traduits en note de bas de page. Sauf précision contraire, les traductions sont faites par mes soins.

A.2 Sources et références bibliographiques

J'indique tout au long de cette thèse, comme il se doit dans tout travail de construction d'un savoir collectif, les références vers les travaux d'où je tire les connaissances et dont j'essaie de restituer au mieux le sens initial. Je présente ci-après quelques conventions concernant l'utilisation des références dans cette thèse. Le tableau A.1 résume les différentes notations utilisées et leurs significations.

J'ai essayé tant que possible de me référer directement aux sources originelles des informations données, mais la science qui est sans doute encore toute jeune est déjà bien vaste. Aussi il arrive fréquemment que je fasse référence à un ouvrage sans l'avoir consulté directement mais dont j'ai eu une description du contenu à travers une autre source (souvent un état de l'art). Je préfère ainsi citer directement le travail initial à l'origine du résultat, afin que les personnes intéressées puissent s'y référer directement. Mais, afin de valoriser le travail de médiation scientifique et par souci de précision concernant les références que j'utilise, j'indique la source intermédiaire (appelée *source médiatrice*) en note de bas de page (cas 2). Parfois la source indirecte ne donne pas elle-même la référence vers le travail originel, mais renvoie vers une autre source intermédiaire (*source médiatrice indirecte*). J'indique alors les deux références en notes de bas de page, en explicitant celle que j'ai consultée, et celle qui est la source médiatrice indirecte (cas 4 et 5). Il arrive aussi que je ne cite pas la source médiatrice indirecte (cas 3). Afin de visibiliser les sources que je n'ai pas consultées directement, les renvois vers les sources indirectes sont précédés de la lettre A¹. Ces conventions permettent deux conclusions qui, je l'espère, faciliteront la lisibilité de mes sources :

- Une information sourcée dans le texte [X], *sans note de bas de page*, justifie directement l'information et a été consultée directement (cas 1).
- La référence suivant directement et *dans le texte* une information est celle justifiant l'information².

1. ↑ Voir tableau A.1 et figure A.1 page 375 et les explications au début de la bibliographie page 425.

2. ↑ Ou censée le faire : en effet, dans le cas où je passe par une ou plusieurs sources médiatrices, je ne peux être certain que la source primaire justifie bien l'information, sinon en faisant confiance aux sources médiatrices.

Cas	Texte du manuscrit	Source consultée	Source(s) médiatrice(s)	Source de « INFORMATION »
1	INFORMATION [X]	[X]	aucune	[X]
2	INFORMATION [Y] ¹	[X]	[X]	[Y]
3	INFORMATION ²	[X]	non donnée	non disponible
4	INFORMATION ³	[X]	[X], [Z]	non disponible
5	INFORMATION [Y] ⁴	[X]	[X], [Z]	[Y]

¹ Citée par [X].

² [X]

³ [Z], citée par [X].

⁴ Citée par [Z], elle-même citée par [X].

TAB. A.1 : **Conventions utilisées dans ce travail pour la notation des références.** [X], [Y] et [Z] sont des références à la bibliographie page 425. Lorsque la source est consultée directement, le numéro n de renvoi est indiqué sous la forme $[n]$ et réfère à la partie C.1 page 425 de la bibliographie. Lorsque la source n'est pas consultée directement, le numéro n de renvoi est indiqué sous la forme $[An]$ et réfère à la partie C.2 page 445 de la bibliographie. Des exemples d'application de ces conventions sont données dans la figure A.1 page 375.

1. « L'intérêt des universitaires se portera par ailleurs vers d'autres approches, notamment synchroniques [353, p. 100]. »
2. « Les méthodes de distance seront ainsi popularisées par les travaux de Fitch et Margoliash [A142]¹ qui introduisent en 1967 le *critère des moindres carrés* [...] »
3. « [Scaliger] pose les premières bases de la méthode comparative² [...] »
4. « Schleicher publie en 1853 deux travaux incluant des arbres de langues³ [...] »
5. « Lachmann [A231]⁴, durant la première moitié du XIX^{ème} siècle, met en avant la nécessaire rigueur méthodologique [...] »

¹ Citée par Felsenstein [136, pp. 131–132].

² Atkinson et Gray [27, p. 515]

³ Koerner [A227], citée par Atkinson et Gray [27, p. 517].

⁴ Citée par Timpanaro [A415, pp. 85–88], elle-même citée par Robins [353, p. 91].

FIG. A.1 : **Exemples d'application des conventions utilisées pour la notation des références.** Les cinq exemples présentés correspondent aux cinq cas du tableau A.1 page 375.

Annexe B

Suppléments

B.1 Chapitre 3 : Citations autour de la notion de *modèle*

Les citations présentées dans les pages suivantes ont toutes un lien avec le terme de *modèle* ou avec l'une des concepts qui en sont issus. Elles me servent à discuter ces différentes notion dans la partie 3.3 pages 118 à 128. Le numéro qui identifie chacune d'entre elles (commençant par *C*, par exemple *C23*) permet de me référer précisément à ces citations lors de la discussion. Le tableau ci-dessous présente un index des identifiant, des références et des pages de cette thèse correspondant à chaque citation.

Citation	Référence	Page
C1	Arom (1985) [13, p. 18]	379
C2	Arom (1985) [13, p. 226]	379
C3	Arom (1985) [13, p. 255]	379
C4	Arom (1985) [13, p. 255]	379
C5	Arom (1985) [13, pp. 272–273]	379
C6	Arom (1985) [13, p. 273]	380
C7	Arom (1987) [14, p. 119]	380
C8	Arom (1991) [15, p. 70]	380
C9	Arom (1991) [15, p. 72]	381
C10	Arom et Alvarez-Pereyre (2007) [20, p. 65]	381
C11	Arom et coll. (2008) [21, p. 296]	381
C12	Bornes Varol (2011) [51, p. 61]	381
C13	Ferran (2010) [138, pp. 42–43]	381
C14	Ferran (2010) [138, p. 214]	382
C15	Ferran (2010) [138, p. 215]	382
C16	Ferran (2010) [138, p. 483]	382
C17	Ferran (2010) [138, p. 483]	382
C18	Ferran (2010) [138, p. 484]	383
C19	Fürniss (1993) [150, p. 104]	383
C20	Fürniss-Yacoubi (2007) [154, p. 38]	383
C21	Fürniss-Yacoubi (2007) [154, p. 38]	383
C22	Le Bomin (2000) [234, p. 137]	383
C23	Le Bomin (2000) [234, p. 138]	384
C24	Le Bomin (2000) [234, p. 141]	384
C25	Le Bomin (2000) [234, p. 278]	384
C26	Le Bomin (2000) [234, p. 354]	384
C27	Le Bomin (2004) [237, p. 179]	385
C28	Le Bomin (2004) [237, p. 182]	385
C29	Le Bomin (2004) [237, p. 182]	385
C30	Le Bomin (2004) [237, p. 187]	385
C31	Le Bomin (2004) [237, p. 188]	385
C32	Nattiez (1991) [303, p. 79].	385
C33	Mifune (2012) [295, p. 503]	386

TAB. B.1 : **Index des citations autour de la notion de modèle.** Ces citations sont discutées en détail dans la partie 3.3 pages 118 à 128.

C1

« Toute musique ethnique — indépendamment de sa fonction sociale — relève d'une systématique. Comme une langue, elle est dotée d'une grammaire et, à ce titre, sanctionnée par des règles que sous-tend une théorie, même si celle-ci est presque toujours *implicite*. »

Arom (1985) [13, p. 18]

C2

« Brăiloiu met ainsi le doigt sur un point capital : la démarche du chercheur consiste à retrouver des propriétés essentielles, c'est-à-dire un *modèle*. Que celui-ci soit ici désigné du terme d'“archétype” indique bien qu'au cours de chaque exécution, les variations qu'il subit ne peuvent s'appliquer simultanément à *tous* ses éléments. L'improvisation, en conséquence, connaît des limites. S'il était en effet possible, à partir d'un modèle, d'improviser à l'infini, la pièce musicale qui l'incarne deviendrait méconnaissable : le modèle est précisément ce qui préserve l'identité de la pièce. »

Arom (1985) [13, p. 226]

C3

« Il est clair, par conséquent, que l'existence d'un code implique celle d'une théorie ; à cet égard, nous avons précisé à plusieurs reprises que dans le cas des musiques traditionnelles africaines, la théorie n'était pas explicite, mais *implicite*. »

Arom (1985) [13, p. 255]

C4

« À un niveau inférieur, la description doit être capable, à l'intérieur de chacune des parties d'une pièce donnée, de caractériser chaque unité opératoire, c'est-à-dire chaque élément minimal de pertinence formelle.

Cette étape du travail de la description, qui procède du tout aux parties, est strictement analytique. Au cours de cette phase, le chercheur tente, pour chacune des pièces de l'ensemble étudié, d'aboutir au *modèle* qui sous-tend chaque partie et, par corollaire, au modèle global de simultanéité qui, de la cohérence de chacune de ses parties, mène à celle de la pièce polyphonique tout entière.

La seconde étape du travail de la description consiste alors à reconstruire, à engendrer, à partir des modèles obtenus, une simulation, un artefact de chacune des pièces de l'ensemble considéré. »

Arom (1985) [13, p. 255]

C5

« Une fois l'analyse faite, il convient d'en valider les résultats. Cette opération exige que soit dégagée le modèle de l'objet considéré, c'est-à-dire le schéma qui sous-tend chacune de ses réalisations, l'épure à laquelle sa substance peut-être ramenée et qui, cependant, contient l'ensemble de ses traits pertinents.

Ce *modèle* — aussi schématique qu'il puisse être — doit encore laisser paraître la singularité de l'objet. Si le modèle répond à cette exigence, il se trouve validé.

Mais il faut alors, afin de vérifier la validité des analyses elles-mêmes, qu'il soit possible, à partir du modèle, de reproduire de nouveaux objets, autrement dit, d'en élaborer de nouvelles réalisations qui, d'un point de vue culturel, soient identiques à l'objet initial.

[...] À la phrase déductive succèdera donc une phase inductive, puisque, après être passé de l'objet à son modèle, on doit maintenant procéder en sens inverse : réaliser l'objet à partir du modèle, afin de vérifier que ce dernier a été correctement déduit. »

Arom (1985) [13, pp. 272–273]

C6

« le modèle constitue une représentation à la fois globale et simplifiée de l'objet. Il le caractérise, tout en condensant, en une représentation qui en accuse la singularité, l'ensemble de ses traits pertinents, et eux seulement.

[...] lorsque toutes les pièces d'un répertoire ont pu être modélisées, lorsque, également, les principes qui déterminent l'engendrement des réalisations de chaque pièce à partir de son modèle ont pu être dégagés, la description détaillée de chacune d'elles devient superflue. »

Arom (1985) [13, p. 273]

C7

« Le modèle est une représentation à la fois globale et simplifiée de l'objet ; il le caractérise tout en condensant, en une représentation qui en accuse la singularité, l'ensemble de ses traits pertinents, et eux seulement.

Dans les musiques centrafricaines le modèle équivaut à la réalisation minimale — c'est-à-dire la plus dépouillée, la référence ultime — d'une entité musicale qui puisse être identifiée comme telle par les détenteurs de la tradition à laquelle elle appartient.

Dans les musiques de tradition orale, la modélisation peut être obtenue, soit par l'émergence — c'est-à-dire la matérialisation — du modèle, soit par une série de déductions successives. »

Arom (1987) [14, p. 119]

C8

« Dans la perspective ainsi définie, le modèle est l'énoncé minimal, la référence ultime de chaque entité musicale — fut-elle strictement rythmique, mélodico-rythmique ou polyphonique — qui est à l'origine de toutes les réalisations qui en sont culturellement admises. Il constitue la réalisation — attestée ou virtuelle — minimale, la plus épurée, d'une entité musicale et à ce titre, contient à l'état implicite, comme en filigrane, toutes ses virtualités.

Par *modèle*, il faut entendre ici une *représentation sonore à la fois globale et simplifiée d'une entité musicale* ; le modèle en accuse la singularité *en ce qu'il condense, en une épure, l'ensemble de ses traits pertinents et eux seuls*.

Le modèle équivaut à la réalisation la plus dépouillée d'un morceau, qui puisse être identifié comme tel par les détenteurs de la tradition à laquelle le morceau considéré appartient. »

Arom (1991) [15, p. 70]

C9

« Non seulement nous sommes bien en présence d'un *modèle* — c'est-à-dire du principe fondamental d'organisation sur lequel s'appuient toutes les réalisations de chacune des parties constitutives d'une pièce donnée — mais, qui plus est, ce modèle est clairement conceptualisé par les tenants de la tradition. » Arom (1991) [15, p. 72]

C10

« [Le modèle constitue la] référence mentale sur laquelle est fondée l'exécution de chaque pièce [...] [la] référence structurelle, commune à toutes les réalisations d'un même morceau, par-delà l'ensemble des variations qu'il admet. » Arom et Alvarez-Pereyre (2007) [20, p. 65]¹

C11

« Qu'est ce qui conduit alors les détenteurs de la tradition à désigner ces pièces par un même nom ? C'est qu'elles procèdent d'une seule et même entité qui se trouve être réalisée de manières différentes. Par entité, il faut comprendre le substrat qui est commun à plusieurs pièces et qui se situe en amont de leur matérialisation. Ce substrat contient les caractéristiques concernant la métrique, la rythmique, l'ordre de succession des degrés et, dans le cas de la plurivocalité, les points d'ancrage et la nature des consonances. L'entité qui constitue le dénominateur commun à toutes les pièces n'est jamais réalisée en tant que telle. Elle correspond en quelque sorte à un schéma qui peut être matérialisé de différentes manières. »

Arom et coll. (2008) [21, p. 296]

C12

« Au cours de cette mise au point, pour le groupe de travail, S. Fűrniß a défini la matrice comme ce qui est transmis lors de l'apprentissage et l'épuration comme le résultat d'un processus de simplification. Elle est le minimum nécessaire pour la reconnaissance d'une pièce ou d'une partie polyphonique ou polyrythmique, dépouillée des ornements. Le modèle, extrait par les ethnomusicologues des épures des pièces représentées, est ce qu'il y a de plus profond au niveau fonctionnel, ce dont on ne peut plus rien retirer. Le modèle ainsi défini se distingue de l'"archétype" en ce que ce dernier prétend être un modèle ancien, archaïque, originel. Nous avons utilisé le terme de "prototype" dans le sens de meilleur exemplaire d'une catégorie considérée, le terme "stéréotype" n'a été employé dans nos études que pour signifier une typification réductrice et figée. »

Bornes Varol (2011) [51, p. 61]

C13

« Contrairement à l'épuration qui est mélodique, le substrat est polyphonique et contient en son sein la totalité des réalisations potentielles de la partie. La pièce maale est donc sous-tendue par un modèle polyphonique, qui ne résulte pas de l'imbrication de différentes épures mélodiques, mais de la superposition des substrats polyphoniques sous-jacents à chacune de ses parties constitutives. »

Ferran (2010) [138, pp. 42-43]

C14

« Tout apprenti musicien se crée ainsi un “stock” de variantes individuelles dans lequel il puisera librement lorsqu’il sera amené à exécuter une pièce. Ces variantes individuelles procèdent toutes d’un même substrat, partagé par l’ensemble des habitants d’une même région. Un substrat polyphonique sert donc de modèle à chacune des parties constitutives des pièces maale.

[...] En termes de démarche, on est passé de la recherche d’un modèle minimaliste, l’épure, à la recherche d’un modèle maximaliste, le substrat. L’épure constitue la version minimale de la partie. Dépourvue de toute variante, on ne peut rien lui retrancher sous peine d’en altérer l’identité. Elle apparaît en filigrane dans la totalité des variantes de la partie. Le substrat, quant à lui, contient en son sein la totalité des réalisations potentielles de la partie. Par la représentation graphique des contraintes et libertés auxquelles les exécutants sont soumis pour exécuter la partie, le substrat permet de visualiser en un seul schéma l’ensemble des réalisations possibles de celle-ci. » Ferran (2010) [138, p. 214]

C15

« On peut se demander quel statut accorder au substrat. C’est avant tout une représentation graphique qui condense, en un seul schéma, l’ensemble des réalisations possibles d’une partie. Mais dans la mesure où ce substrat est obtenu sur la base du jugement d’équivalence des exécutants et qu’il a fait l’objet d’une validation par les membres de la communauté étudiée, il possède un fondement cognitif (Molino 1995 : 201). Comme le souligne Jean Molino (op. cit. : 202), au sujet du modèle reconstruit par Simha Arom, il est clair que nous sommes loin de savoir comment ce substrat est inscrit dans l’esprit des musiciens maale, quelle espèce d’objet mental (Changeux 1983) il constitue. Il n’en reste pas moins que ce substrat relève des capacités cognitives des musiciens étudiés. Il constitue un type particulier de modèle mental (Johnson-Laird 1983), auquel les musiciens se réfèrent pour guider leurs actions musicales, et notamment l’exécution et la reconnaissance de la partie. Le substrat procure une certaine liberté aux exécutants et aux auditeurs tout en contraignant leurs actes. Tant que les interprètes ne transgressent pas les limites imposées par le substrat, l’exécution est correcte et reconnaissable. Au-delà, c’est l’erreur, aussitôt signalée d’une manière ou d’une autre par les musiciens – il est rare, en effet, que la faute passe inaperçue. Violer les règles, c’est enfin prendre le risque de dénaturer la partie et de la rendre méconnaissable. » Ferran (2010) [138, p. 215]

C16

« **Canevas** : modèle d’une pièce pluripartite. Le canevas résulte de l’imbrication des substrats sous-jacents à chacune des parties constitutives d’une pièce composée d’au moins deux parties. » Ferran (2010) [138, p. 483]

C17

« **Entité** : succession cyclique de points d’ancrages sur lesquels se trouvent un ou plusieurs degrés pouvant faire l’objet de commutations. Une entité peut être réalisée de différentes manières. Parmi les multiples réalisations possibles d’une entité, seules certaines sont utilisées par la communauté étudiée : il convient donc de distinguer les modes de réalisation potentiels et fonctionnels. » Ferran (2010) [138, p. 483]

C18

« **Substrat** : le substrat est un modèle, une référence mentale, de type maximaliste. Il convient de distinguer deux types de substrat :

- substrat de l'entité : succession cyclique de points d'ancrage, sur lesquels se trouvent un ou plusieurs degrés pouvant faire l'objet de commutations.
- substrat d'une partie constitutive : succession cyclique de points d'ancrages, dont certains sont obligatoires et d'autres facultatifs, et sur lesquels se trouvent un ou plusieurs degrés pouvant faire l'objet de commutations. »

Ferran (2010) [138, p. 484]

C19

« Chacune des quatre parties constitutives possède pour chaque chant une mélodie propre qui en est la référence pour toutes ses réalisations. Cette référence, le patron en quelque sorte, est la version la plus épurée et non-variée de cette partie. Il est la base des réalisations [...] »

Fürniss (1993) [150, p. 104]

C20

« C'est ce qui constitue le substrat, l'essence du chant polyphonique en question. On est ici à un niveau d'abstraction hors conceptualisation verbalisée par les Aka. Le substrat est ce qui est en amont de toute réalisation d'un chant. Il inclut chacun des modèles mélodico-rythmiques des parties constitutives de la polyphonie. »

Fürniss-Yacoubi (2007) [154, p. 38]

C21

« La notion de substrat [...] a ensuite [...] été associée à celle, plus large, d'entité". Ce terme désigne le dénominateur commun à toutes les réalisations variées d'un même matériel musical de base. Elle est devenue un outil important pour la reconnaissance, par l'ethnomusicologue, de l'identité de pièces musicales portant un même nom, mais étant réalisées par des effectifs musicaux et avec des techniques compositionnelles variables. L'entité ou, dans le cas de musiques polyphoniques, le substrat, permet de comprendre un phénomène récurrent en Afrique centrale, à savoir que des pièces portant le même nom sont associées à des catégories musicales différentes. »

Fürniss-Yacoubi (2007) [154, p. 38]

C22

« Comme dans la majeure partie des musiques traditionnelles où une certaine liberté est offerte aux interprètes, chaque réalisation d'une même pièce donne à entendre quelque chose de nouveau. Cependant, que ces versions soient écoutées par les gens qui la pratiquent habituellement ou des oreilles exercées, elles sont toujours reconnaissable d'une version à l'autre. Il existe ainsi une proportion d'éléments variables et d'autres stables qui permettent de maintenir l'identité de la pièce. En donnant accès à la version minimale d'une pièce ou d'une de ses parties constitutives, la modélisation permet de dégager la référence à partir de laquelle les musiciens effectuent des variations. »

Le Bomin (2000) [234, p. 137]

C23

« En effet, l’acquisition d’une technique instrumentale nécessite l’acquisition de plusieurs modalités de réalisation, chacune d’elles permettant de matérialiser un état d’une pièce ou d’une de ces [*sic*] parties à un moment donné de ses possibilités de réalisation. C’est-à-dire qu’à chaque modalité de réalisation correspond une formule minimale caractéristique de chacune des étapes. Ces modalités étant toujours d’une difficulté croissante, [...] la première d’entre elle [...] est l’épure de la pièce s’il s’agit d’une musique pour soliste ou d’une de ses parties constitutives pour une musique d’ensemble.

La modélisation demande aux musiciens de restituer leur parcours d’apprentissage. La reconstruction des pièces va alors de soi dans le va et vient perpétuel qu’effectuent les musiciens entre les versions conventionnelles dont ils doivent matérialiser les versions minimales de chaque partie, et les versions minimales à partir desquelles ils reconstruisent une version conventionnelle étape par étape. Apparaissent ainsi différentes références, collectives ou individuelles. »

Le Bomin (2000) [234, p. 138]

C24

« Le jeu sur le xylophone soliste s’effectue à l’aide de trois modalités de réalisation acquises lors de l’apprentissage et qui constituent des “jeux de base” : il s’agit de trois manières de jouer du xylophone soliste, chacune d’elles correspondant à des contraintes de réalisation [...] spécifiques.

[...] L’utilisation *stricte* de ces modalités engendre des *structures sonore strictement répétitives*, sous forme de versions minimales pour chacune des modalités de jeu, et que nous nommerons désormais *matrices* pour les besoins de l’analyse. En effet, comme nous le verrons par la suite, la réalisation d’une pièce consiste à se servir des matrices pour point de départ d’un plus ou moins grand nombre de variations. »

Le Bomin (2000) [234, p. 141]

C25

« [Le xylophone soliste] fait appel à quatre niveaux de référence, trois matrices et un énoncé type. Ces énoncés n’ont pas cependant une origine unique qui serait la matrice 1, la version minimale, mais s’inscrivent dans un continuum où chaque énoncé est la propre référence de celui qui lui succède[.] [...] Ainsi, chaque référence peut être considérée comme enrichissement du matériau mélodico-rythmique de la référence précédente. Dans ce sens, on peut se demander si les matrices 2 et 3 ne seraient pas des énoncés types d’innovations techniques s’étant cristallisés. Peut-être existera-t-il un jour une matrice avec deux lames jouées en commun par les deux mains. »

Le Bomin (2000) [234, p. 278]

C26

« Ces références que nous avons qualifiées de *prototypiques* sont à distinguer de l’épure qui, correspondant à une modalité de réalisation jamais utilisée en contexte mais uniquement dans l’apprentissage, ne constitue pas une référence dans l’interprétation. »

Le Bomin (2000) [234, p. 354]

C27

« Nous entendrons ici par référence, un matériau métrique et mélodico-rythmique d'une pièce dans une présentation minimale et à laquelle le musicien se réfère dans son interprétation. »

Le Bomin (2004) [237, p. 179]

C28

« Une matrice est une structure sonore strictement répétitive issue de l'utilisation stricte d'une modalité de jeu. Elle se présente sous la forme d'une réalisation minimale adaptée aux contraintes de réalisation de chaque modalité de jeu. Pour chaque pièce, il en existe autant que de modalités de jeu. »

Le Bomin (2004) [237, p. 182]

C29

« Cette version minimale pourrait être qualifiée de "référence neutre" en ce qu'elle ne fait l'objet d'aucune marque d'interprétation d'aucun ordre de la part de l'instrumentiste qui la réalise. Il s'agit en quelque sorte d'une référence abstraite, enfouie au fond de la mémoire de l'instrumentiste, et que seule une sollicitation particulière permet de faire ressurgir à la surface. Pour une pièce donnée, et sur un même instrument, cette référence sera en tout point réalisée de façon identique quel que soit l'instrumentiste qui la réalise. »

Le Bomin (2004) [237, p. 182]

C30

« Il est alors possible de distinguer des références culturelles, les matrices qui se transmettent au cours de l'apprentissage et qui contiennent simultanément les traits identitaires des pièces et des modalités de jeu qui les matérialisent, des références de nature individuelle, véritables signatures musicales et techniques des instrumentistes qui les ont élaborées. »

Le Bomin (2004) [237, p. 187]

C31

« Cela signifie que l'énoncé de référence pour le xylophoniste n'est pas celui de la matrice, mais l'adaptation qu'il en effectue. Au cours de l'interprétation, les variations ne seront pas construites à partir de la matrice, mais à partir de cet énoncé que nous appellerons dorénavant *énoncé type*. Celui-ci est l'énoncé qui, au cours d'une interprétation, revient le plus souvent non modifié et que l'instrumentiste joue le plus fréquemment. »

Le Bomin (2004) [237, p. 188]

C32

« Arom parle constamment, à juste titre, de la reconstitution du système (émique) sous-jacent aux manifestations sonores. Nous sommes moins d'accord avec lui lorsqu'il qualifie ce système de "théorie" sous-jacente (pp. 18, 255, 877), dans la mesure où le terme de théorie connote, dans l'usage scientifique occidental, la verbalisation explicite et abstraite d'un fonctionnement quelconque. Il est d'autant plus important de faire la différence entre système et théorie que l'ethnomusicologie s'est récemment tournée, avec les travaux de Zemp et de Feld, vers les ethno-théories, vers les théories musicales explicitement proposées par les autochtones. »

Nattiez (1991) [303, p. 79, note 3].

C33

« **Conception** état abstrait de la pièce musicale qui renvoie à l'énoncé mélodico-rythmique de référence qui préside la réalisation de la pièce. » Mifune (2012) [295, p. 503]

B.2 Chapitre 5 : Inférence phylogéographique

Les données présentées dans cette annexes viennent compléter celles présentées dans le chapitre 5. Les explications concernant la construction et la signification des figures présentées sont données dans ce chapitre².

B.2.1 Fichiers informatiques

Les fichiers informatiques reproduits dans cette section sont ceux utilisés pour lancer l'analyse de phyloplastie présentée dans la partie 5.1 page 157. Le lancement de cette analyse nécessite trois fichiers d'entrée et se fait grâce à la commande suivante :

```
bayestraits Arbre.nex Coordonnees.csv < BayesTraits.input > BayesTraits.output
```

Voici l'explication des différents éléments de cette commande :

bayestraits	Instruction de lancement du logiciel BayesTraits
Arbre.nex	Topologie de l'arbre décrivant les relations entre les taxons étudiés. Ce fichier est décrit plus en détail dans la partie B.2.1.2 page 389.
Coordonnees.csv	Coordonnées géographiques des taxons étudiés. Ce fichier est décrit plus en détail dans la partie B.2.1.1 page 388.
<	Signal de prendre les instructions d'entrée dans le fichier suivant, c'est-à-dire « BayesTraits.input ».
BayesTraits.input	Liste d'instructions à envoyer au logiciel BayesTraits. Ce fichier est décrit plus en détail dans la partie B.2.1.3 page 391.
>	Signal d'écrire les résultats de l'analyse dans le fichier suivant, c'est-à-dire « BayesTraits.output ».
BayesTraits.output	Fichier dans lequel les résultats sont enregistrés.

2. ↑ Voir chapitre 5 pages 155 à 185.

B.2.1.1 Coordonnées des taxons de l'arbre

Le fichier « Coordonnees.csv » donne les coordonnées des 58 taxons étudiés. Il comprend trois colonnes donnant pour chaque taxon 1) son identifiant, 2) sa longitude et 3) sa latitude.

Fichier B.1 : « Coordonnees.csv »

akelesekas	11.57097	-1.63993
tsoghoseka	11.57097	-1.63993
akelemimon	10.38885	-0.86186
babongoytr	11.06231	-1.03987
babongotch	11.09453	-1.04299
babongonyo	11.05321	-1.03131
tsoghonyoy	11.05321	-1.03131
tsoghomata	10.67501	-1.00146
baghamadem	10.79282	-3.53151
punudemand	11.05106	-2.49692
punudedoua	11.22366	-2.64393
akelemoual	11.0433	-1.80188
babongonga	11.04605	-1.72944
akelebolap	12.02958	-2.0129
babongomoy	11.87384	-1.90336
apinjidest	10.31656	-0.66205
akelemanam	12.87796	-1.04093
babongonan	12.87796	-1.04093
bawandjlip	12.836	-1.02422
awandjidum	12.96242	-0.84329
babongodou	12.96242	-0.84329
barimbeman	11.05106	-2.49692
gisiramand	10.61073	-1.28079
gaka	10.04563	-0.9212
orungu	8.81336	-0.80338
nkmi	9.28819	-1.58161
teke	14.43757	-1.26049
tekedeobia	13.82301	-1.2144
tekedeobeg	13.7714	-1.35253
obambadeon	13.77501	-1.56531
obambadeey	13.77637	-1.55593
obambadeom	13.77389	-1.33366
bakanighi	13.60357	-1.81875
eshira	10.65247	-1.445
sango	11.92016	-1.33689
sangodedya	11.72901	-1.51569
pvi	11.90722	-1.28437
simba	11.97136	-1.18439
nhani	11.50098	2.08745
mweng	12.06158	2.16762
clek	11.56451	1.6006
fwg	12.0911	2.25736
haka	12.18762	2.10727
bakoyekata	14.29502	0.67584
longm	13.99517	1.03385
bakoyaimbo	13.99517	1.03385
mwssa	13.99517	1.03385
kotadeegop	14.12067	1.25998
kweledemaz	14.08794	1.37649
babongoeka	14.13539	-1.36775
babongobou	13.58038	-2.24155
hahambu	13.58038	-2.24155
mbahouins	13.73587	-1.73863
babongodei	12.69159	-1.84558
babongomog	12.65128	-1.80453
babongomak	12.42503	-1.33933
babongomid	12.3644	-1.44842
tsoghoestu	9.67268	0.37055
extragroup	-	-

B.2.1.2 Topologie de l'arbre

Le fichier « Arbre.nex » fournit l'arbre décrivant les relations entre les taxons étudiés. Il est composé de :

- Un en-tête contenant le format de fichier (nexus).
- Une instruction signifiant le début d'une ensemble d'arbre (« Begin trees »).
- Un ensemble de lignes définissant, pour chaque population, un numéro pour l'identifier (ligne « translate » et les 59 lignes suivantes).
- Une ligne (commençant par « tree ») définissant les relations entre les taxons, c'est-à-dire la topologie de l'arbre, au format nexus [A278].
- Une instruction de fin de l'ensemble d'arbres (« End ; »).

Fichier B.2 : « Arbre.nex »

```

#NEXUS

Begin trees; [Treefile saved Mon Feb 17 13:26:40 2014]
translate
1 akelebolap,
2 akelemanam,
3 akelemimon,
4 akelemoual,
5 akelesekas,
6 apinjidest,
7 awandjidum,
8 babongobou,
9 babongodei,
10 babongodou,
11 babongoeka,
12 babongomak,
13 babongoman,
14 babongomid,
15 babongomog,
16 babongomoy,
17 babongonga,
18 babongonyo,
19 babongotch,
20 babongoytr,
21 baghamadem,
22 bahumbu____,
23 baka_____,
24 bakanighi____,
25 bakoyaimbo,
26 bakoyekata,
27 barimbeman,
28 bawandjiip,
29 bongam_____,
30 eshira_____,
31 fing_____,
32 gaba_____,
33 gisiramand,
34 kotadeegop,
35 kweledemaz,
36 mbahouins____,
37 mweng_____,
38 mwessa_____,
39 nkami_____,
40 ntumu_____,
41 obambadeey,
42 obambadeom,
43 obambadeon,
44 ckek_____,
45 orungu_____,
46 pumudedoua,
47 pumudemand,
48 pavi_____,
49 sangro_____,
50 sangodedya,
51 simba_____,
52 teke_____,
53 tekedeobeg,
54 tekedeobia,
55 tsoghoestu,
56 tsoghomata,
57 tsoghonyoy,
58 tsoghoseka,
59 extragroup;

tree LLHmulti-ConsMaj50-LB = [&R] (59:12,(((5:2,1:7):16,(3:14,4:27):12):19,((16:46,(((52:8,53:1):3,54:1):7,((43:1,42:0):2,41:2):3):4,11:0):15):12,(((24:3,22:1):10,36:10):7,8:18):6):8):2,(((2:4,13:1):24,(28:8,(7:3,10:5):7):26):13,(9:15,15:17,(12:2,14:10):7):9):7,(((40:0,37:2,44:2,31:1):25,23:19):12,(((26:18,25:13):3,29:1):10,(38:6,(34:4,35:5):2):9):5):11):8):12,(((((((58:2,(20:0,19:5):6,(18:3,(57:0,56:2):3):3):7,17:8):8,6:16):10,(((27:15,55:2):6,50:6):3,48:3):11):3,33:2,30:1,51:0):10,((32:7,45:7):3,39:6):17):2,49:4):8,(47:7,46:6):9):3,21:7):11);

End;

```

B.2.1.3 Fichier d'instructions pour BayesTraits

Le fichier « BayesTraits.input » sert à donner les différentes instructions au logiciel BayesTraits pour lancer l'analyse. Il permet en particulier de spécifier les paramètres utilisés. Chaque ligne correspond à une instruction, et l'analyse n'est lancée qu'à la réception de la dernière instruction « Run ». Les instructions données sont les suivantes :

7 Choix du modèle d'évolution utilisé. Le choix 7 correspond au modèle des contrastes indépendants³⁴.

2 Choix de la méthode d'analyse utilisée. Le choix 2 correspond à la méthode MCMC, décrite en détails dans le chapitre 5⁵.

Iterations Choix du nombre d'itérations. 1,1 milliards d'itérations seront réalisées.

Burnin Choix du nombre d'itérations pendant la période d'allumage de la chaîne. Les 100 premiers millions d'itérations seront considérées faire parties de la période d'allumage.

Sample Choix de l'intervalle d'échantillonnage. Les valeurs des paramètres estimés seront échantillonnés une fois toutes les 100 000 itérations.

VarRates Sélectionne le modèle à taux variables⁶.

AddMRCA [Nœud] [TAXON 1] [...] [TAXON N]

Chacune de ces lignes commençant par l'instruction « AddMRCA » permet d'ajouter un **clade** pour lequel on veut inférer la position ancestrale⁷. Le premier paramètre [Nœud] permet de spécifier le nom du nœud. Les paramètres suivants ([TAXON 1], ..., [TAXON N]) correspondent à une liste des taxons inclus dans ce clade, ce qui permet de définir précisément le clade dont il s'agit. Ainsi le clade *Ngom*⁸ est celui regroupant les taxons « akelesekas » (Ngom de Seka seka), « akelebolap » (Ngom de Bolapessa), « akelemimon » (Ngom de Mimongo) et « akelemoual » (Ngom de Moualo).

Run Instruction pour lancer l'analyse

3. ↑ *Independent contrasts*, en anglais.

4. ↑ Venditti et coll. [430, supplément, pp. 12–16].

5. ↑ Voir [partie 5.2.3](#) pages 162 à 164.

6. ↑ *Variable rates*, en anglais.

7. ↑ Le sigle *MRCA* signifie *plus récent ancêtre commun* (*most recent common ancestor*, en anglais).

8. ↑ Voir l'arbre dans la [partie 5.1](#) page 160 pour visualiser la position de ce clade.

Fichier B.3 : « BayesTraits.input »

```

7
2
Iterations 110000000
Burnin 10000000
Sample 100000
VarRates
AdMRCA      Ngo  akelesekas  akelebolap  akelemimon  akelemoual
AdMRCA      Tek  teke_____  tekedeobeg  tekedeobia  tekedeobia
AdMRCA      Oba  obambadeon  obambadeom  obambadeey  obambadeey  babongobou
AdMRCA      Bo   bakanighi_  bahambu____  mbahouins_  mbahouins_  babongobou
AdMRCA      Ho   babongomoy  teke_____  tekedeobeg  tekedeobia  obambadeon  obambadeom  obambadeey
  babongoeka  bakanighi_  bahambu____  mbahouins_  babongobou
AdMRCA      La   akelemanam  babongoman  bawandjlip  bawandjlip  babongodou
AdMRCA      Bko  babongodei  babongomog  babongomak  babongomid  babongodou  babongodei  babongomog
AdMRCA      Ol   akelemanam  babongoman  bawandjlip  bawandjlip  babongodou  babongodei  babongomog
  babongomak  babongomid
AdMRCA      F    nhamu_____  mweg_____  dek_____  fng_____
AdMRCA      Wn   nhamu_____  mweg_____  dek_____  fng_____  baka_____
AdMRCA      Oi   bakoyekata  bakoyaimbo  bongm_____  mwesa_____  kotadeegop  kweledemaz
AdMRCA      N    nhamu_____  mweg_____  dek_____  fng_____  baka_____  bakoyekata  bakoyaimbo
  bongm_____  mwesa_____  kotadeegop  kweledemaz
AdMRCA      Pa   akelesekas  akelebolap  akelemimon  akelemoual  babongomoy  teke_____  tekedeobeg
  tekedeobia  obambadeon  obambadeom  obambadeey  babongoeka  bakanighi_  bahambu____  mbahouins_
  babongobou  akelemanam  babongoman  bawandjlip  bawandjlip  babongodou  babongodei  babongomog
  babongomak  babongomid  nhamu_____  mweg_____  dek_____  fng_____  baka_____  bakoyekata
  bakoyaimbo  bongm_____  mwesa_____  kotadeegop  kweledemaz
AdMRCA      Bt   babongoytr  babongotch  babongotch  babongonyo  tsoghonyoy  tsoghomata  babongonga
AdMRCA      Ts   tsoghoseka  babongoytr  babongotch  babongonyo  tsoghonyoy  tsoghomata  babongonga
AdMRCA      Ng   tsoghoseka  babongoytr  babongotch  babongonyo  tsoghonyoy  tsoghomata  babongonga
  apinjidest  barimbeman  tsoghoestu  sangodedya  pvi_____  gisiramand  eshira_____  simba_____
AdMRCA      My   gaka_____  orungu_____  nkmi_____
AdMRCA      Pu   punudemand  punudedoua
AdMRCA      Ma   tsoghoseka  babongoytr  babongotch  babongonyo  tsoghonyoy  tsoghomata  babongonga
  apinjidest  barimbeman  tsoghoestu  sangodedya  pvi_____  gisiramand  eshira_____  simba_____
  gaka_____  orungu_____  nkmi_____  punudemand  punudedoua  baghamadem
AdMRCA      Root akelesekas  akelebolap  akelemimon  akelemoual  babongomoy  teke_____  tekedeobeg
  tekedeobia  obambadeon  obambadeom  obambadeey  babongoeka  bakanighi_  bahambu____  mbahouins_
  babongobou  akelemanam  babongoman  bawandjlip  bawandjlip  babongodou  babongodei  babongomog
  babongomak  babongomid  nhamu_____  mweg_____  dek_____  fng_____  baka_____  bakoyekata
  bakoyaimbo  bongm_____  mwesa_____  kotadeegop  kweledemaz  tsoghoseka  babongoytr  babongotch
  babongonyo  tsoghonyoy  tsoghomata  babongonga  apinjidest  barimbeman  tsoghoestu  sangodedya
  pvi_____  gisiramand  eshira_____  simba_____  gaka_____  orungu_____  nkmi_____  sangp_____
  punudemand  punudedoua  baghamadem
Run

```

B.2.2 Synthèse par nœud

Je donne ici le détail des résultats de l'inférence phylogéographique présentée dans le chapitre 5 pages 155 à 185 pour chaque nœud étudié. On se référera à la partie 5.3 pages 165 à 177 pour des explications détaillées concernant la construction et l'interprétation de ces figures. Le tableau B.2 ci-dessous donne, pour chaque nœud étudié, le numéro de la figure de synthèse correspondante et le numéro de la page où se trouve cette figure.

Nœud	Figure	Page
Bko	B.1	394
Bo	B.2	395
Bt	B.3	396
F	B.4	397
Ho	B.5	398
La	B.6	399
Ma	B.7	400
My	B.8	401
N	B.9	402
Ng	B.10	403
Ngo	B.11	404
Oba	B.12	405
Oi	B.13	406
Ol	B.14	407
Pa	B.15	408
Pu	B.16	409
Root	B.17	410
Tek	B.18	411
Ts	B.19	412
Wn	B.20	413

TAB. B.2 : Référence des figures de synthèse pour chaque nœud.

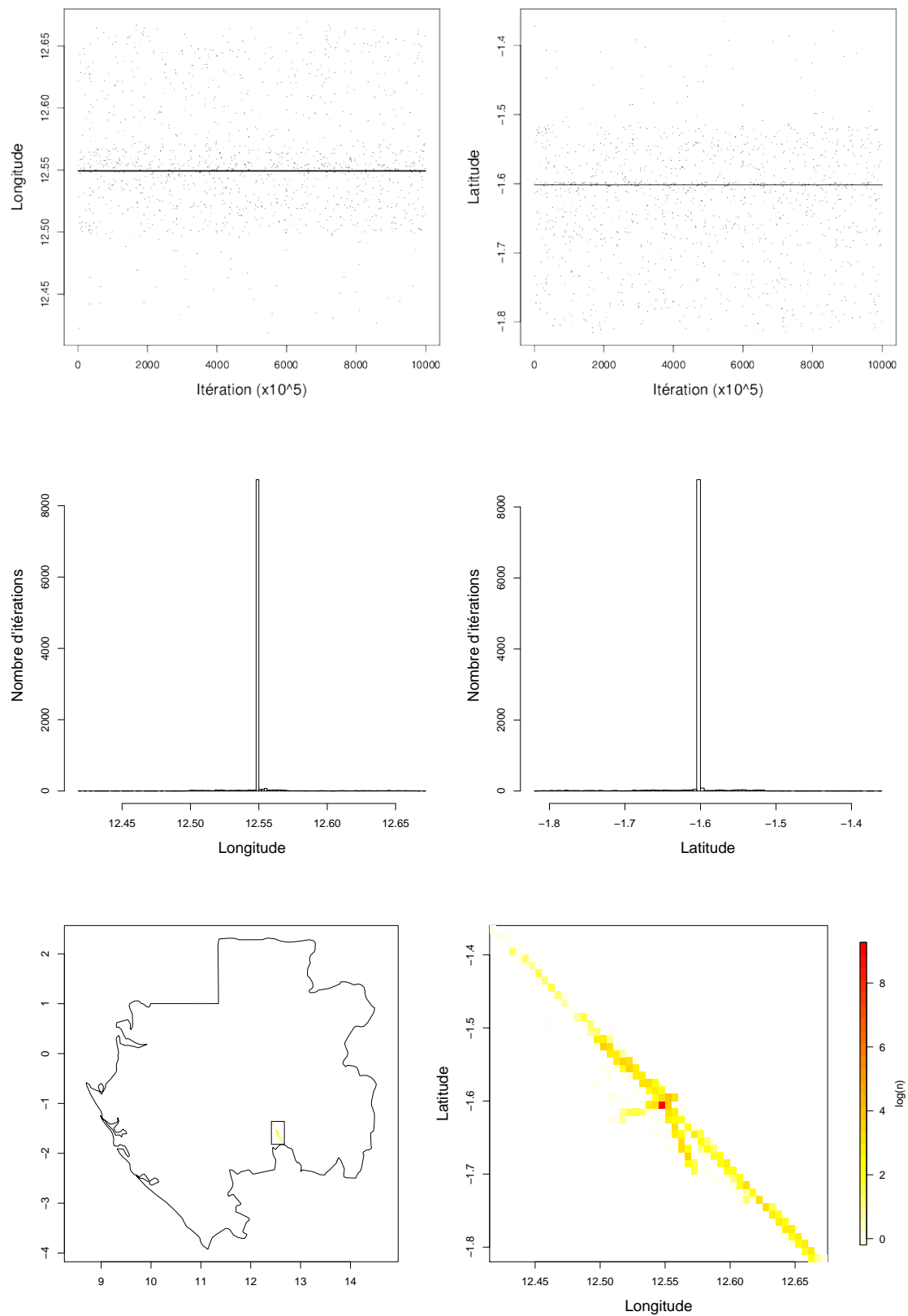


FIG. B.1 : Distribution des localisations échantillonnées pour le nœud Bko. Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

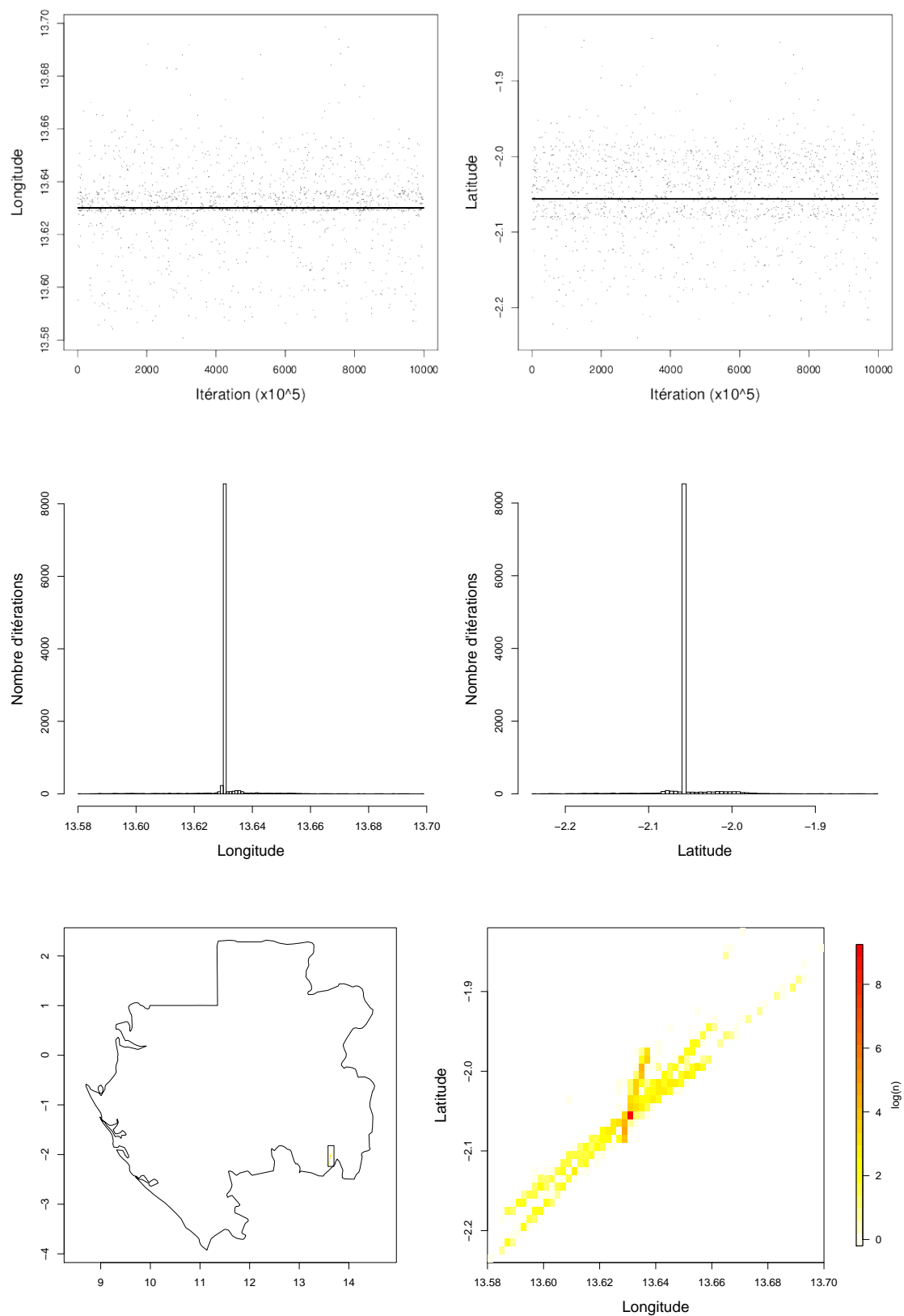


FIG. B.2 : Distribution des localisations échantillonnées pour le nœud Bo. Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

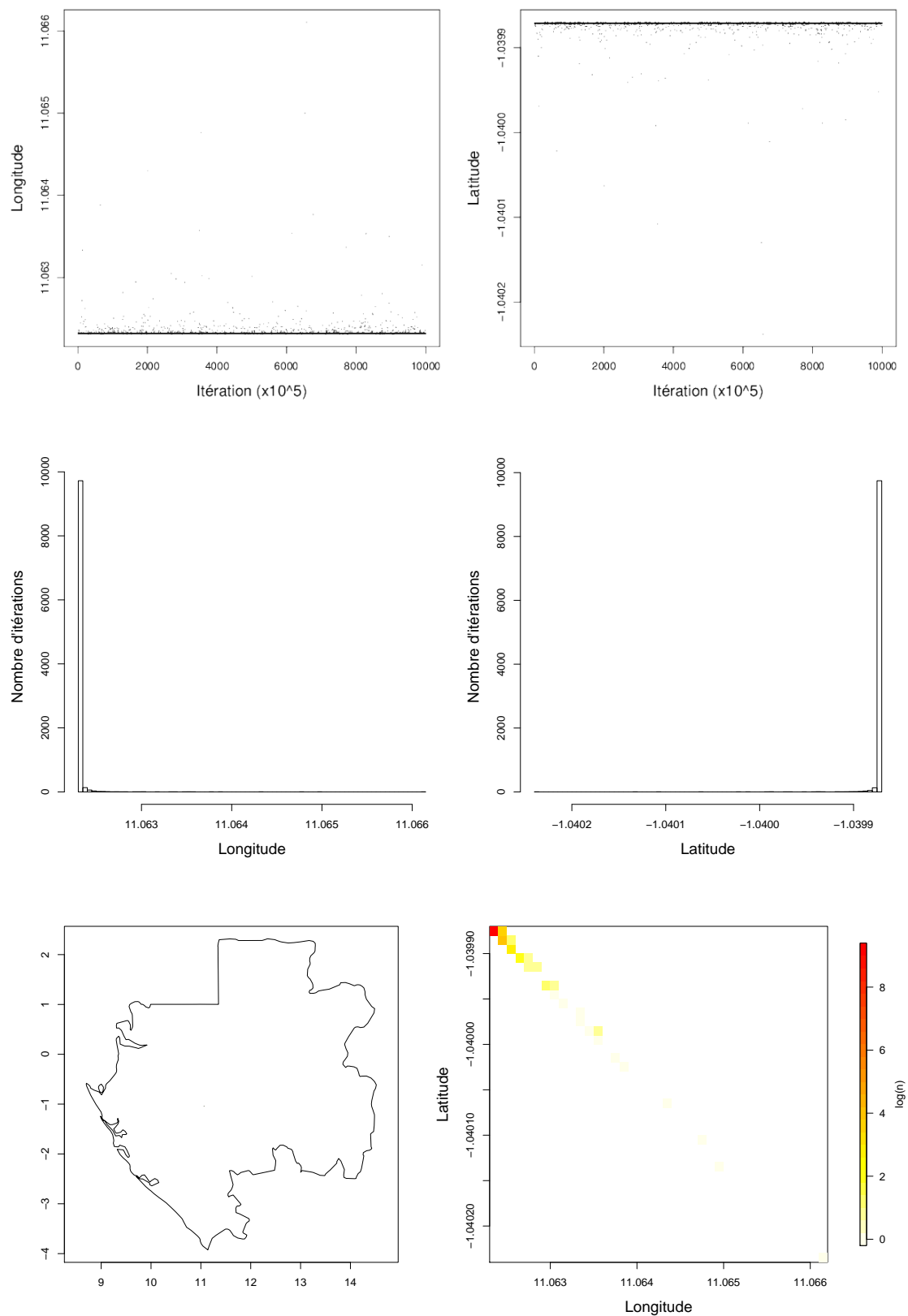


FIG. B.3 : Distribution des localisations échantillonnées pour le nœud Bt. Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

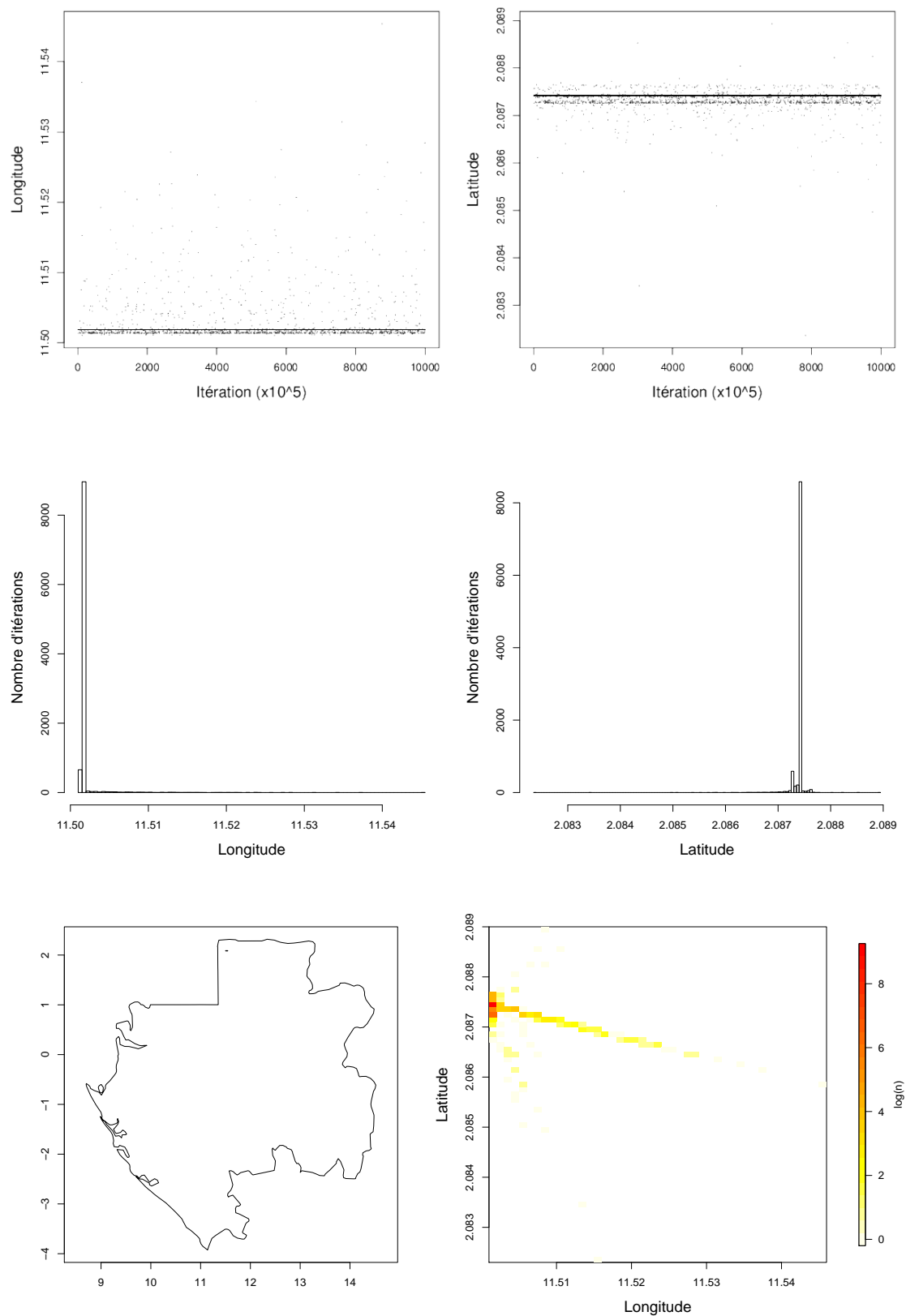


FIG. B.4 : Distribution des localisations échantillonnées pour le nœud F. Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

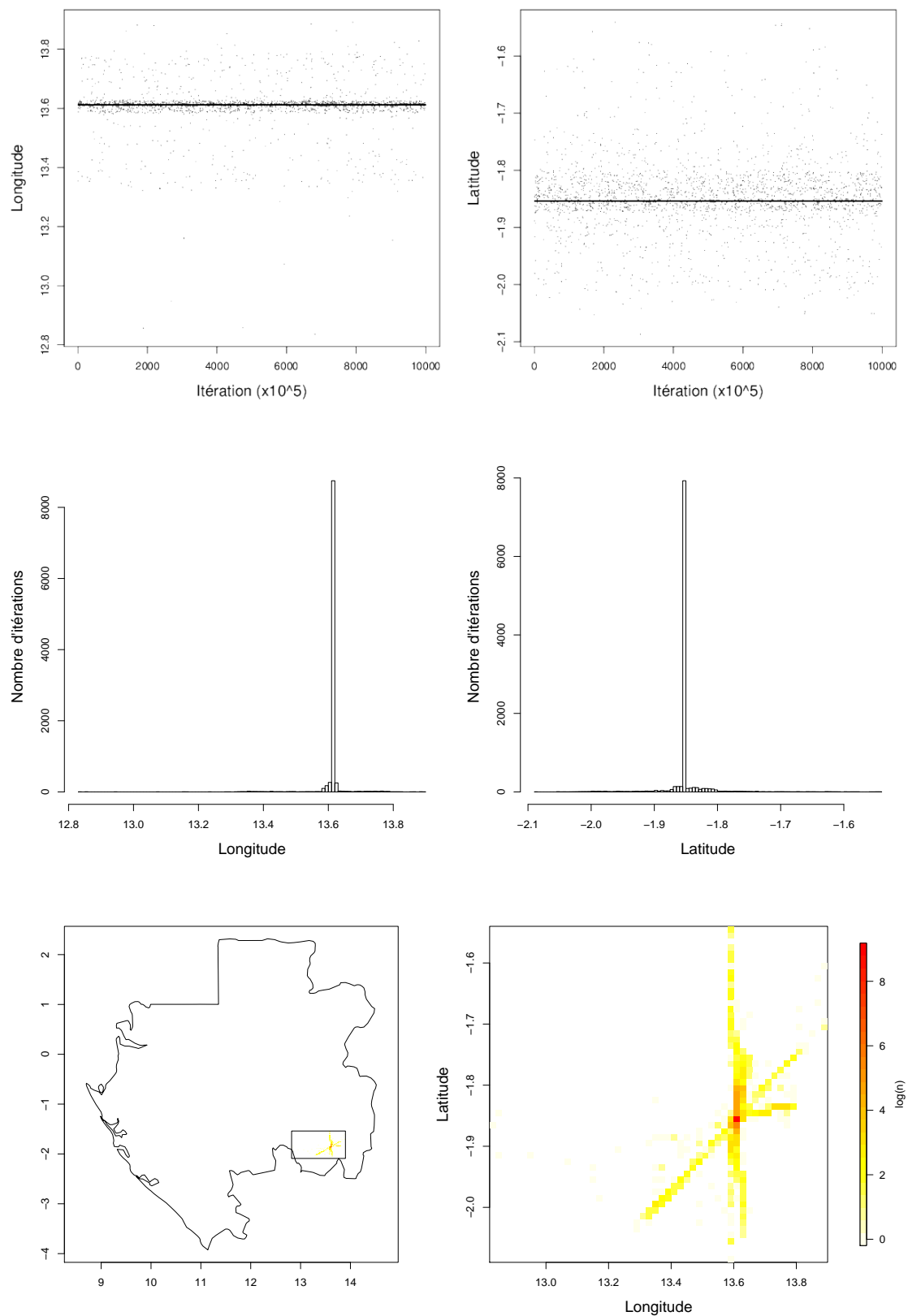


FIG. B.5 : **Distribution des localisations échantillonnées pour le nœud Ho.** Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

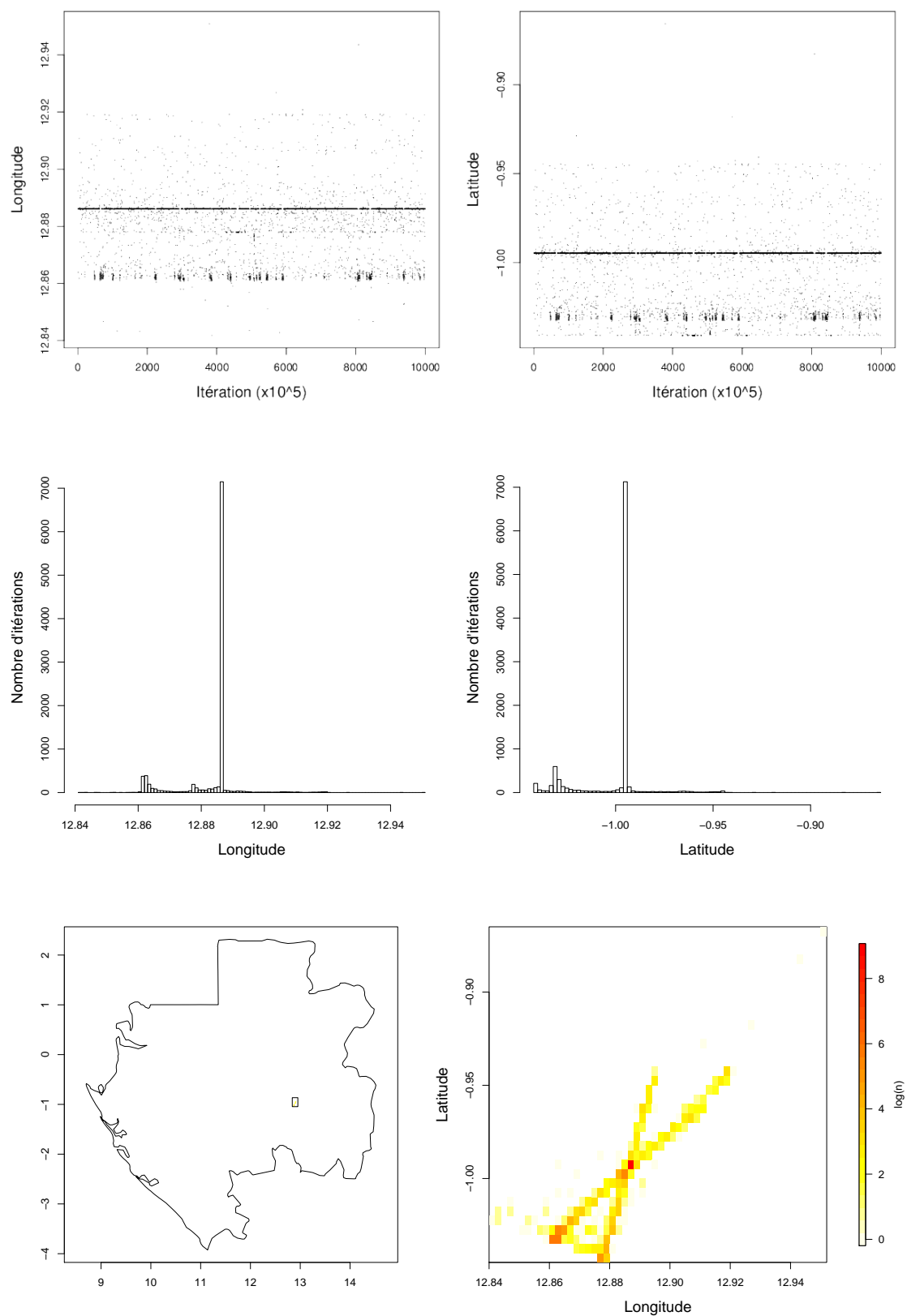


FIG. B.6 : Distribution des localisations échantillonnées pour le nœud La. Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

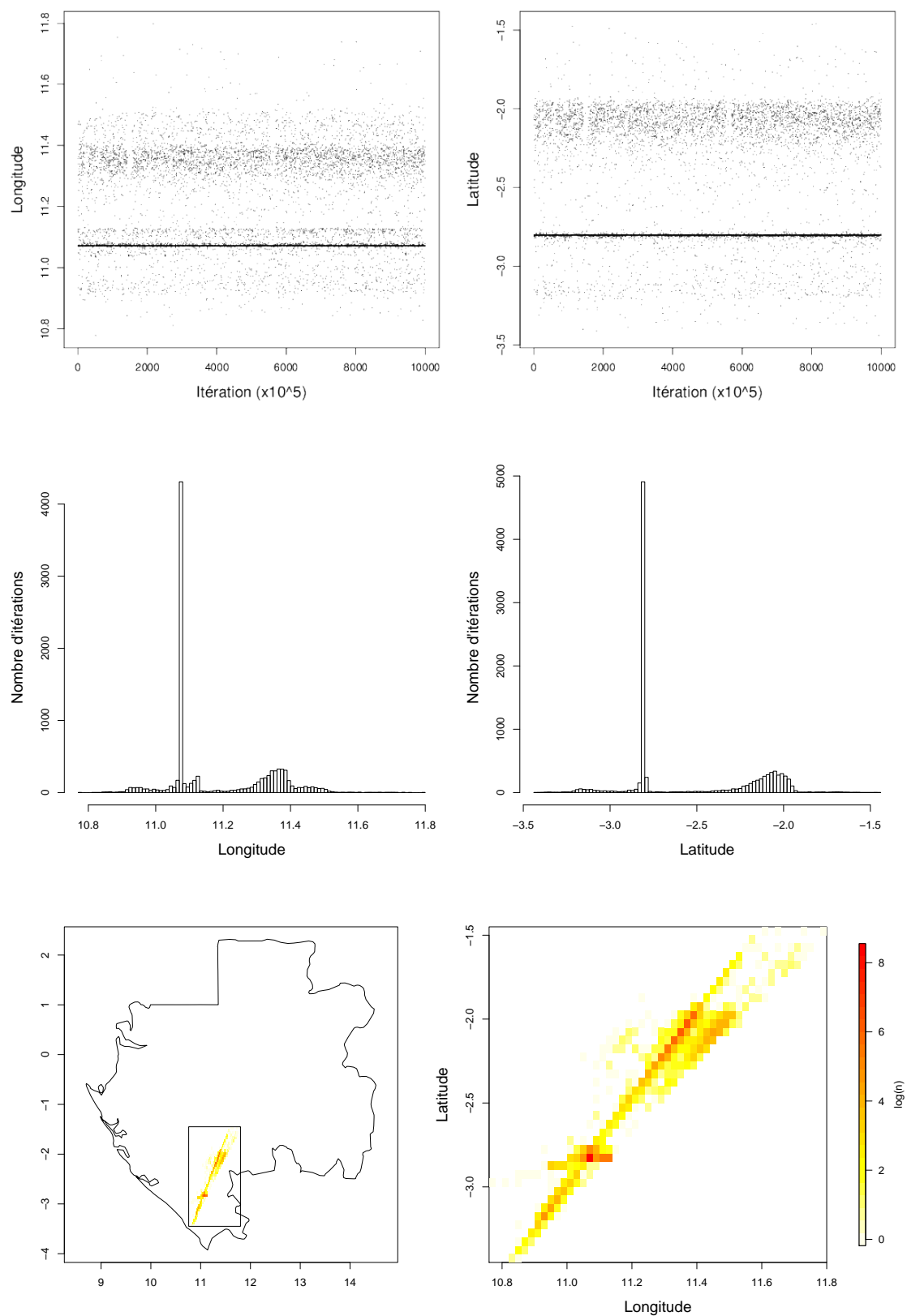


FIG. B.7 : **Distribution des localisations échantillonnées pour le nœud Ma.** Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

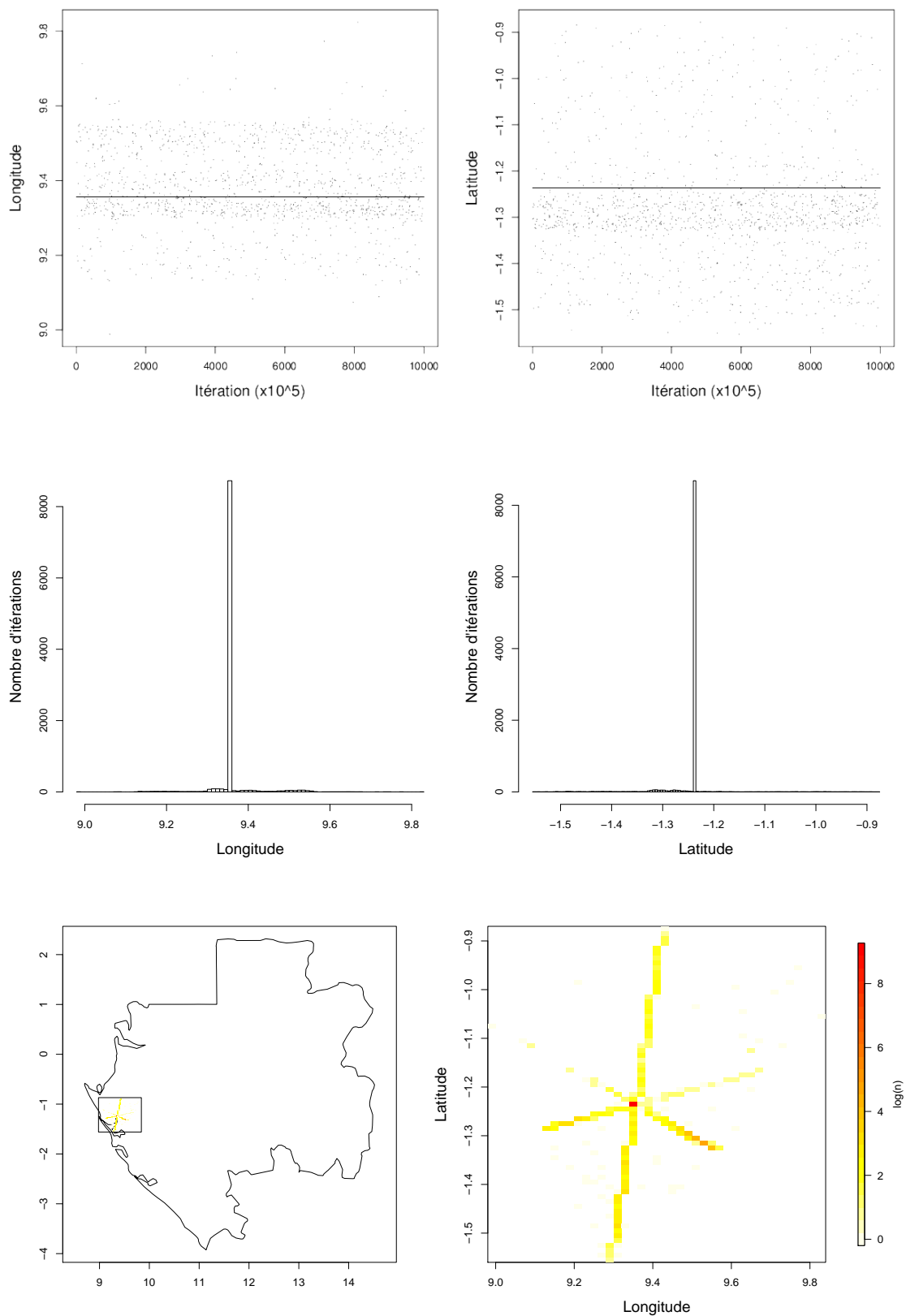


FIG. B.8 : **Distribution des localisations échantillonnées pour le nœud My.** Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

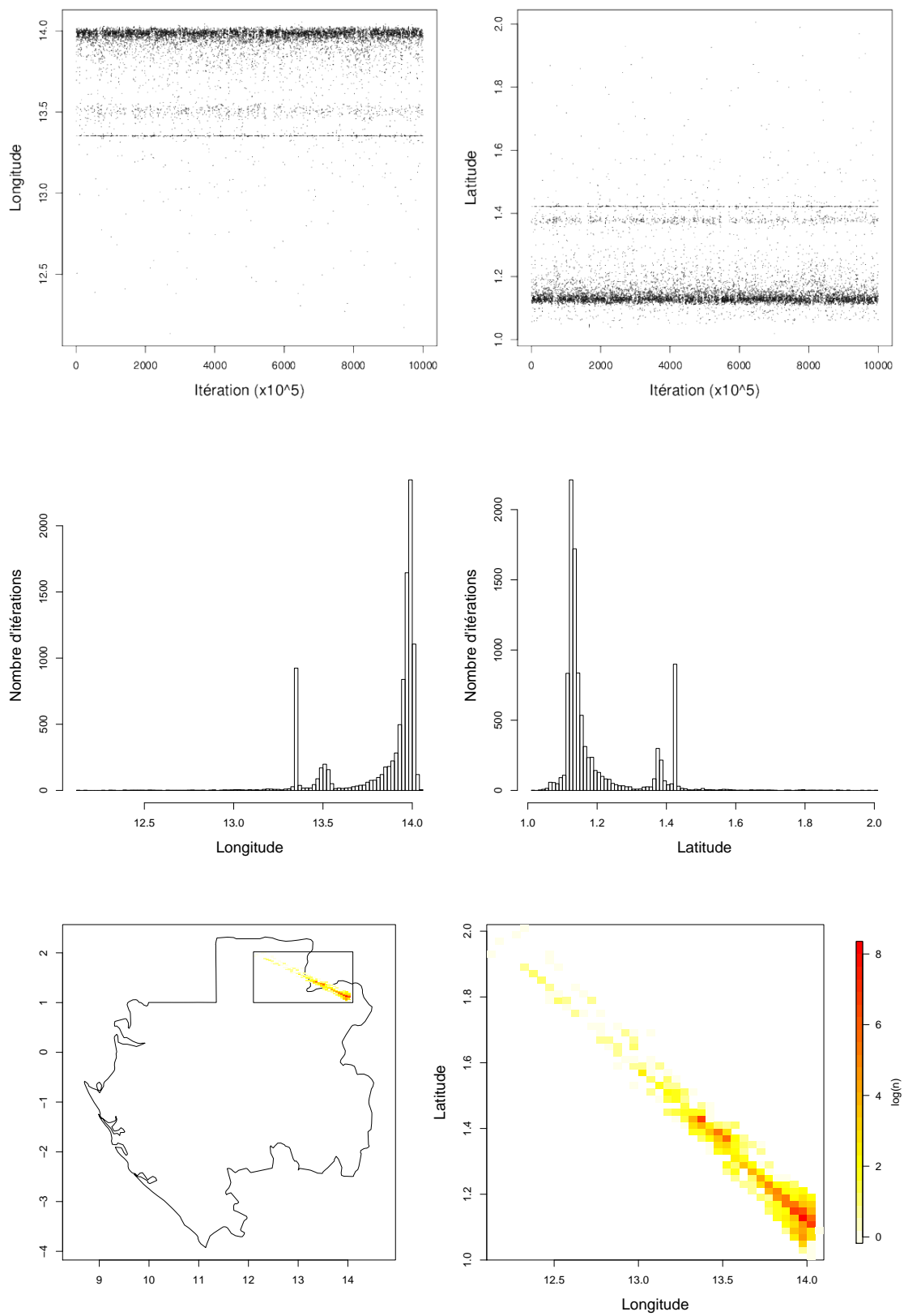


FIG. B.9 : Distribution des localisations échantillonnées pour le nœud N. Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

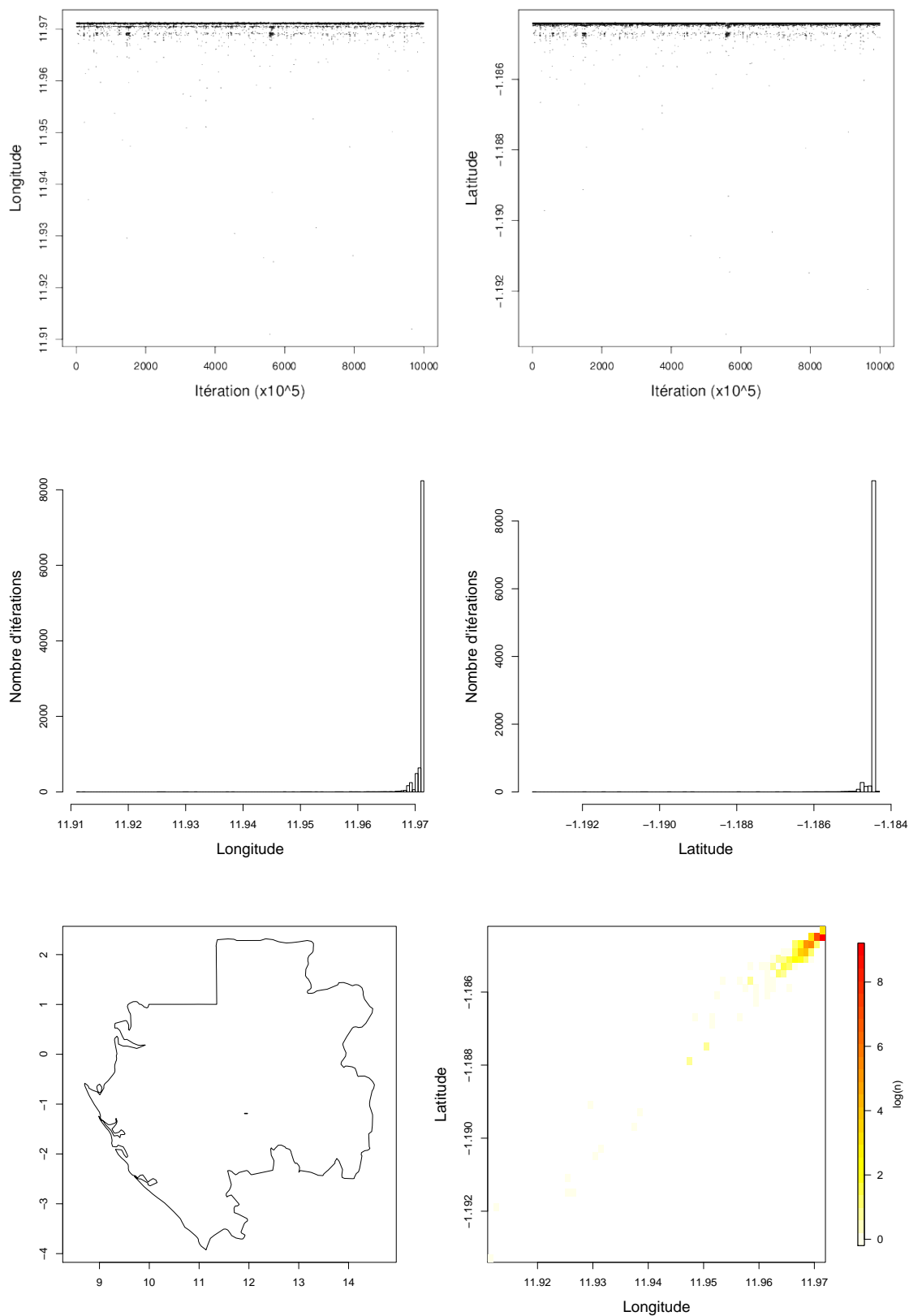


FIG. B.10 : Distribution des localisations échantillonnées pour le nœud Ng. Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

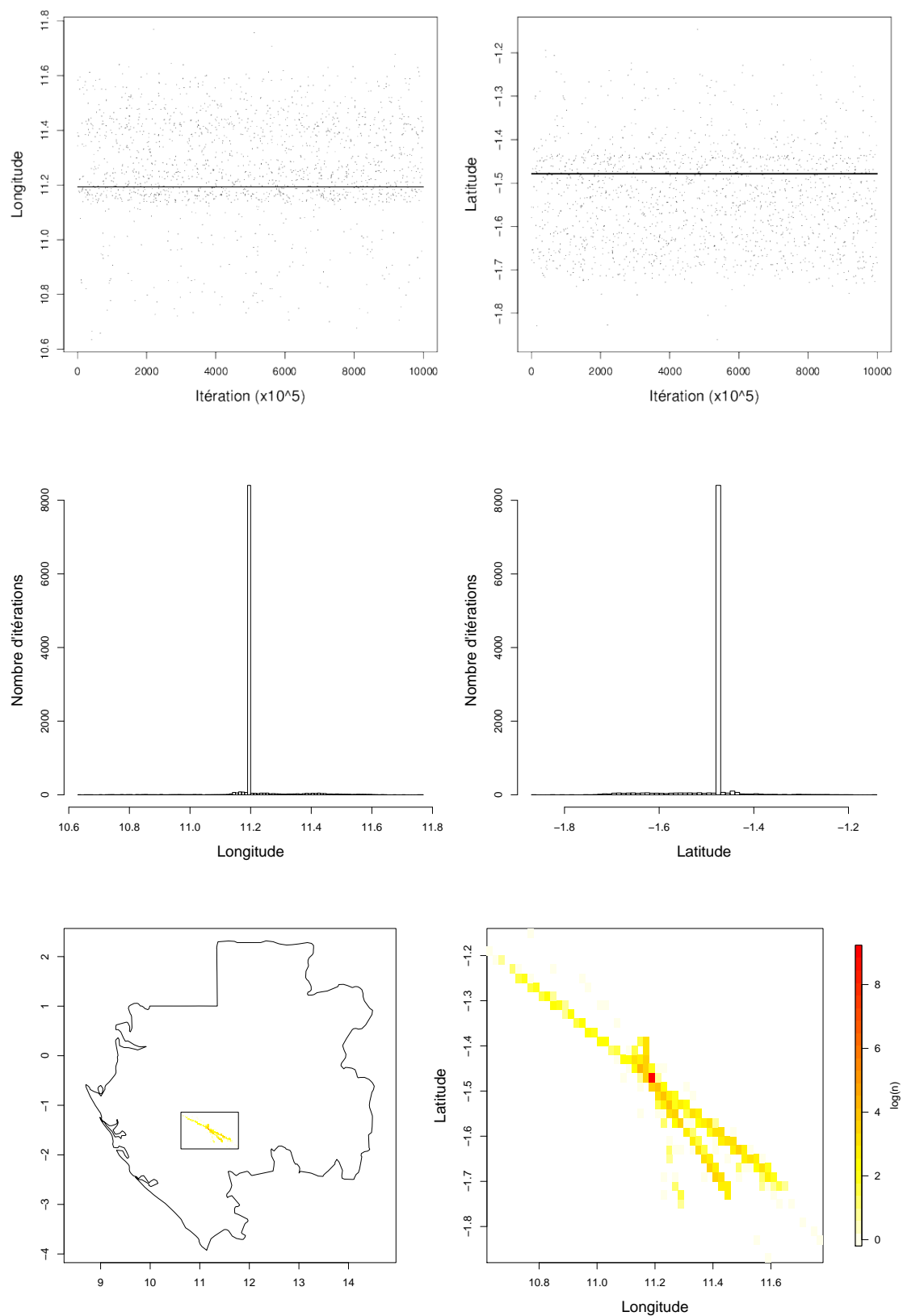


FIG. B.11 : **Distribution des localisations échantillonnées pour le nœud Ngo.** Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

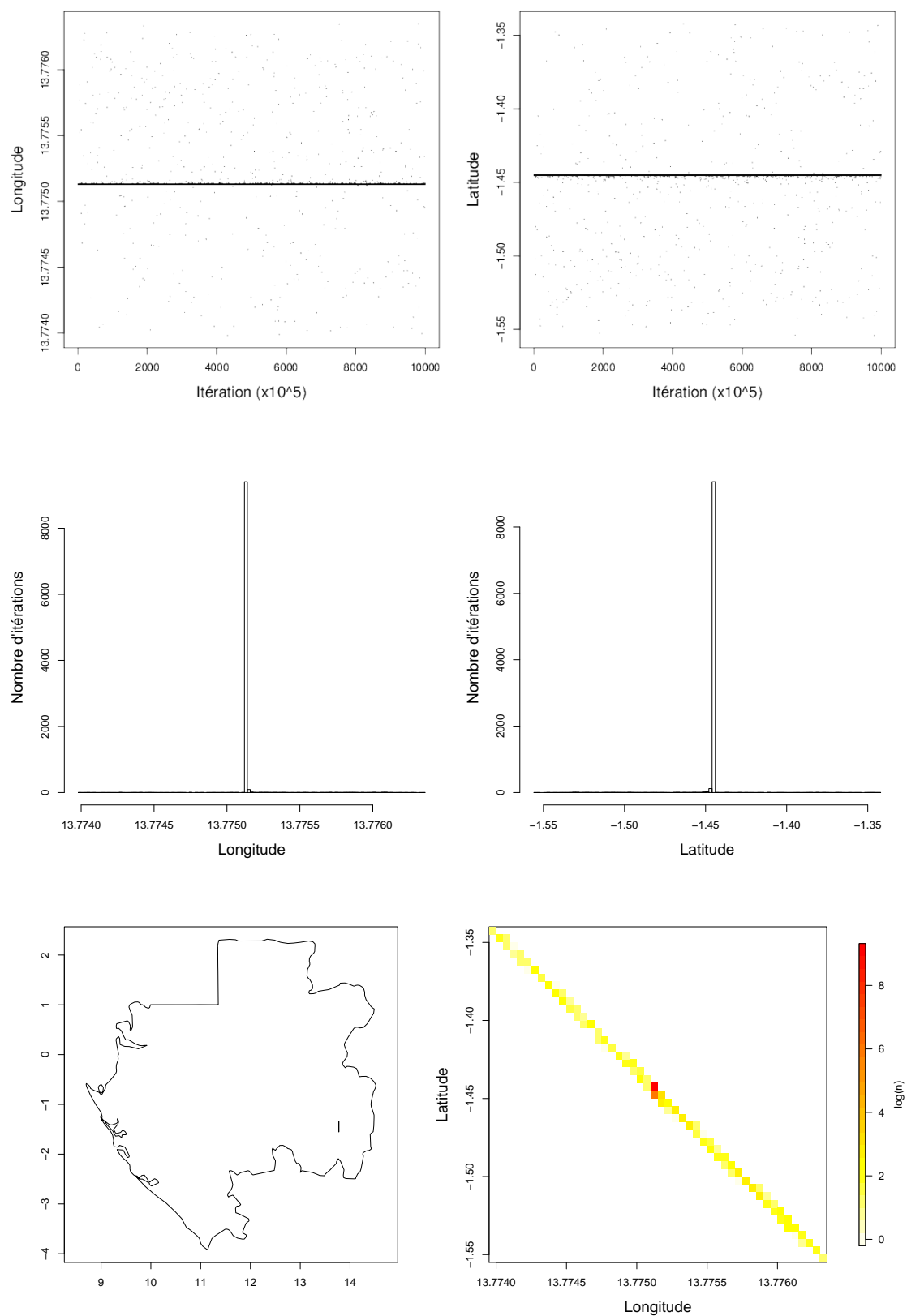


FIG. B.12 : **Distribution des localisations échantillonnées pour le nœud Oba.** Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

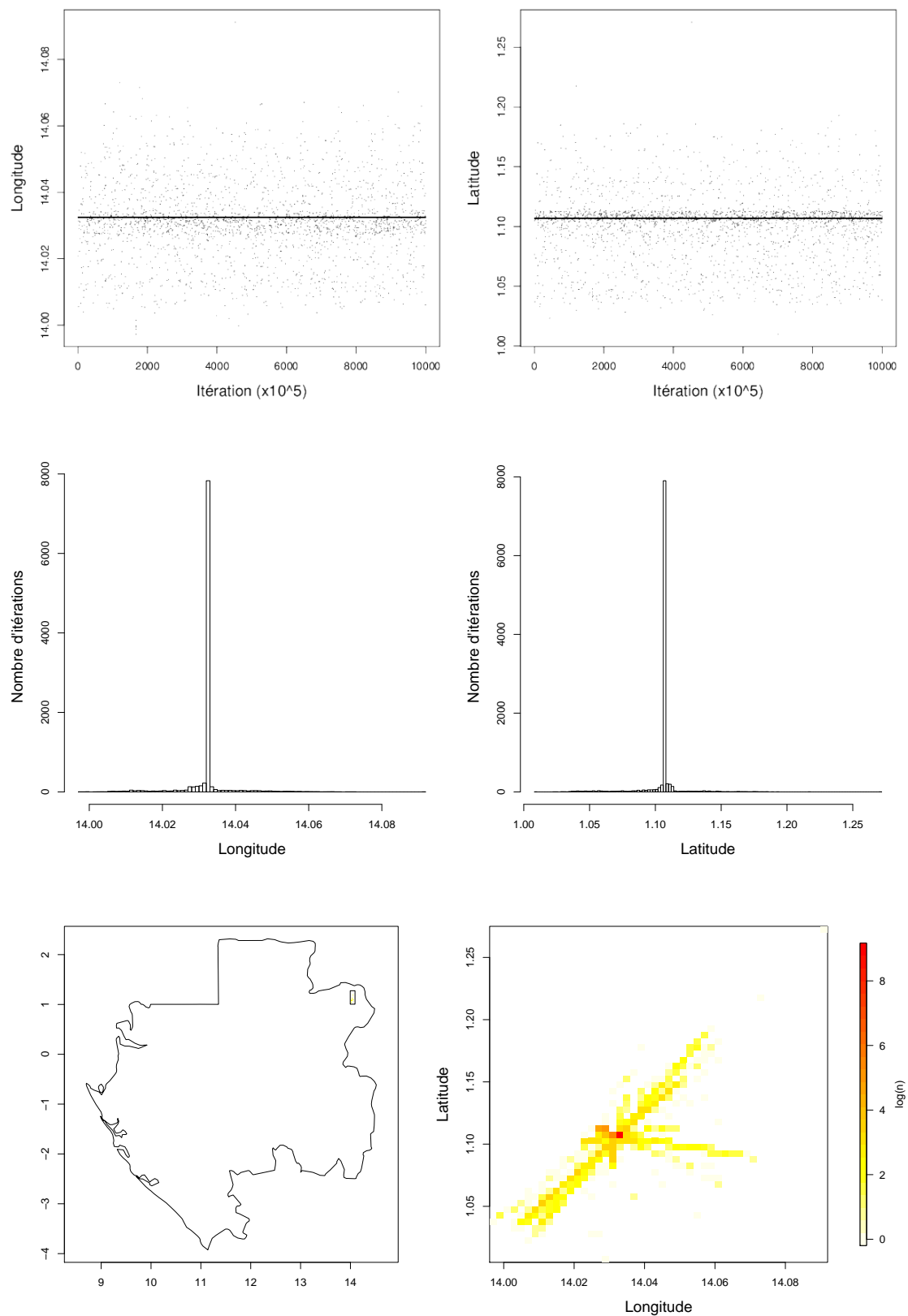


FIG. B.13 : Distribution des localisations échantillonnées pour le nœud Oi. Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

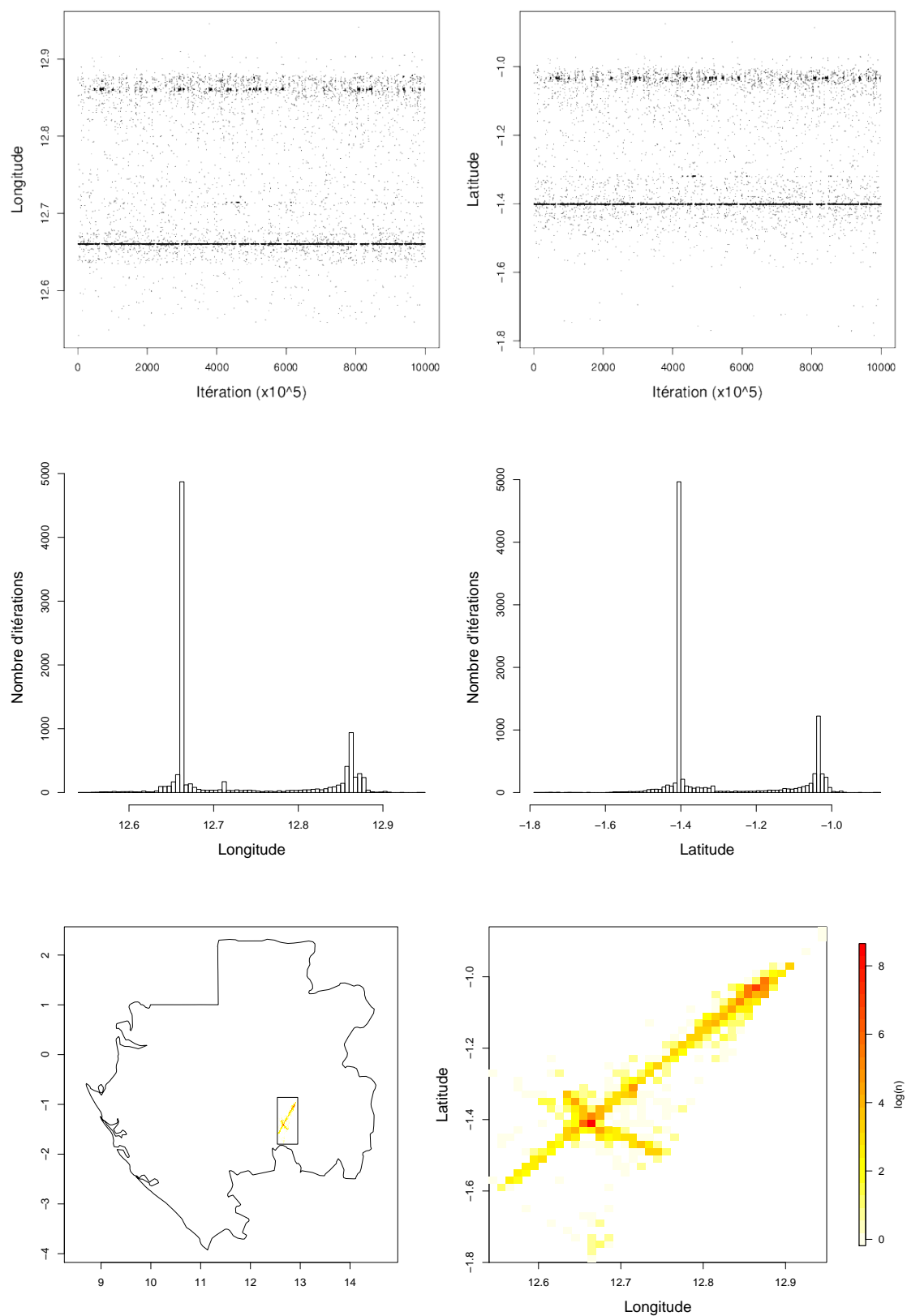


FIG. B.14 : Distribution des localisations échantillonnées pour le nœud O1. Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

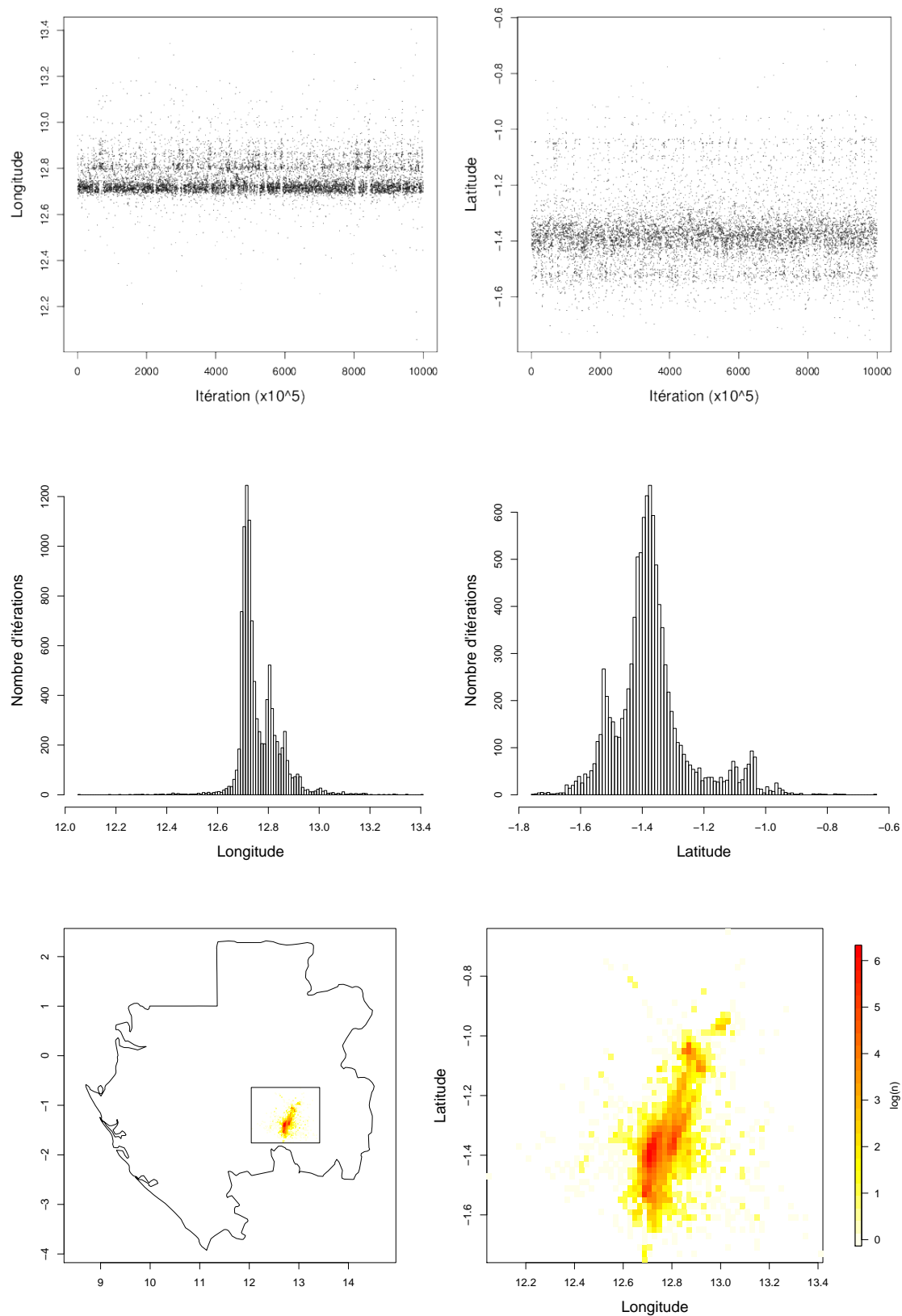


FIG. B.15 : **Distribution des localisations échantillonnées pour le nœud Pa.** Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

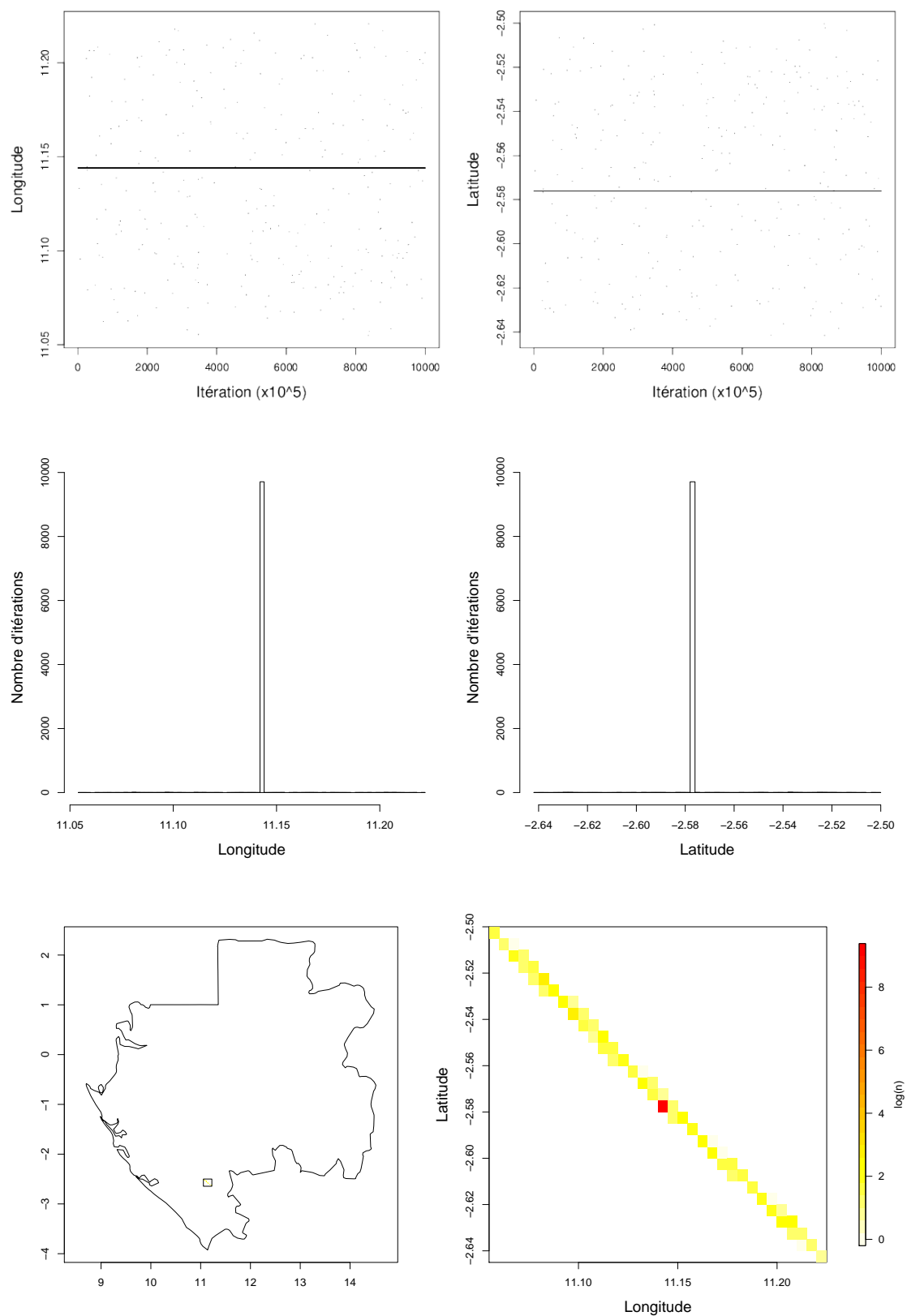


FIG. B.16 : Distribution des localisations échantillonnées pour le nœud Pu. Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

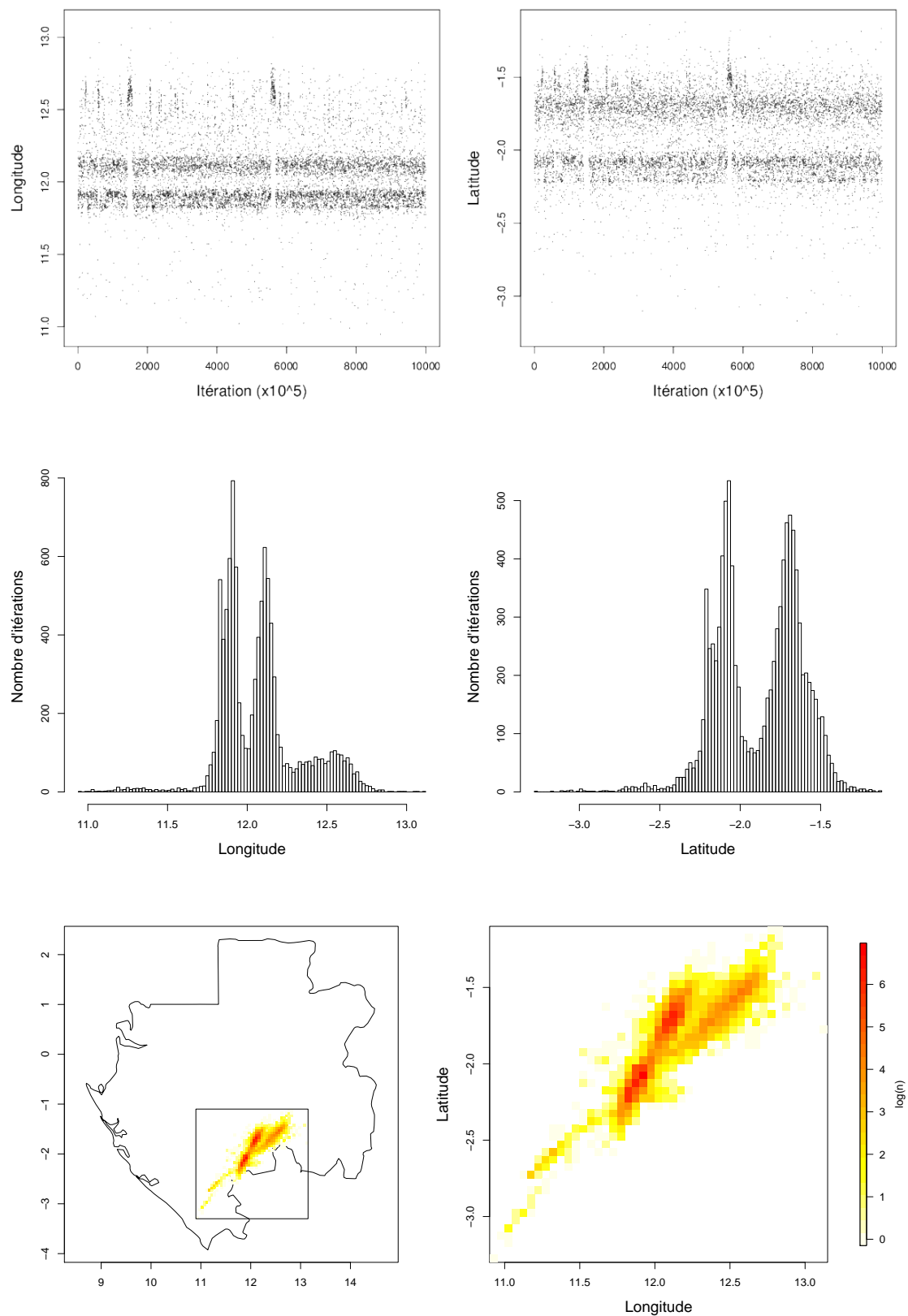


FIG. B.17 : **Distribution des localisations échantillonnées pour le nœud Root.** Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

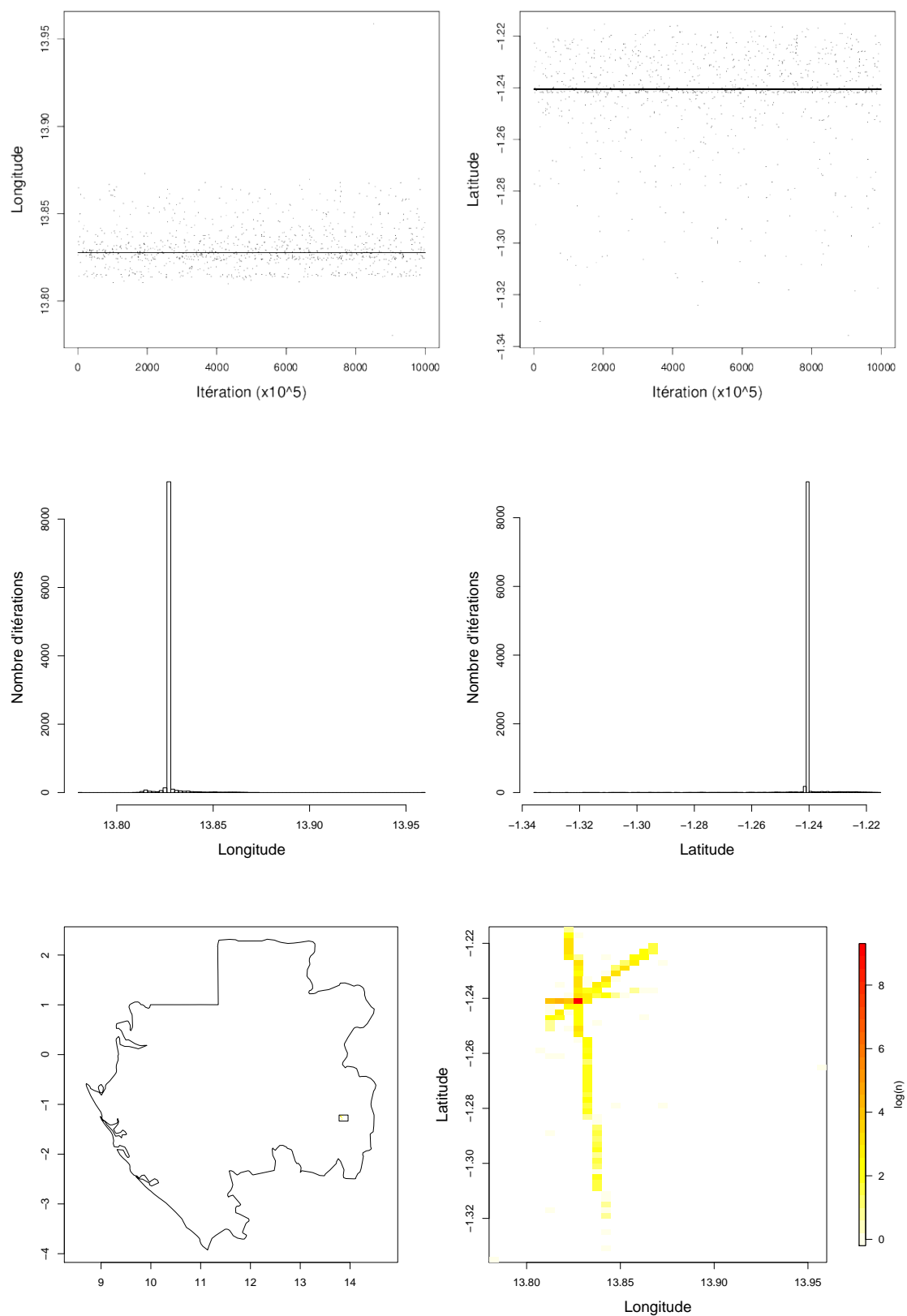


FIG. B.18 : **Distribution des localisations échantillonnées pour le nœud Tek.** Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

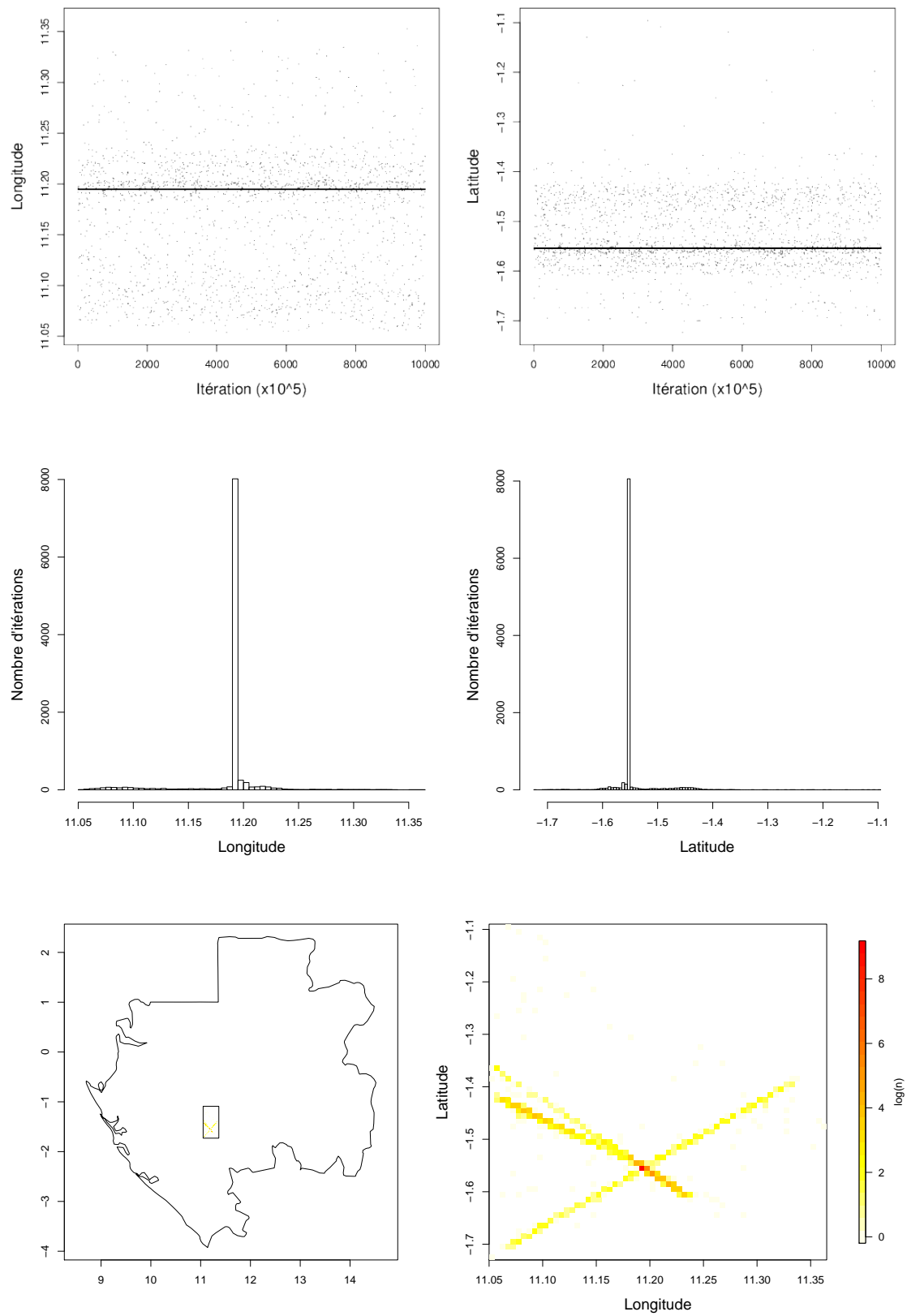


FIG. B.19 : **Distribution des localisations échantillonnées pour le nœud Ts.** Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

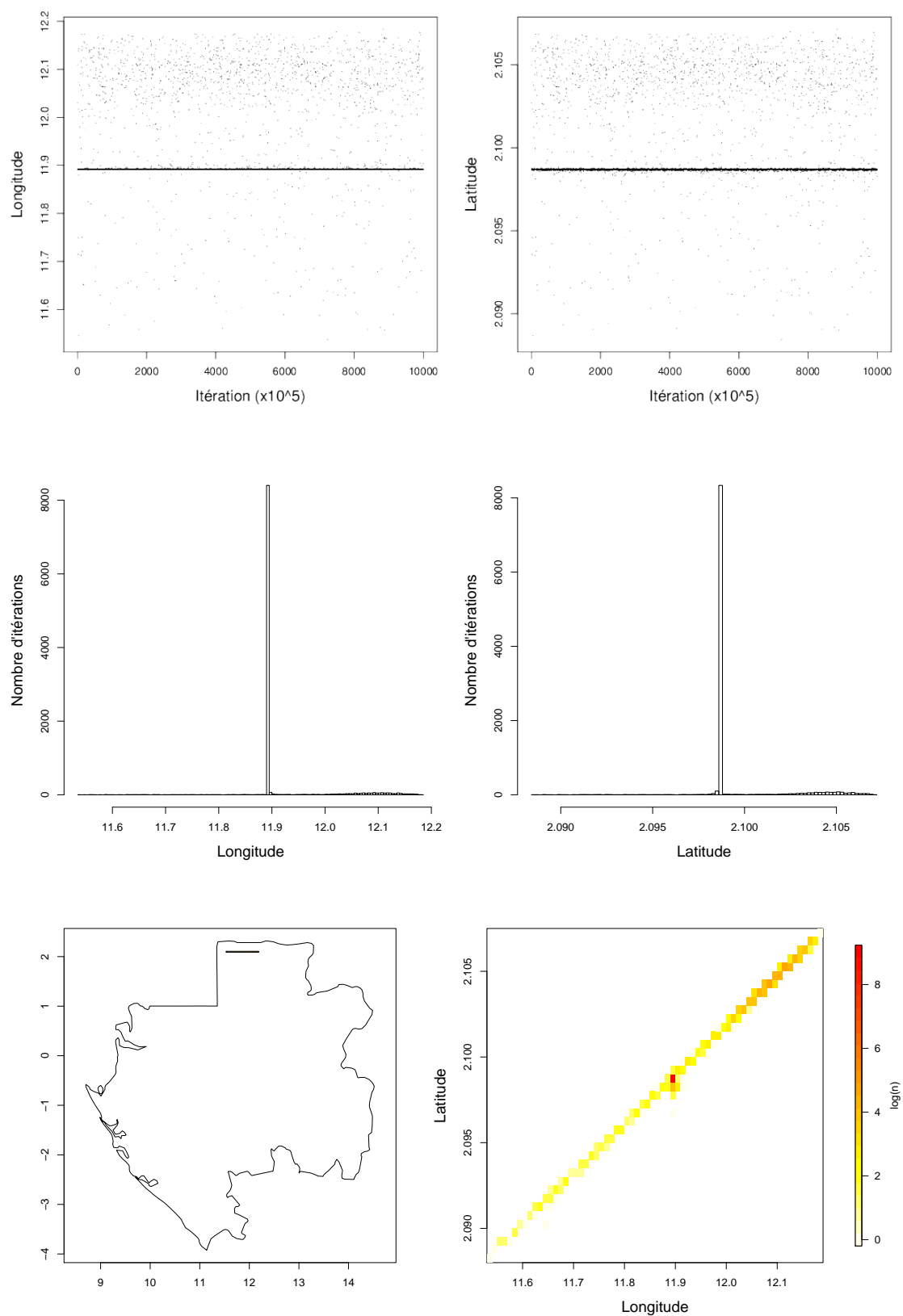


FIG. B.20 : **Distribution des localisations échantillonnées pour le nœud Wn.** Voir partie 5.3 pages 165 à 177 pour des explications concernant la construction et l'interprétation de ces figures.

B.2.3 Dépendance des positions incertaines de différents nœuds

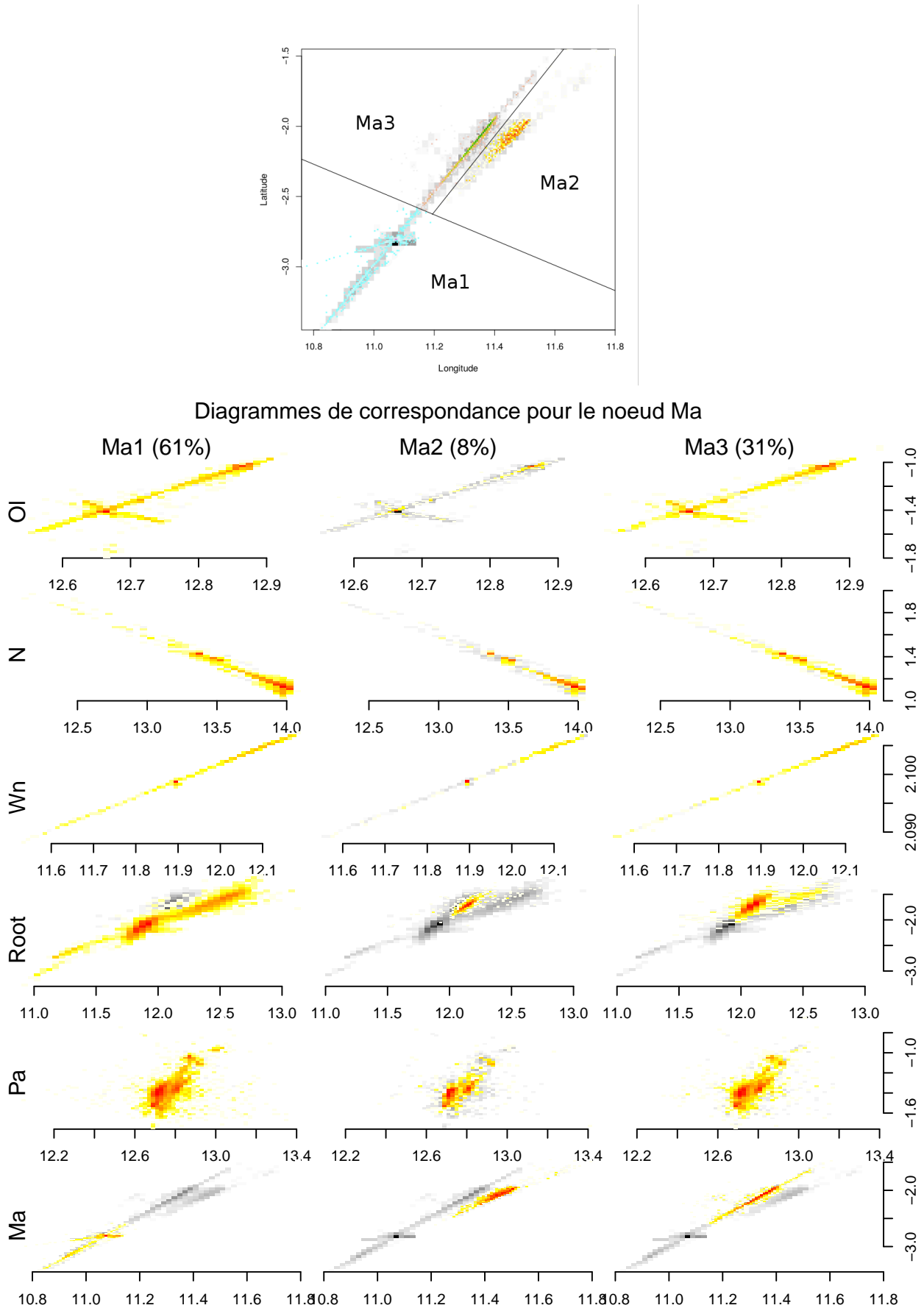


FIG. B.21 : Dépendance à la position des autres nœuds pour le nœud Ma. Voir partie 5.3.3 page 168 pour des explications concernant la construction et l'interprétation de ces figures.

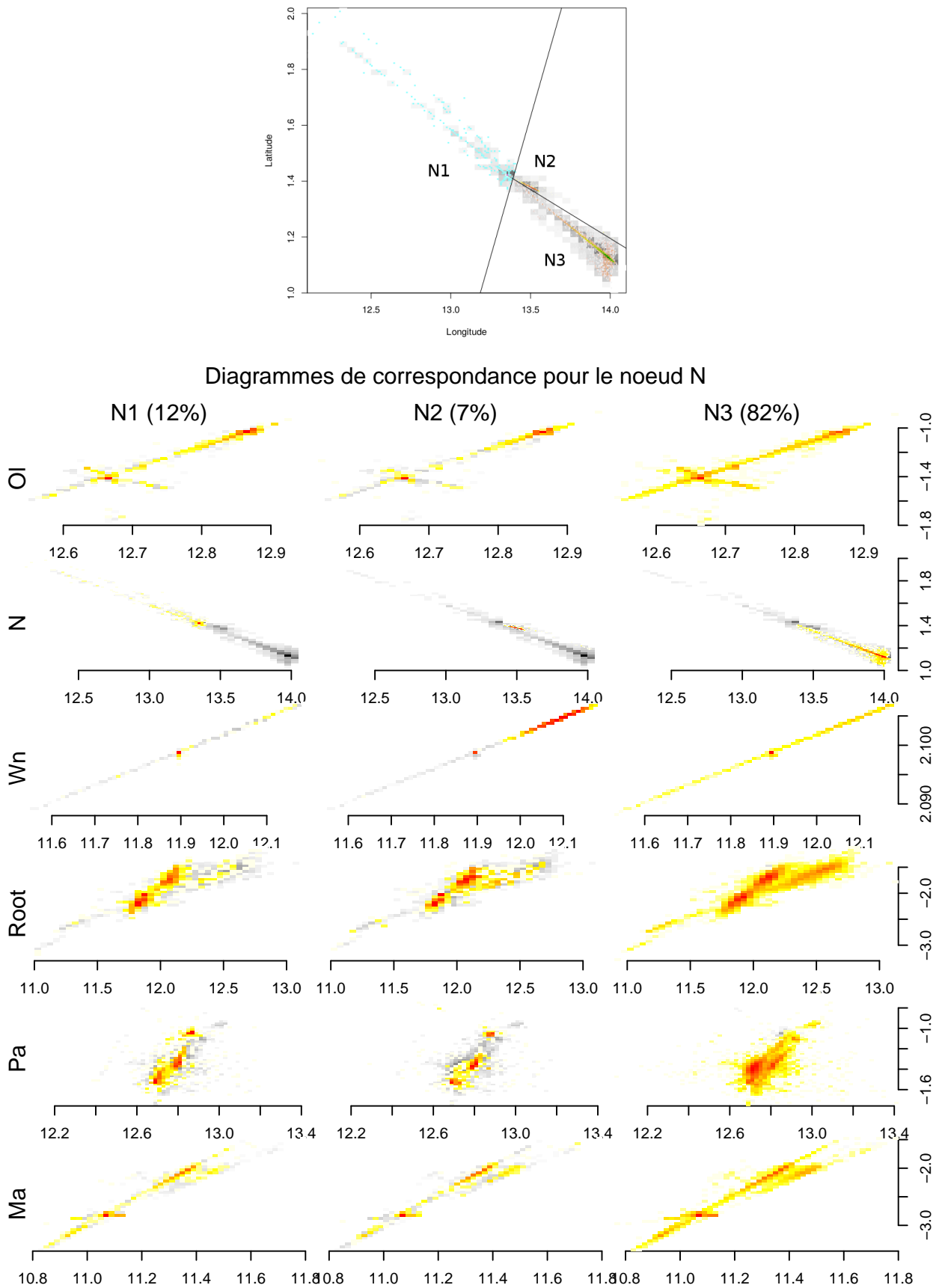
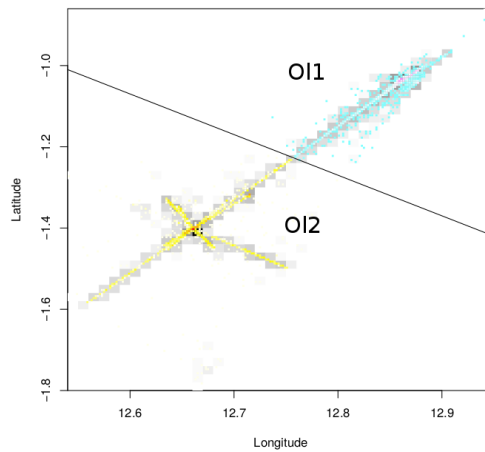


FIG. B.22 : Dépendance à la position des autres nœuds pour le nœud N. Voir partie 5.3.3 page 168 pour des explications concernant la construction et l'interprétation de ces figures.



Diagrammes de correspondance pour le noeud OI
 OI1 (32%) OI2 (68%)

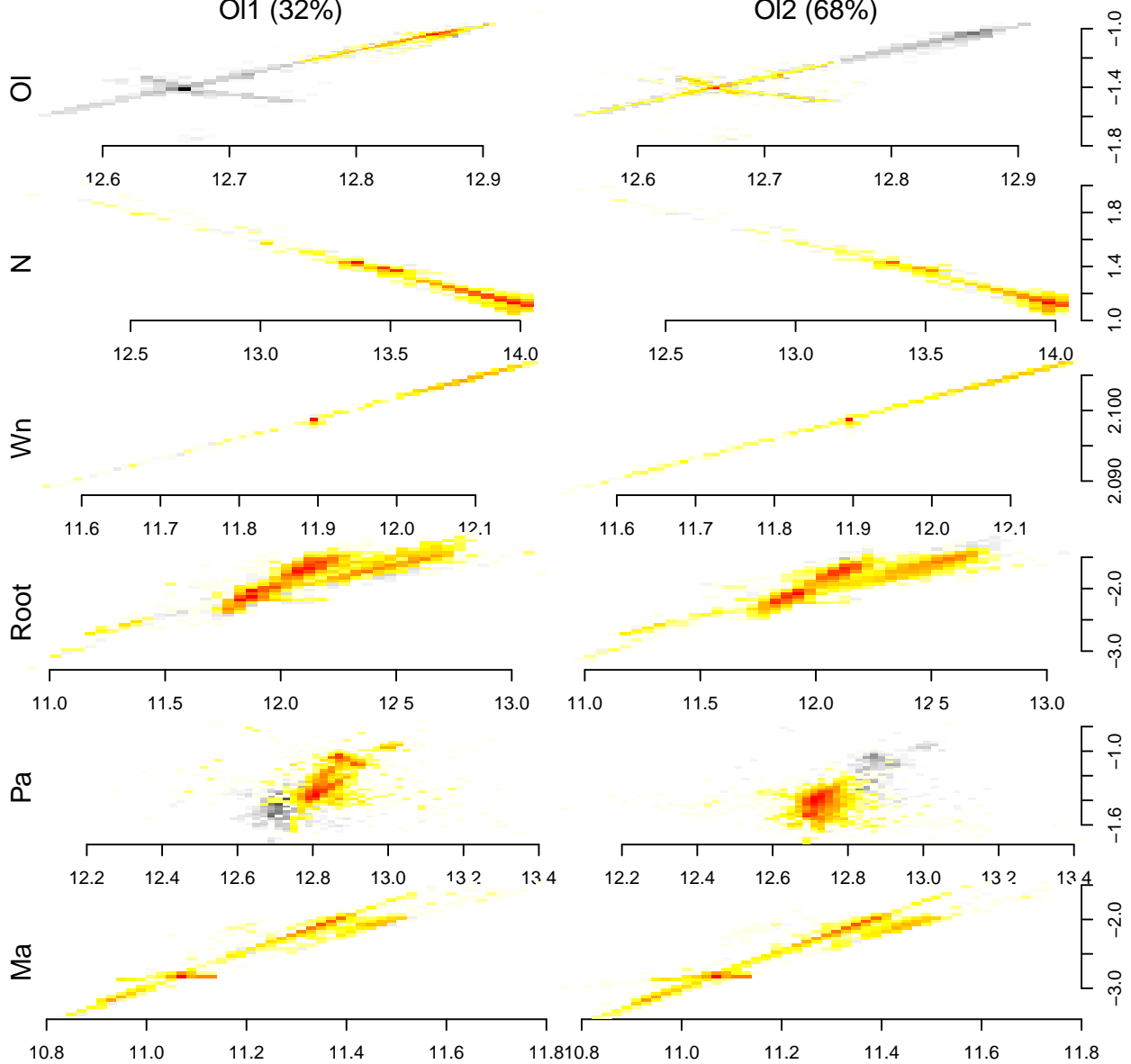


FIG. B.23 : Dépendance à la position des autres nœuds pour le nœud OI. Voir partie 5.3.3 page 168 pour des explications concernant la construction et l'interprétation de ces figures.

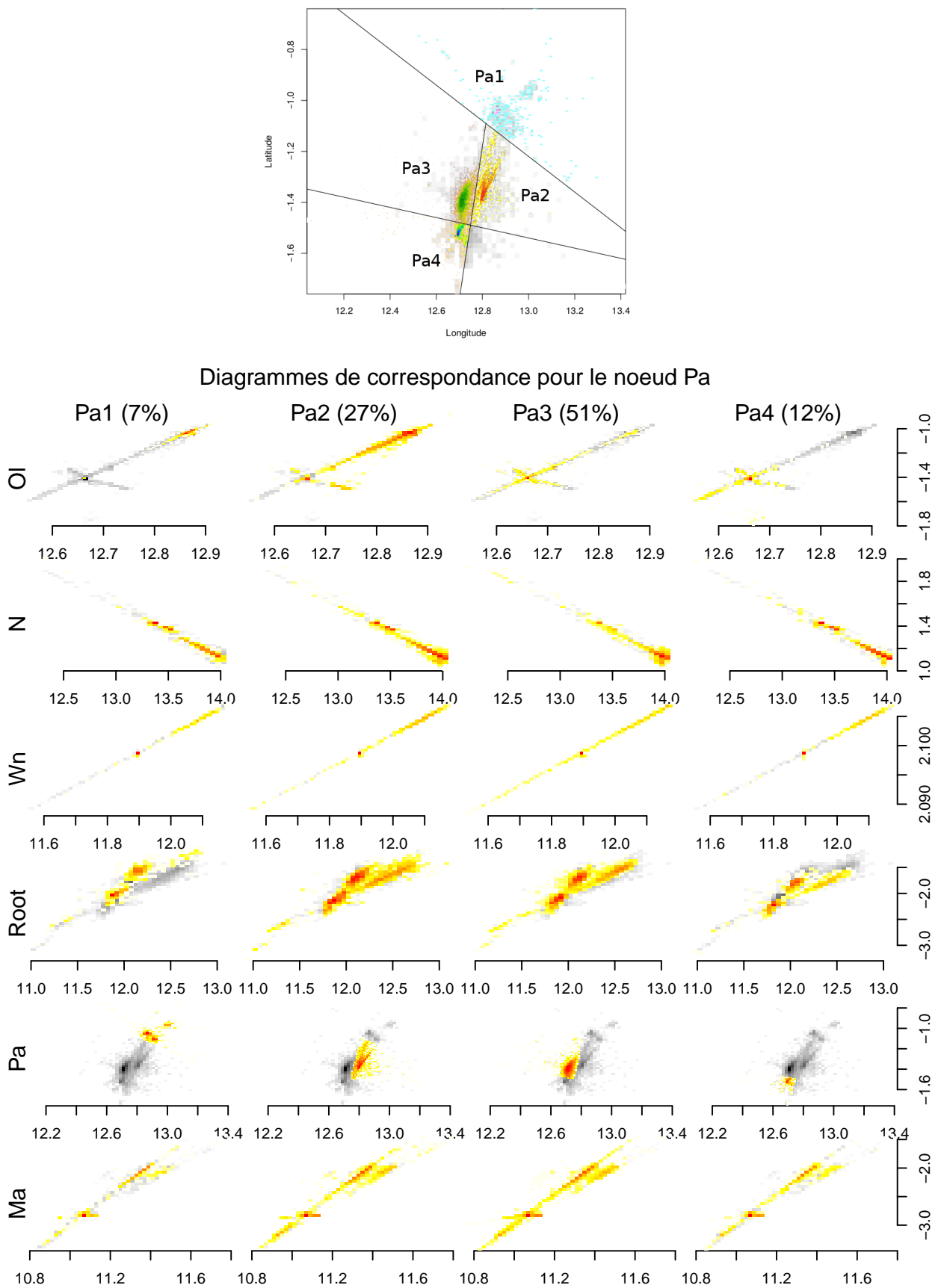


FIG. B.24 : Dépendance à la position des autres nœuds pour le nœud Pa. Voir partie 5.3.3 page 168 pour des explications concernant la construction et l'interprétation de ces figures.

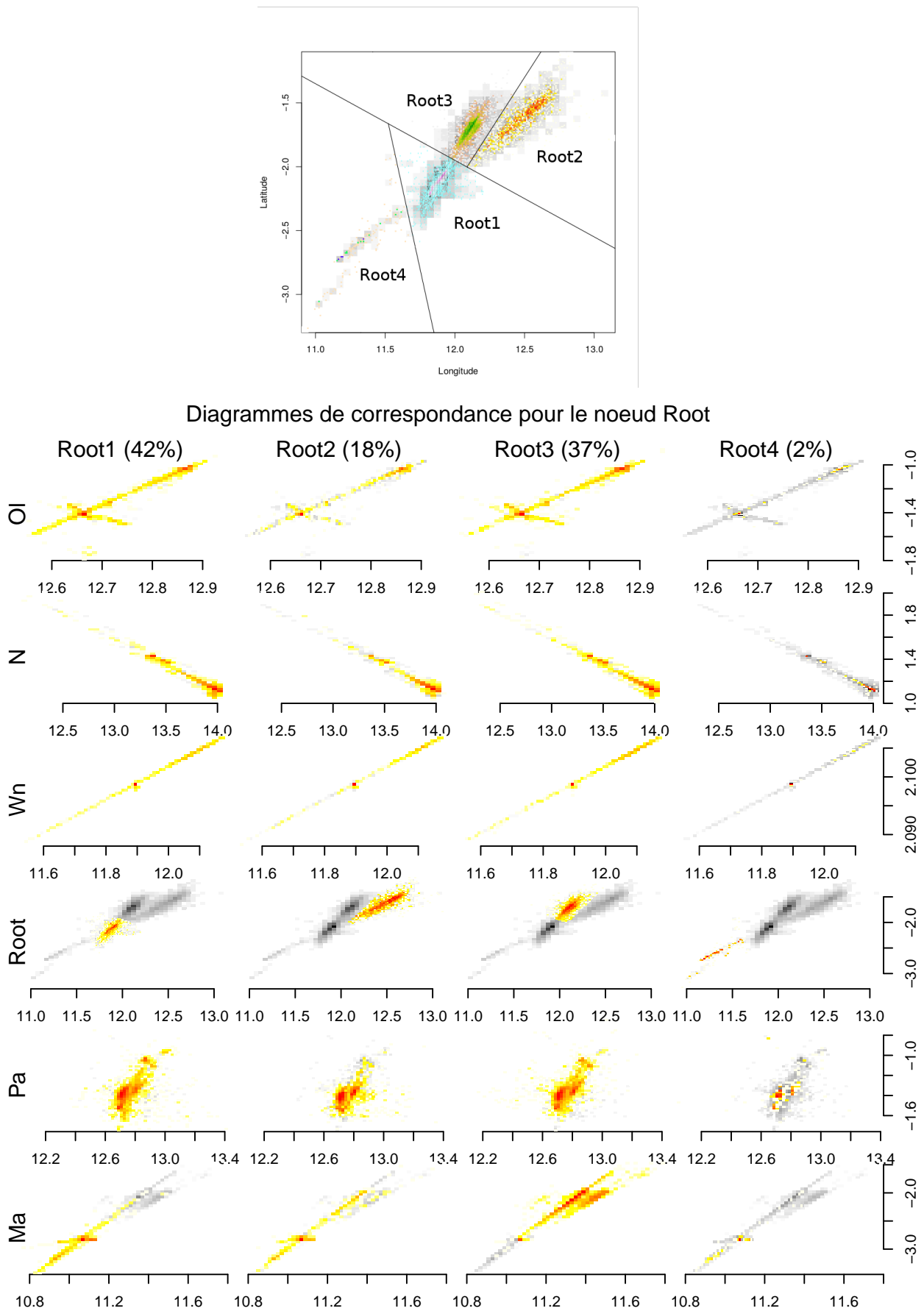


FIG. B.25 : Dépendance à la position des autres nœuds pour le nœud Root. Voir partie 5.3.3 page 168 pour des explications concernant la construction et l'interprétation de ces figures.

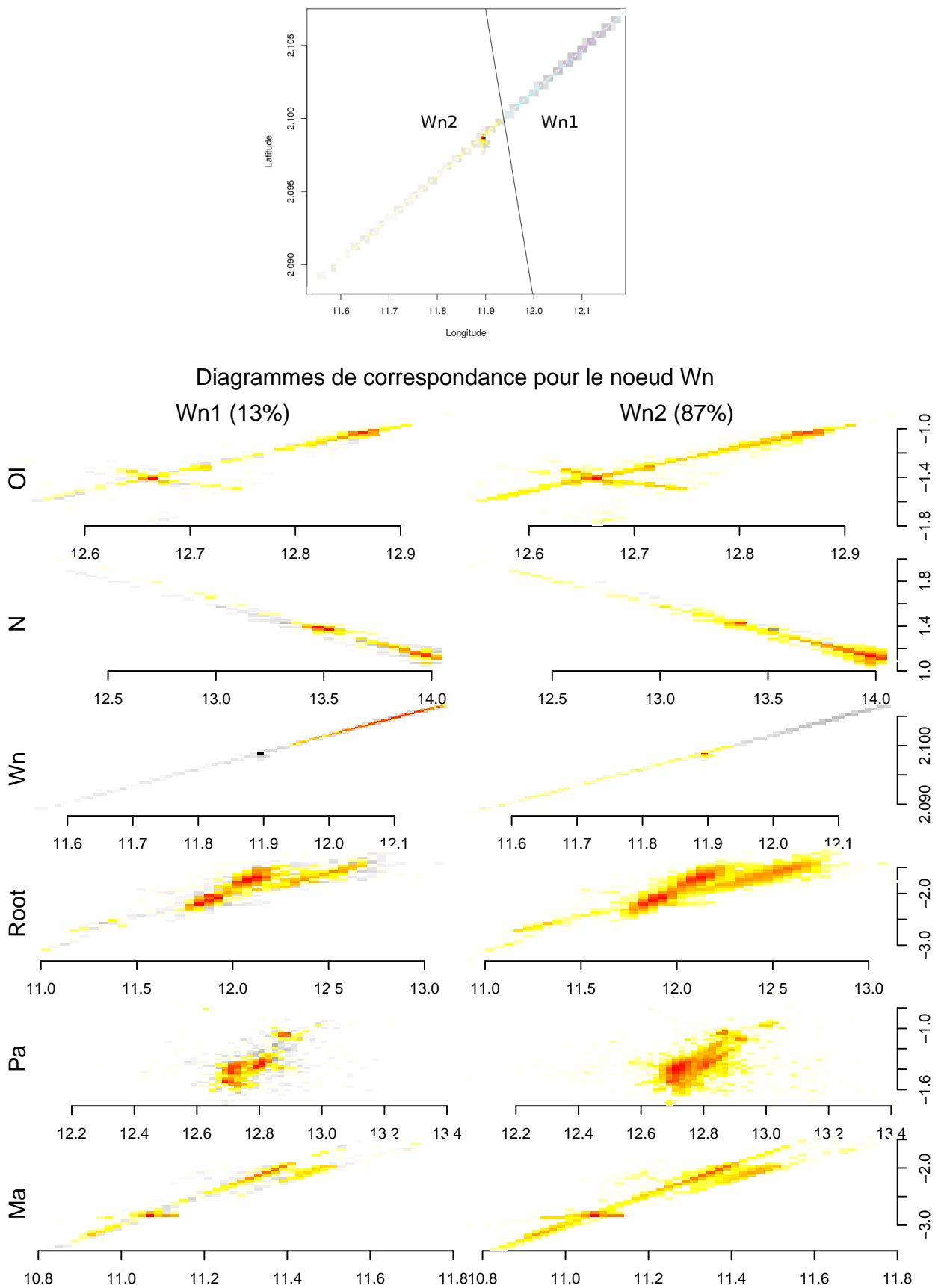


FIG. B.26 : Dépendance à la position des autres nœuds pour le nœud Wn. Voir partie 5.3.3 page 168 pour des explications concernant la construction et l'interprétation de ces figures.

B.3 Test de Mantel

B.3.1 Le test de Mantel

Le test de Mantel est un test statistique permettant d'évaluer l'existence d'une corrélation entre deux matrices de tailles identiques, et représentant typiquement deux mesures de distances différentes pour un même ensemble d'objets. Il est par exemple fréquemment utilisé en génétique des populations pour tester la corrélation entre la distance géographique et la distance génétique au sein d'une même population. Les mesures de distances présentes dans chacune des matrices n'étant pas indépendantes entre elles, les tests de corrélation habituels supposant une indépendance des données ne sont pas applicables. Le test de Mantel permet de traiter ce cas.

Le test de Mantel considère l'hypothèse nulle (H_0) selon laquelle les deux matrices ne sont pas corrélées. Une statistique t est ensuite calculée à partir des deux matrices. Les données présentes dans les matrices sont ensuite permutées un grand nombre de fois, et la comparaison de la statistique observée et des statistiques obtenues après permutation permet d'évaluer si l'hypothèse nulle peut être rejetée.

La statistique t . Dans son article originel, Mantel [279] utilise comme statistique du test la somme des produits des distances correspondantes dans chaque matrice [255, p. 832]. Ainsi, si a et b sont deux matrices de distances, où a_{ij} et b_{ij} représentent deux mesures de distances différentes entre les deux mêmes éléments i et j :

$$t_M = \sum_{j>i} a_{ij} \times b_{ij} \quad (\text{B.1})$$

Aujourd'hui, c'est une statistique de corrélation calculée à partir des données centrées et réduites qui est généralement utilisée [255], permettant d'obtenir une statistique comprise entre -1 et 1, sans influencer sur le résultat du test. Ainsi, en notant n le nombre d'éléments dans chaque matrice de distance⁹, \bar{a} et \bar{b} la moyenne de l'ensemble des valeurs respectivement des matrices a et b et σ_a et σ_b les écarts types des valeurs respectivement des matrices a et b , la statistique vaut alors :

$$r_M = \frac{1}{n-1} \sum_{j>i} \frac{a_{ij} - \bar{a}}{\sigma_a} \times \frac{b_{ij} - \bar{b}}{\sigma_b} \quad (\text{B.2})$$

Le test statistique. La procédure de permutation permet de tester la significativité de la corrélation alors que les données de distance ne sont pas indépendantes. À chaque permutation, les colonnes et les lignes de l'une des matrices sont réarrangées aléatoirement, et la statistique r_M est calculée. La procédure de permutation est répétée un grand nombre de fois (par exemple 1000 fois si le taux d'erreur accepté est de 5% [425]).

9. [↑] Si les matrices sont semi-triangulaires, et en notant m le nombre d'objets comparés deux à deux, on a : $n = \frac{m(m-1)}{2}$

On regarde alors combien de fois on obtient dans les données permutées une valeur égale ou supérieure à la statistique calculée sur les données de départ. Si le pourcentage d'obtention d'une valeur égale ou supérieur est inférieur au risque d'erreur que l'on s'est fixé au départ (par exemple 5%), alors l'hypothèse nulle est rejetée, et l'on admet l'existence de la corrélation entre les deux matrices, avec le risque d'erreur choisi.

B.3.2 Dérivés du test de Mantel

Test partiel de Mantel. Ce test, introduit par Smouse et coll. [385] en 1986, cherche à évaluer la capacité relative de deux distances X_1 et X_2 à expliquer une troisième distance observée Y . Elle permet donc d'estimer l'importance de la corrélation entre Y et X_1 (respectivement X_2) en retirant les effets de corrélation dus à X_2 (respectivement X_1). Smouse et coll. [385] montrent par exemple que si les distances géographique (X_1) et linguistique (X_2) observées dans des populations Yanomama d'Amérique du Sud sont toutes deux corrélées à la distance génétique (Y), elles sont aussi fortement corrélées entre elles, et la corrélation entre distances linguistique et génétique n'est plus significative lorsque les effets de la corrélation entre distances géographique et linguistique ont été pris en compte.

Corrélogramme de Mantel. [323, pp. 615–616] décrivent en 1986 cette méthode permettant de tester l'autocorrélation spatiale de données de distances observées (Y) dans une population. Les distances géographiques sont divisées en plusieurs classes, et pour chaque classe une matrice binaire X_i est construite : la valeur pour cette classe pour un couple d'objets est de 1 si leur distance géographique est incluse dans la classe, et 0 sinon. Pour chacune de ces matrices, un test est réalisé entre la matrice de classe de distance géographique (X_i) et la matrice de distance testée (Y). On peut alors savoir pour chaque classe de distance si les matrices sont significativement corrélées, permettant de détecter une autocorrélation spatiale des données à différentes échelles [256, pp. 819–821].

B.3.3 Critiques

Généralité. Le test de Mantel est fréquemment utilisé pour tester la corrélation statistique entre deux matrices quelconques et sans conditions sur les données. Cependant la généralité de ce test a été mise en doute. Il ne devrait être utilisé que lorsque les paramètres mesurés ne peuvent s'exprimer qu'en terme de distance [255]. De plus il n'est applicable que sous certaines conditions [180], et requiert en particulier l'absence d'auto-corrélation spatiale dans les données utilisées.

Performances. Le test de Mantel et le test partiel de Mantel ont été testés sur des données issues de simulations, et les résultats ont été plusieurs fois peu concluant. En général, ces études

révèlent un excès d'erreurs de type I¹⁰ et une faible puissance du test¹¹ [180, p. 4].

Procédure de permutation. La procédure de permutation utilisée aussi bien pour le test simple que le test partiel de Mantel est controversée, et serait responsable de l'excès d'erreurs de type I [180]. Une synthèse de la controverse sur les procédures de permutation dans les tests de Mantel est présentée dans l'article de Legendre et Fortin [255, annexe 4].

10. ↑ Les erreurs de type I sont celles correspondant au rejet de l'hypothèse nulle (revenant ici à accepter la corrélation entre les matrices) alors que celle-ci est vraie.

11. ↑ La puissance $(1 - \beta)$ d'un test statistique correspond à sa capacité à rejeter l'hypothèse nulle lorsque celle-ci est effectivement fausse.

Annexe C

Bibliographie

J'ai choisi, comme expliqué en détail dans la partie A.2, de séparer dans deux bibliographies distinctes les travaux que j'ai pu consulter directement et ceux dont j'ai pris connaissance par l'intermédiaire d'autres sources. Ceci me permet à la fois de clarifier l'origine des informations rapportées et de valoriser aussi bien les publications à la source de ces informations que les travaux de médiation m'ayant permis d'en prendre connaissance. Les études que j'ai pu lire sont référencées dans la partie *Publications consultées* (pages 425 et suivantes) et leur référence dans le texte est un simple numéro (par exemple [12]). Celles que je n'ai pas consultées directement sont référencées dans la partie *Autres publications citées* (pages 445 et suivantes) et le numéro de leur référence dans le texte est précédé d'un « A » (par exemple [A13]).

C.1 Publications consultées

- [3] ALVAREZ-PEREYRE, Frank et AROM, Simha. « La problématique étique/émique et l'ethnomusicologie ». In : *La boîte à outils d'un ethnomusicologue*. Textes réunis et présentés par Nathalie Fernando. Les Presses de l'université de Montréal, 2007, pp. 77–105 (citée pp. 109, 110).
- [7] APPLE, James B. « Fragments and Phylogeny of the Tibetan Version of the Mañjuśrī-vihārasūtra : A Case Study in the Genealogy of Tibetan Kanjurs ». In : *Annual Report of the International Research Institute for Advanced Buddhology (ARIRIAB) at Soka University for the Academic Year 2013* 17 (2014), pp. 293–335 (citée pp. 55, 60, 93).
- [9] ARCHIE, James W. « Homoplasy Excess Ratios : New Indices for Measuring Levels of Homoplasy in Phylogenetic Systematics and a Critique of the Consistency Index ». In : *Systematic Biology* 38.3 (1989), pp. 253–269 (citée p. 38).
- [11] AROM, Simha. « The Use of Play-Back Techniques in the Study of Oral Polyphonies ». In : *Ethnomusicology* 20.3 (1976), pp. 483–519 (citée pp. 113, 359).
- [12] AROM, Simha. « Structuration du temps dans les musiques d'Afrique centrale : périodicité, mètre, rythmique et polyrythmie ». In : *Revue de Musicologie* 70.1 (1984), pp. 5–36 (citée p. 373).

- [13] AROM, Simha. *Polyphonies et polyrythmies instrumentales d'Afrique Centrale : Structure et méthodologie*. Société d'Études Linguistiques et Anthropologiques de France (SELAF), 1985 (citée pp. 108, 112, 113, 119, 120, 124, 298, 337, 338, 341, 358, 359, 378–380).
- [14] AROM, Simha. « Réalisations, variations, modèles dans les musiques traditionnelles centrafricaines ». In : *L'improvisation dans les musiques de tradition orale*. Sous la dir. de LORTAT-JACOB, Bernard. T. 4. Ethnomusicologie. Société d'Études Linguistiques et Anthropologiques de France (SELAF), 1987, pp. 119–122 (citée pp. 112, 120, 124, 125, 378, 380).
- [15] AROM, Simha. « Modélisation et modèles dans les musiques de tradition orale ». In : *Analyse musicale* 22 (1991), pp. 67–78 (citée pp. 111, 112, 119–122, 350, 365, 378, 380, 381).
- [16] AROM, Simha. *La boîte à outils d'un ethnomusicologue*. Textes réunis et présentés par Nathalie Fernando. Les Presses de l'université de Montréal, 2007 (citée pp. 336, 354, 358, 369).
- [17] AROM, Simha. « L'organisation du temps musical : essai de typologie ». In : *Musiques. Une encyclopédie pour XXI^e siècle*. Sous la dir. de NATTIEZ, Jean-Jacques. Actes Sud / Cité de la Musique, 2007, pp. 928–944 (citée pp. 350, 361).
- [18] AROM, Simha. « Nouvelles perspectives dans la description des musiques de tradition orale ». In : *La boîte à outils d'un ethnomusicologue*. Textes réunis et présentés par Nathalie Fernando. Les Presses de l'université de Montréal, 2007, pp. 59–75 (citée p. 120).
- [19] AROM, Simha. « Structuration du temps dans les musiques d'Afrique centrale : périodicité, mètre, rythmique et polyrythmie ». In : *La boîte à outils d'un ethnomusicologue*. Textes réunis et présentés par Nathalie Fernando. Les Presses de l'université de Montréal, 2007, pp. 285–318 (citée p. 341).
- [20] AROM, Simha et ALVAREZ-PEREYRE, Frank. *Précis d'ethnomusicologie*. CNRS Éditions, 2007 (citée pp. 124, 351, 364, 365, 378, 381).
- [21] AROM, Simha, FERNANDO, Nathalie, FÜRNISS, Susanne, LE BOMIN, Sylvie, MARANDOLA, Fabrice et MOLINO, Jean. « La catégorisation des patrimoines musicaux dans les sociétés de tradition orale ». In : *Catégories et catégorisations : une perspective interdisciplinaire*. Sous la dir. d'ALVAREZ-PEREYRE, Frank. Peeters, 2008, pp. 273–313 (citée pp. 113, 118, 249, 332, 333, 353, 357, 358, 360, 378, 381).
- [22] AROM, Simha, FERNANDO, Nathalie, FÜRNISS, Susanne, LE BOMIN, Sylvie, MARANDOLA, Fabrice, OLIVIER, Emmanuelle, RIVIÈRE, Hervé et TOURNY, Olivier. « Typologie des techniques polyphoniques ». In : *Musiques. Une encyclopédie pour XXI^e siècle*. Sous la dir. de NATTIEZ, Jean-Jacques. Actes Sud / Cité de la Musique, 2007, pp. 1088–1109 (citée pp. 350, 358).
- [23] AROM, Simha, FERNANDO, Nathalie et MARANDOLA, Fabrice. « Une méthode innovante pour l'étude des échelles musicales africaines : aspects cognitifs et techniques ». In : *La boîte à outils d'un ethnomusicologue*. Textes réunis et présentés par Nathalie Fernando. Les Presses de l'université de Montréal, 2007, pp. 379–403 (citée p. 120).
- [24] AROM, Simha et MARTIN, Denis-Constant. *L'enquête en ethnomusicologie : préparation, terrain, analyse*. VRIN, 2015 (citée p. 331).

- [25] ATKINSON, Quentin D. et GRAY, Russel D. « Are Accurate Dates an Intractable Problem for Historical Linguistics ? » In : *Mapping Our Ancestors : Phylogenetic Approaches in Anthropology and Prehistory*. Sous la dir. de LIPO, Carl P., O'BRIEN, Michael J., COLLARD, Mark et SHENNAN, Stephen J. Transaction Publishers, 2006. Chap. 16, pp. 269–296 (citée pp. 64, 78).
- [26] ATKINSON, Quentin Douglas. « From Species to Languages : a Phylogenetic Approach to Human Prehistory ». Thèse de doctorat. University of Auckland, 2006 (citée pp. 64, 93, 251).
- [27] ATKINSON, Quentin et GRAY, Russell. « Curious Parallels and Curious Connections — Phylogenetic Thinking in Biology and Historical Linguistics ». In : *Systematic Biology* 54.4 (2005), pp. 513–526 (citée pp. 18, 22–25, 341, 346, 375).
- [28] ATKINSON, Quentin, NICHOLLS, Geoff, WELCH, David et GRAY, Russell. « From Words to Dates : Water into Wine, Mathemagic or Phylogenetic Inference ? » In : *Transactions of the Philological Society* 103.2 (2005), pp. 193–219 (citée pp. 64, 78, 81, 87, 93, 340).
- [33] BARBROOK, Adrian C., HOWE, Christopher J., BLAKE, Norman et ROBINSON, Peter. « The Phylogeny of The Canterbury Tales ». In : *Nature* 394 (1998), pp. 839–839 (citée pp. 54, 91, 93, 356).
- [34] BARET, Philippe V., MACÉ, Caroline et ROBINSON, Peter. « Testing Methods on an Artificially Created Textual Tradition ». In : *The Evolution of Texts : Confronting Stemmatological and Genetical Methods : Proceedings of the International Workshop Held in Louvain-la Neuve on September 1–2, 2004*. T. 24. *Linguistica Computazionale*. Roma–Pisa : Istituti editoriali e poligrafici internazionali, 2006, pp. 255–283 (citée pp. 54, 60, 93, 368).
- [40] BEL, Bernard. « Acquisition et représentation de connaissances en musique ». Thèse de doctorat. Université d'Aix-Marseille III, 1990 (citée p. 124).
- [41] BEN HAMED, Mahé. « La linguistique historique, nouveau terrain d'expérimentation de la phylogénie ». In : *Les mondes darwiniens*. Sous la dir. d'HEAMS, Thomas, HUNEMAN, Philippe, LECOINTRE, Guillaume et SILBERSTEIN, Marc. Syllepse, 2009. Chap. 44, pp. 975–995 (citée pp. 23, 25).
- [43] BENNETT, Charles H., LI, Ming et MA, Bin. « Chain Letters and Evolutionary Histories ». In : *Scientific American* 288.6 (2003), pp. 76–81 (citée pp. 55, 93).
- [45] BLANCHARD, Nolwenn. « Identité culturelle et patrimoine immatériel : la collection sonore constituée par Herbert Pepper au Gabon (1954-1966) ». Thèse de doctorat. Université Lumière Lyon 2, 2011 (citée p. 281).
- [46] BLAND, Mark. « Francis Beaumont's Verse Letters to Ben Jonson and 'The Mermaid Club' ». In : *English Manuscript Studies 1100-1700* 12 (2005), pp. 139–179 (citée pp. 54, 93).
- [47] BOC, Alix. « Détection des transferts horizontaux de gènes : modèles et algorithmes appliqués à l'évolution des espèces et des langues ». Thèse de doctorat. Université de Québec à Montréal, 2012 (citée pp. 91, 93).
- [49] BORDALEJO, Barbara. « The Phylogeny of the Tale-Order in the Canterbury Tales ». Thèse de doctorat. New York University, 2003 (citée pp. 18, 19, 25, 54–57, 348, 349, 356).
- [51] BORNES VAROL, Marie-Christine, éd. *Chocs de langues et de cultures ? Un discours de la méthode*. Presses universitaires de Vincennes, 2011 (citée pp. 123, 378, 381).

- [52] BORNES VAROL, Marie-Christine et FÜRNISS, Susanne. « Une méthode pour l'étude des situations de contact ». In : *Chocs de langues et de cultures ? un discours de la méthode*. Sous la dir. de BORNES VAROL, Marie-Christine. Presses universitaires de Vincennes, 2011, pp. 421–492 (citée p. 123).
- [53] BORTOLINI, Eugenio. « Isolation by Distance, Demic Diffusion, and Cultural Transmission in the Genomic Era : A Case Study Based on Folktale Distribution over Eurasia and Africa ». 6th International Conference on DNA Polymorphisms in Human Populations, Paris. 2016 (citée p. 74).
- [54] BOUCKAERT, Remco, LEMEY, Philippe, DUNN, Michael, GREENHILL, Simon J., ALEKSEYENKO, Alexander V., DRUMMOND, Alexei J., GRAY, Russell D., SUCHARD, Marc A. et ATKINSON, Quentin D. « Mapping the Origins and Expansion of the Indo-European Language Family ». In : *Science* 337.6097 (2012), pp. 957–960 (citée pp. 65, 93).
- [55] BOURDIEU, Pierre. *Les usages sociaux de la science. Pour une sociologie clinique du champ scientifique*. Institut national de la recherche agronomique (INRA), 1997 (citée pp. 310, 311).
- [56] BOWERN, Claire. « Historical Linguistics in Australia : Trees, Networks and Their Implications ». In : *Philosophical Transactions of the Royal Society B : Biological Sciences* 365.1559 (2010), pp. 3845–3854 (citée p. 93).
- [57] BOWERN, Claire, EPPS, Patience, GRAY, Russell, HILL, Jane, HUNLEY, Keith, McCONVELL, Patrick et ZENTZ, Jason. « Does Lateral Transmission Obscure Inheritance in Hunter-Gatherer Languages ? » In : *PLOS ONE* 6.9 (2011), e25195 (citée p. 81).
- [61] BOYD, Robert, BORGERHOFF-MULDER, Monique, DURHAM, William H. et RICHERSON, Peter J. « Are Cultural Phylogenies Possible ». In : *Human By Nature : Between Biology and the Social Sciences*. Sous la dir. de WEINGART, Peter, MITCHELL, Sandra D., RICHERSON, Peter J. et MAASEN, Sabine. Mahwah, New Jersey : Lawrence Erlbaum Associates, 1997, pp. 355–386 (citée pp. 260, 261, 263, 266).
- [65] BROWN, Steven, SAVAGE, Patrick E., KO, Albert Min-Shan, STONEKING, Mark, KO, Ying-Chin, LOO, Jun-Hun et TREJAUT, Jean A. « Correlations in the Population Structure of Music, Genes and Language ». In : *Proceedings of the Royal Society B : Biological Sciences* 281.20132072 (2014), pp. 1–7 (citée pp. 252–258, 262).
- [66] BRYANT, David, FILIMON, Flavia et GRAY, Russell D. « Untangling our Past : Languages, Trees, Splits and Networks ». In : *The Evolution of Cultural Diversity : a Phylogenetic Approach*. Sous la dir. de MACE, Ruth, HOLDEN, Clare J. et SHENNAN, Stephen. Left Coast Press, 2005. Chap. 5, pp. 67–84 (citée p. 94).
- [67] BRYANT, David et MOULTON, Vincent. « NeighborNet : An Agglomerative Method for the Construction of Planar Phylogenetic Networks ». In : *Algorithms in Bioinformatics*. Springer, 2002, pp. 375–391 (citée pp. 80, 90, 360).
- [68] BRYANT, David et MOULTON, Vincent. « Neighbor-Net : An Agglomerative Method for the Construction of Phylogenetic Networks ». In : *Molecular Biology and Evolution* 21.2 (2004), pp. 255–265 (citée p. 360).
- [69] BUCHANAN, Briggs et COLLARD, Mark. « Investigating the Peopling of North America Through Cladistic Analyses of Early Paleoindian Projectile Points ». In : *Journal of Anthropological Archaeology* 26.3 (2007), pp. 366–393 (citée pp. 68, 69, 88, 94).

- [70] BUCHANAN, Briggs et COLLARD, Mark. « Phenetics, Cladistics, and the Search for the Alaskan Ancestors of the Paleoindians : a Reassessment of Relationships Among the Clovis, Nenana, and Denali Archaeological Complexes ». In : *Journal of Archaeological Science* 35.6 (2008), pp. 1683–1694 (citée pp. 66, 69, 94).
- [71] BUCKLEY, Christopher D. « Investigating Cultural Evolution Using Phylogenetic Analysis : The Origins and Descent of the Southeast Asian Tradition of Warp Ikat Weaving ». In : *PLOS ONE* 7.12 (2012), e52064 (citée pp. 70, 81, 94).
- [72] CABRERA, Juan J., DÍAZ-BÁÑEZ, Jose Miguel, ESCOBAR-BORREGO, Francisco J., GÓMEZ, Emilia, GOMEZ, Francisco et MORA, Joaquín. « Comparative Melodic Analysis of A Cappella Flamenco *Cantes* ». In : *Proceedings of the fourth Conference on Interdisciplinary Musicology (CIM08)*. Ministerio de Cultura, 2008 (citée pp. 71, 73, 94).
- [73] CALUDE, Andreea S. et VERKERK, Annemarie. « The Typology and Diachrony of Higher Numerals in Indo-European : a Phylogenetic Comparative Study ». In : *Journal of Language Evolution* 1.2 (2016), pp. 91–108 (citée p. 260).
- [75] CAMPBELL, Lyle. *Historical Linguistics : An Introduction*. Edinburgh University Press, 1999 (citée pp. 28, 335).
- [76] CAMPBELL, Lyle. *Languages and Genes in Collaboration : Some Practical Matters*. Paper for the Languages and Genes Conference, Santa Barbara, September 8–10. University of California, 2006 (citée p. 251).
- [77] CAMPBELL, Lyle et MIXCO, Mauricio J. *A Glossary of Historical Linguistics*. Edinburgh University Press, 2007 (citée pp. 337, 346, 368).
- [78] CARLSON, Stephen C. « The Origin(s) of the “Caesarean” Text ». In : *Annual Meeting of the Society of Biblical Literature, San Antonio*. T. 20. 2004 (citée pp. 55–57, 94).
- [79] CARPENTER, James M. « Cladistics of Cladists ». In : *Cladistics* 3.4 (1987), pp. 363–375 (citée pp. 74, 90, 94).
- [80] CARTLIDGE, Neil. « The Canterbury Tales and Cladistics ». In : *Neuphilologische Mitteilungen* 102.2 (2001), pp. 135–150 (citée p. 58).
- [83] CAVALLI-SFORZA, Luigi Luca, PIAZZA, Alberto, MENOZZI, Paolo et MOUNTAIN, Joanna. « Reconstruction of Human Evolution : Bringing Together Genetic, Archaeological, and Linguistic Data ». In : *Proceedings of the National Academy of Sciences (PNAS)* 85.16 (1988), pp. 6002–6006 (citée p. 251).
- [87] CHARLES-DOMINIQUE, Luc. « Ethnomusicologie et histoire : deux *artes memoriae* ». In : *Cahiers d’ethnomusicologie* 22 (2009), pp. 15–34. URL : <http://ethnomusicologie.revues.org/900> (citée pp. 138–140).
- [89] CHEMILLIER, Marc. « Représentations musicales et représentations mathématiques ». In : *L’Homme* 171–172.3 (2004), pp. 265–283 (citée p. 124).
- [92] COCHRANE, Ethan E. et LIPO, Carl P. « Phylogenetic Analyses of Lapita Decoration Do Not Support Branching Evolution or Regional Population Structure During Colonization of Remote Oceania ». In : *Philosophical Transactions of the Royal Society B : Biological Sciences* 365.1559 (2010), pp. 3889–3902 (citée pp. 69, 91, 94).
- [93] COLLARD, Mark et SHENNAN, Stephen J. « Processes of Culture Change in Prehistory : A Case Study from the European Neolithic ». In : *Archaeogenetics : DNA and the Population Prehistory of Europe*. Sous la dir. de RENFREW, Colin et BOYLE, K. McDonald Institute for Archaeological Research, 2000, pp. 89–97 (citée pp. 66, 67, 94, 339).

- [94] COLLARD, Mark, SHENNAN, Stephen J. et TEHRANI, Jamshid J. « Branching, Blending, and the Evolution of Cultural Similarities and Differences Among Human Populations ». In : *Evolution and Human Behavior* 27.3 (2006), pp. 169–184 (citée pp. 80, 262).
- [97] COWARD, Fiona, SHENNAN, Stephen, COLLEDGE, Sue, CONOLLY, James et COLLARD, Mark. « The Spread of Neolithic Plant Economies from the Near East to Northwest Europe : a Phylogenetic Analysis ». In : *Journal of Archaeological Science* 35.1 (2008), pp. 42–56 (citée pp. 74, 94).
- [98] COWDERY, James R. « A Fresh Look at the Concept of Tune Family ». In : *Ethnomusicology* 28.3 (1984), pp. 495–504 (citée pp. 137, 141).
- [100] CRISCI, Jorge V. « Phylogeography ». In : *Historical Biogeography : An Introduction*. Harvard University Press, 2003. Chap. 9, pp. 118–126 (citée pp. 40–45).
- [101] CRISCI, Jorge Victor, KATINAS, Liliana et POSADAS, Paula. *Historical Biogeography : An Introduction*. Harvard University Press, 2003 (citée p. 345).
- [102] CURRIE, Thomas E., GREENHILL, Simon J. et MACE, Ruth. « Is Horizontal Transmission Really a Problem for Phylogenetic Comparative Methods? A Simulation Study Using Continuous Cultural Traits ». In : *Philosophical Transactions of the Royal Society B : Biological Sciences* 365.1559 (2010), pp. 3903–3912 (citée p. 79).
- [103] CURRIE, Thomas E., MEADE, Andrew, GUILLON, Myrtille et MACE, Ruth. « Cultural Phylogeography of the Bantu Languages of Sub-Saharan Africa ». In : *Proceedings of the Royal Society B : Biological Sciences* 280.1762 (2013), p. 20130695 (citée pp. 65, 94, 157, 158).
- [106] DARLU, Pierre. « L'évolution des langues : les méthodes en question ? » In : *Biosystema* 27 (2009), pp. 59–76 (citée pp. 62, 63, 284).
- [107] DARLU, Pierre et TASSY, Pascal. *La reconstruction phylogénétique : concepts et méthodes*. Masson, 1993 (citée pp. 32–34, 36, 38, 88–90, 329, 330, 334, 342–345, 348, 352, 355, 356, 366).
- [108] DARWENT, John et O'BRIEN, Michael J. « Using Cladistics to Construct Lineages of Projectile Points from Northeastern Missouri ». In : *Mapping Our Ancestors : Phylogenetic Approaches in Anthropology and Prehistory*. Sous la dir. de LIPO, Carl P., O'BRIEN, Michael J., COLLARD, Mark et SHENNAN, Stephen J. Transaction Publishers, 2006. Chap. 12, pp. 185–208 (citée p. 94).
- [112] DAYRAT, Benoît. « Ancestor-Descendant Relationships and the Reconstruction of the Tree of Life ». In : *Paleobiology* 31.3 (2005), pp. 347–353 (citée p. 56).
- [114] DEHOUX, Vincent. *Chants à penser Gbaya (Centrafrique)*. Paris : Société d'Études Linguistiques et Anthropologiques de France (SELAF), 1986, p. 219 (citée p. 121).
- [115] DEHOUX, Vincent, FÜRNISS, Susanne, LE BOMIN, Sylvie, OLIVIER, Emmanuelle, RIVIÈRE, Hervé et VOISIN, Frédéric, eds. *Ndroje balendro. Musiques, terrains et disciplines. Textes offerts à Simha Arom*. Peeters, 1995.
- [116] DELSUC, Frédéric et DOUZERY, Emmanuel J. P. « Les méthodes probabilistes en phylogénie moléculaire. (1) Les modèles d'évolution des séquences et le maximum de vraisemblance ». In : *Biosystema* 22 (2004), pp. 59–74 (citée pp. 39, 40, 370).
- [117] DELSUC, Frédéric et DOUZERY, Emmanuel J. P. « Les méthodes probabilistes en phylogénie moléculaire. (2) L'approche bayésienne ». In : *Biosystema* 22 (2004), pp. 75–86 (citée pp. 39, 164).

- [120] D'HUY, Julien. « Polyphemus (Aa. Th. 1137) : A Phylogenetic Reconstruction of a Pre-historic Tale ». In : *Nouvelle Mythologie Comparée/New Comparative Mythology* 1.1 (2013), pp. 3–18 (citée pp. 74, 95).
- [121] D'HUY, Julien. « A Cosmic Hunt in the Berber Sky : A Phylogenetic Reconstruction of Palaeolithic Mythology ». In : *Les Cahiers de l'AARS* 16 (2013), pp. 93–106 (citée pp. 74, 88, 95, 352).
- [122] D'HUY, Julien. « A Phylogenetic Approach of Mythology and its Archaeological Consequences ». In : *Rock Art Research* 30.1 (2013), pp. 115–118 (citée pp. 74, 95).
- [123] D'HUY, Julien et LE QUELLEC, Jean-Loïc. « Comment reconstruire la préhistoire des mythes ? Applications d'outils phylogénétiques à une tradition orale ». In : *Apparenter la pensée ? Vers une phylogénie des concepts savants*. Sous la dir. de CHARBONNAT, Pascal, BEN HAMED, Mahé et LECOINTRE, Guillaume. Éditions Matériologiques, 2014. Chap. 3, pp. 145–185 (citée p. 95).
- [124] DÍAZ-BÁÑEZ, Miguel, FARIGU, Giovanna, GÓMEZ, Francisco, RAPPAPORT, David et TOUSSAINT, Godfried T. « El Compás Flamenco : A Phylogenetic Analysis ». In : *Proceedings of Bridges : Mathematical Connections in Art, Music and Science*. 2004, pp. 61–70 (citée pp. 71, 73, 95).
- [127] DUNN, Michael, TERRILL, Angela, REESINK, Ger, FOLEY, Robert A. et LEVINSON, Stephen C. « Structural Phylogenetics and the Reconstruction of Ancient Language History ». In : *Science* 309.5743 (2005), pp. 2072–2075 (citée pp. 62, 95).
- [128] EAGLETON, Catherine et SPENCER, Matthew. « Copying and Conflation in Geoffrey Chaucer's *Treatise on the Astrolabe* : A Stemmatic Analysis Using Phylogenetic Software ». In : *Studies in History and Philosophy of Science Part A* 37.2 (2006), pp. 237–268 (citée pp. 54, 59, 95).
- [129] EBACH, Malte C., MORRONE, Juan J. et WILLIAMS, David M. « A new Cladistics of Cladists ». In : *Biology and Philosophy* 23.1 (2008), pp. 153–156 (citée pp. 74, 75, 95).
- [131] EEROLA, Tuomas, JÄRVINEN, Topi, LOUHIVUORI, Jukka et TOIVIAINEN, Petri. « Statistical Features and Perceived Similarity of Folk Melodies ». In : *Music Perception* 18.3 (2001), pp. 275–296 (citée p. 123).
- [132] ERDEM, Esra, LIFSCHITZ, Vladimir, NAKHLEH, Luay et RINGE, Donald. « Reconstructing the Evolutionary History of Indo-European Languages Using Answer Set Programming ». In : *International Symposium on Practical Aspects of Declarative Languages*. Springer. 2003, pp. 160–176 (citée pp. 63, 89, 95).
- [134] FELD, Steven. « Linguistic Models in Ethnomusicology ». In : *Ethnomusicology* 18.2 (1974), pp. 197–217 (citée pp. 307, 308).
- [135] FELSENSTEIN, Joe. *PHYLIP (Phylogeny Inference Package)*. URL : <http://evolution.genetics.washington.edu/phylip/> (consultée le 24/07/2013) (citée p. 90).
- [136] FELSENSTEIN, Joseph. *Inferring Phylogenies*. Sunderland : Sinauer Associates, 2004 (citée pp. 21, 22, 32–34, 54, 88–91, 345, 357, 369, 375).
- [137] FERNANDO, Nathalie. « Typologie des échelles et des modes ». In : *Musiques. Une encyclopédie pour XXI^e siècle*. Sous la dir. de NATTIEZ, Jean-Jacques. Actes Sud / Cité de la Musique, 2007, pp. 945–977 (citée p. 337).
- [138] FERRAN, Hugo. « Offrandes et bénédictions. Une anthropologie musicale du culte des ancêtres chez les Maale d'Éthiopie ». Thèse de doctorat. École des hautes études en sciences sociales (EHESS), 2010 (citée pp. 121, 123, 124, 332, 378, 381–383).

- [141] FISLER, Marie et LECOINTRE, Guillaume. « Categorizing Ideas about Trees : A Tree of Trees ». In : *PLOS ONE* 8.8 (2013), e68814 (citée pp. 74, 86, 89, 95).
- [143] FORBER, Patrick et GRIFFITH, Eric. « Historical Reconstruction : Gaining Epistemic Access to the Deep Past ». In : *Philosophy and Theory in Biology* 3.20160405 (2011), e203 (citée p. 273).
- [144] FORSTER, Peter et RENFREW, Colin, éd. *Phylogenetic Methods and the Prehistory of Languages*. McDonald Institute for Archaeological Research, 2006.
- [145] FORSTER, Peter et TOTH, Alfred. « Toward a Phylogenetic Chronology of Ancient Gaulish, Celtic, and Indo-European ». In : *Proceedings of the National Academy of Sciences (PNAS)* 100.15 (2003), pp. 9079–9084 (citée pp. 62, 96).
- [146] FORSTER, Peter, TOTH, Alfred et BANDELT, Hans-Jürgen. « Evolutionary Network Analysis of Word Lists : Visualising the Relationships Between Alpine Romance Languages ». In : *Journal of Quantitative Linguistics* 5.3 (1998), pp. 174–187 (citée pp. 62, 89, 96).
- [147] FORTUNATO, Laura et MACE, Ruth. « Testing Functional Hypotheses About Cross-Cultural Variation : A Maximum-Likelihood Comparative Analysis of Indo-European Marriage Practices ». In : *Pattern and Process in Cultural Evolution*. Sous la dir. de SHENNAN, Stephen. University of California Press, 2009 (citée pp. 62, 96, 260).
- [150] FÜRNISS, Susanne. « Rigueur et liberté. La polyphonie vocale des pygmées Aka (Centrafrique) ». In : *Polyphonies de tradition orale : Histoire et traditions vivantes. Actes du colloque de Royaumont — 1990*. Sous la dir. d'HUGLO, Michel et PÉRÈS, Marcel. Éditions Créaphis, 1993, pp. 101–131 (citée pp. 121, 124, 378, 383).
- [151] FÜRNISS, Susanne. « Les musiques du Sud-Est Cameroun, reflets d'une histoire mouvementée ». In : *Journal des Africanistes* 84.2 (2014), pp. 8–46 (citée pp. 139, 142, 264, 265, 284–286, 292, 302, 332, 365).
- [152] FÜRNISS, Susanne. « La technique du jodel chez les Pygmées Aka (Centrafrique). Étude phonétique et acoustique ». In : *Cahiers de musiques traditionnelles* 4 (1991), pp. 167–187 (citée p. 345).
- [153] FÜRNISS, Susanne et OLIVIER, Emmanuelle. « Systématique musicale pygmée et bochimán : Deux conceptions africaines du contrepoint ». In : *Musurgia* 4.3 (1997), pp. 9–30 (citée pp. 114–118, 143).
- [154] FÜRNISS-YACOUBI, Susanne. *Approche interdisciplinaire des musiques pygmées*. Université de Paris X. Synthèse en vue de l'obtention de l'habilitation à diriger des recherches. 2007 (citée pp. 331, 332, 378, 383).
- [156] GARCÍA RIVERO, Daniel. « Darwinian Archaeology and Cultural Phylogenetics ». In : *Cultural Phylogenetics — Concepts and Applications in Archaeology*. Sous la dir. de STRAFFON, Larissa Mendoza. Interdisciplinary Evolution Research 4. Springer, 2016, pp. 43–72 (citée pp. 65, 66).
- [157] GARCÍA RIVERO, Daniel et O'BRIEN, Michael J. « Phylogenetic Analysis Shows That Neolithic Slate Plaques from the Southwestern Iberian Peninsula are Not Genealogical Recording Systems ». In : *PLOS ONE* 9.2 (2014), pp. 1–18 (citée p. 96).
- [159] GINGRAS, Yves. *Sociologie des sciences*. Paris : Presses universitaires de France, 2013 (citée p. 311).

- [160] GOEBEL, Ted, POWERS, Roger et BIGELOW, Nancy. « The Nenana Complex of Alaska and Clovis Origins ». In : *Clovis : Origins and Adaptations*. Sous la dir. de BONNICHSEN, Robson et TURNMIRE, Karen L. Oregon : Center for the Study of the First Americans, Oregon State University, Corvallis, 1991, pp. 49–79 (citée pp. 66, 69).
- [162] GRAÇA DA SILVA, Sara et TEHRANI, Jamshid J. « Comparative Phylogenetic Analyses Uncover the Ancient Roots of Indo-European Folktales ». In : *Royal Society Open Science* 3.1 (2016), p. 150645 (citée pp. 81, 261).
- [163] GRAUER, Victor. « “Echoes of Our Forgotten Ancestors” : Author’s Reply ». In : *The World of Music* 48.2 (2006), pp. 101–134 (citée p. 143).
- [164] GRAUER, Victor. « New Perspectives on the Kalahari Debate : A Tale of Two ‘Genomes’ ». In : *Before Farming* 2007.2 (2007), pp. 1–14 (citée pp. 252, 258).
- [165] GRAUER, Victor A. « Echoes of Our Forgotten Ancestors ». In : *The World of Music* 2 (2006), pp. 5–58 (citée pp. 116, 142, 143, 252, 258, 259, 265, 266).
- [166] GRAY, R. D., DRUMMOND, A. J. et GREENHILL, S. J. « Language Phylogenies Reveal Expansion Pulses and Pauses in Pacific Settlement ». In : *Science* 323.5913 (2009), pp. 479–483 (citée pp. 65, 96, 254).
- [167] GRAY, Russel D., BRYANT, David et GREENHILL, Simon J. « On the Shape and Fabric of Human History ». In : *Philosophical Transactions of the Royal Society B : Biological Sciences* 365.1559 (2010), pp. 3923–3933 (citée pp. 80, 96, 260, 263, 264).
- [168] GRAY, Russell D. et ATKINSON, Quentin D. « Language-Tree Divergence Times Support the Anatolian Theory of Indo-European Origin ». In : *Nature* 426.6965 (2003), pp. 435–439 (citée pp. 64, 96).
- [169] GRAY, Russell D. et JORDAN, Fiona M. « Language Trees Support the Express-Train Sequence of Austronesian Expansion ». In : *Nature* 405.6790 (2000), pp. 1052–1055 (citée pp. 62, 96).
- [172] GREENHILL, Simon J., CURRIE, Thomas E. et GRAY, Russell D. « Does Horizontal Transmission Invalidate Cultural Phylogenies? » In : *Proceedings of the Royal Society B : Biological Sciences* 276.1665 (2009), pp. 2299–2306 (citée pp. 76–78, 96).
- [174] GRIFFITH, John G. « A Taxonomic Study of the Manuscript Tradition of Juvenal ». In : *Museum Helveticum* 25.2 (1968), pp. 101–138 (citée pp. 53, 55, 59, 96).
- [175] GRIFFITH, John G. « Numerical Taxonomy and Some Primary Manuscripts of the Gospels ». In : *The Journal of Theological Studies* 20.2 (1969), pp. 389–406 (citée p. 54).
- [176] GRIFFITHS, Graham C. D. « On the Foundations of Biological Systematics ». In : *Acta Biotheoretica* 23.3-4 (1974), pp. 85–131 (citée p. 363).
- [179] GROLLEMUND, Rebecca, BRANFORD, Simon, BOSTOEN, Koen, MEADE, Andrew, VENDITTI, Chris et PAGEL, Mark. « Bantu Expansion Shows That Habitat Alters the Route and Pace of Human Dispersals ». In : *Proceedings of the National Academy of Sciences (PNAS)* 112.43 (2015), pp. 13296–13301 (citée pp. 65, 96).
- [180] GUILLOT, Gilles et ROUSSET, François. « Dismantling the Mantel Tests ». In : *Methods in Ecology and Evolution* 4.4 (2013), pp. 336–344 (citée pp. 422, 423).
- [181] HANNA, Ralph. « The Application of Thought to Textual Criticism in All Modes — With Apologies to A. E. Housman ». In : *Studies in Bibliography* 53 (2000), pp. 163–172 (citée p. 58).

- [182] HARMON, Marcel J., VANPOOL, Todd L., LEONARD, Robert D., VANPOOL, Christine S. et SALTER, Laura A. « Reconstructing the Flow of Information across Time and Space : A Phylogenetic Analysis of Ceramic Traditions from Prehispanic Western and Northern Mexico and the American Southwest ». In : *Mapping Our Ancestors : Phylogenetic Approaches in Anthropology and Prehistory*. Sous la dir. de LIPO, Carl P., O'BRIEN, Michael J., COLLARD, Mark et SHENNAN, Stephen J. Transaction Publishers, 2006. Chap. 13, pp. 209–229 (citée pp. 69, 96).
- [184] HARTIGAN, J. A. et WONG, M. A. « Algorithm AS 136 : A *K*-Means Clustering Algorithm ». In : *Journal of the Royal Statistical Society : Series C (Applied Statistics)* 28.1 (1979), pp. 100–108 (citée pp. 254, 255).
- [186] HEAMS, Thomas, HUNEMAN, Philippe, LECOINTRE, Guillaume et SILBERSTEIN, Marc, éd. *Les mondes darwiniens. L'évolution de l'évolution*. Syllepse, 2009.
- [187] HELMLINGER, Aurélie. « Mémoriser à plusieurs. Expérience sur l'effet du groupe dans les steelbands (Trinidad et Tobago). » In : *Annales Fyssen* 24 (2010), pp. 216–235. URL : http://halshs.archives-ouvertes.fr/docs/00/53/57/57/PDF/2010_FYSSEN_24_HELMLINGER.pdf (citée p. 123).
- [192] HICKERSON, Michael J., CARSTENS, Bryan Charles, CAVENDER-BARES, Jeannine, CRANDALL, Keith A., GRAHAM, Catherine H., JOHNSON, Jerry B., RISSLER, Leslie, VICTORIANO, Pedro F. et YODER, Anne D. « Phylogeography's Past, Present, and Future : 10 Years After Avise, 2000 ». In : *Molecular Phylogenetics and Evolution* 54.1 (2010), pp. 291–301 (citée p. 41).
- [193] HOENIGSWALD, Henry M. et WIENER, Linda F., éd. *Biological Metaphor and Cladistic Classification : An Interdisciplinary Perspective*. Philadelphia : University of Pennsylvania Press, 1987.
- [194] HOLDEN, Clare J. et GRAY, Russel D. « Rapid Radiation, Borrowing and Dialect Continua in the Bantu Languages ». In : *Phylogenetic Methods and the Prehistory of Languages*. Sous la dir. de FORSTER, Peter et RENFREW, Colin. McDonald Institute for Archaeological Research, 2006, pp. 19–31 (citée p. 97).
- [195] HOLDEN, Clare J., MEADE, Andrew et PAGEL, Mark. « Comparison of Maximum Parsimony and Bayesian Bantu Language Trees ». In : *The Evolution of Cultural Diversity : a Phylogenetic Approach*. Sous la dir. de MACE, Ruth, HOLDEN, Clare J. et SHENNAN, Stephen. Left Coast Press, 2005. Chap. 4, pp. 53–66 (citée pp. 63, 97).
- [196] HOLDEN, Clare Janaki. « Bantu Language Trees Reflect the Spread of Farming Across Sub-Saharan Africa : a Maximum-Parsimony Analysis ». In : *Proceedings of the Royal Society B : Biological Sciences* 269.1493 (2002), pp. 793–799 (citée pp. 62, 96, 348).
- [201] VON HORNBOSTEL, Erich Moritz. « Die Probleme der vergleichenden Musikwissenschaft / The Problems of Comparative Musicology ». In : *Hornbostel Opera Omnia*. Trad. par CAMPBELL, Richard. Martinus Nijhoff, 1975 [1905], pp. 247–270 (citée pp. 130, 140, 351).
- [203] HOWE, Christopher J., CONNOLLY, Ruth et WINDRAM, Heather F. « Responding to Criticisms of Phylogenetic Methods in Stemmatology ». In : *Studies in English Literature (SEL) 1500-1900* 52.1 (2012), pp. 51–67 (citée pp. 56–58).
- [204] HUGOT, Jean-Pierre. « Biogéographie historique. Étude de la diversité biologique dans l'espace et dans le temps ». In : *Biosystema* 20 (2007), pp. 27–39 (citée pp. 40, 41).

- [205] HURLES, Matthew E., MATISOO-SMITH, Elizabeth, GRAY, Russell D. et PENNY, David. « Untangling Oceanic Settlement : the Edge of the Knowable ». In : *Trends in Ecology & Evolution* 18.10 (2003), pp. 531–540 (citée p. 97).
- [207] HUSON, Daniel H. et BRYANT, David. *User Manual for SplitsTree4 V4.14.4*. <http://ab.inf.uni-tuebingen.de/data/software/splitstree4/download/manual.pdf>. 2016. (Consultée le 16/04/2017) (citée p. 72).
- [209] HUSON, Daniel H., RUPP, Regula et SCORNAVACCA, Celine. *Phylogenetic Networks. Concepts, Algorithms and Applications*. Cambridge University Press, 2011 (citée p. 54).
- [210] IRFAN, Mohammad Tanvir, AKHTARUZAMAN, Md., ISLAM, Md. Saiful et ALAM, Md. Mahabubul. « Mathematical Representation and Analysis of Rhythms from Various Regions ». In : *International Conference on Information and Communication Technology (ICICT'07)*. Institute of Electrical and Electronics Engineers (IEEE). 2007, pp. 13–16 (citée p. 71).
- [214] JORDAN, Fiona Marie. « A Comparative Phylogenetic Approach to Austronesian Cultural Evolution ». Thèse de doctorat. University College London, 2007 (citée pp. 63, 97).
- [215] JORDAN, Peter et MACE, Thomas. « Tracking Culture-Historical Lineages : Can “Descent with Modification” be Linked to “Association by Descent” ? » In : *Mapping Our Ancestors : Phylogenetic Approaches in Anthropology and Prehistory*. Sous la dir. de LIPO, Carl P., O'BRIEN, Michael J., COLLARD, Mark et SHENNAN, Stephen J. Transaction Publishers, 2006. Chap. 10, pp. 149–167 (citée pp. 67, 97, 260, 262, 264).
- [216] JORDAN, Peter et O'NEILL, Sean. « Untangling Cultural Inheritance : Language Diversity and Long-House Architecture on the Pacific Northwest Coast ». In : *Philosophical Transactions of the Royal Society B : Biological Sciences* 365.1559 (2010), pp. 3875–3888 (citée pp. 67, 97, 261, 366).
- [217] JORDAN, Peter et SHENNAN, Stephen J. « Cultural Transmission, Language, and Basketry Traditions Amongst the California Indians ». In : *Journal of Anthropological Archaeology* 22.1 (2003), pp. 42–74 (citée pp. 67, 79, 97, 260).
- [223] KIPPEN, James. « An Ethnomusicological Approach to the Analysis of Musical Cognition ». In : *Music Perception* 5.2 (1987), pp. 173–195 (citée p. 124).
- [225] KITCHEN, Andrew, EHRET, Christopher, ASSEFA, Shiferaw et MULLIGAN, Connie J. « Bayesian Phylogenetic Analysis of Semitic Languages Identifies an Early Bronze Age Origin of Semitic in the Near East ». In : *Proceedings of the Royal Society B : Biological Sciences* 276.1668 (2009), pp. 2703–2710 (citée pp. 65, 97).
- [229] KRUMHANSL, Carol L., TOIVANEN, Pekka, EEROLA, Tuomas, TOIVAINEN, Petri, JÄRVINEN, Topi et LOUHIVUORI, Jukka. « Cross-Cultural Music Cognition : Cognitive Methodology Applied to North Sami Yoiks ». In : *Cognition* 76.1 (2000), pp. 13–58 (citée p. 123).
- [232] LACOMBE, Claire. « Les archives sonores comme terrain d'investigations : L'exemple des xylophones fang ». In : *Journal des Africanistes* 84.2 (2014), pp. 80–105 (citée pp. 137, 139, 281, 282).
- [233] LANTIN, Anne-Catherine, BARET, Philippe V. et MACÉ, Caroline. « Phylogenetic Analysis of Gregory of Nazianzus' Homily 27 ». In : *Le poids des mots : Actes des 7^{èmes} journées internationales d'analyse statistique des données textuelles (JADT04)*. T. 2. 2004, pp. 700–707 (citée pp. 55, 61, 97).

- [234] LE BOMIN, Sylvie. « Le patrimoine musical des Banda Gbambiya (République Centrafricaine). Catégorisation — Systématique de la musique pour xylophones ». Thèse de doctorat. École des hautes études en sciences sociales (EHESS), 2000 (citée pp. 122, 270, 284, 378, 383, 384).
- [235] LE BOMIN, Sylvie. « Raison morphologique et langage musical. Musiques de xylophone en Afrique centrale ». In : *Cahiers de Musiques traditionnelles* 14 (2001), pp. 203–219. URL : <http://ethnomusicologie.revues.org/172> (citée p. 284).
- [236] LE BOMIN, Sylvie. *Musiques Bateke. Mpa Atege. Gabon*. Saint-Maur-des-Fossés : Sépia, 2004 (citée pp. 286, 292).
- [237] LE BOMIN, Sylvie. « Étudier une musique instrumentale par ses processus d'apprentissage ». In : *Revue de Musicologie* 90.2 (2004), pp. 179–192 (citée pp. 122, 123, 284, 378, 385).
- [238] LE BOMIN, Sylvie et BIKOMA, Florence. *Musiques Myènè. De Port-Gentil à Lambaréné. Gabon*. Saint-Maur-des-Fossés : Sépia, 2005 (citée pp. 113, 137, 138, 141, 270–272, 287, 290–292).
- [239] LE BOMIN, Sylvie, LECHAUX, Emeline et MIFUNE, Marie-France. « Ce que “faire ensemble” peut vouloir dire en musique ». In : *Cahiers d'ethnomusicologie* 21 (2008), pp. 175–204 (citée p. 267).
- [240] LE BOMIN, Sylvie, LECOINTRE, Guillaume et HEYER, Evelyne. « The Evolution of Musical Diversity : The Key Role of Vertical Transmission ». In : *PLOS ONE* 11.3 (2016), e0151570 (citée pp. 73, 75, 79, 80, 84, 97, 142, 157, 158, 160, 170, 173, 180–182, 184, 189, 266, 267, 272, 275, 284, 286, 288, 317, 318, 353, 354).
- [241] LE BOMIN, Sylvie et MBOT, Jean-Émile. « Identité musicale des Pygmées du Gabon et absence de territoire ». In : *Territoires musicaux mis en scène*. Sous la dir. de DESROCHES, Monique. Presses de l'université de Montréal, 2011, pp. 279–293 (citée p. 267).
- [242] LE BOMIN, Sylvie et MBOT, Jean-Émile. « The Musical Heritage of the Gabon Bongo : Heritage Under Influence ». In : *Before Farming. The Archaeology and Anthropology of Hunter-Gatherers* 1.3 (2012), pp. 1–19 (citée p. 267).
- [243] LE BOMIN, Sylvie et MBOT, Jean-Émile. « Sur les traces de l'histoire des Pygmées du Gabon : résultats de cinq ans de prospection ». In : *Journal des africanistes* 82.1–2 (2012), pp. 277–318 (citée pp. 108, 267).
- [244] LE GONIDEC, Marie-Barbara. « Pierre Bec : Les instruments de musique d'origine arabe. Sens et histoire de leurs désignations ». In : *Cahiers d'ethnomusicologie* 18 (2005), pp. 302–304. URL : <http://ethnomusicologie.revues.org/376> (citée p. 140).
- [245] LE QUELLEC, Jean-Loïc. « Aréologie, phénétique et art rupestre : l'exemple des théranthropes du Sahara central ». In : *Les Cahiers de l'AARS* 16 (2013), pp. 155–176 (citée pp. 70, 98).
- [248] LECHAUX, Émeline. *Inventaire détaillé des archives sonores de Pierre Sallée*. Centre de recherche en ethnomusicologie (CREM), 2009. URL : <http://archives.crem-cnrs.fr> (citée p. 281).
- [249] LECHAUX, Émeline. « Tisser le fil de la mémoire. Contribution à l'histoire des répertoires musicaux des cérémonies de *bwétè* chez les Mitsogo du Gabon ». Thèse de doctorat. Paris : École des hautes études en sciences sociales (EHESS), 2015 (citée pp. 113, 136, 139, 281, 292, 298, 337, 338).

- [250] LECOINTRE, Guillaume. « Filiation ». In : *Les mondes darwiniens*. Sous la dir. d'HEAMS, Thomas, HUNEMAN, Philippe, LECOINTRE, Guillaume et SILBERSTEIN, Marc. Syllepse, 2009. Chap. 8, pp. 175–223 (citée pp. 20, 56).
- [251] LECOINTRE, Guillaume et LE GUYADER, Hervé. *Classification phylogénétique du vivant*. 3^e éd. Belin, 2001 (citée pp. 56, 304, 329, 334–336, 343, 355, 356, 363, 364).
- [252] LEE, Arthur R. « Numerical Taxonomy Revisited : John Griffith, Cladistic Analysis and St. Augustine's Quaestiones in Heptateuchum ». In : *Studia Patristica* 20 (1989), pp. 24–32 (citée pp. 55, 59, 97).
- [253] LEE, S. et HASEGAWA, T. « Bayesian Phylogenetic Analysis Supports an Agricultural Origin of Japonic Languages ». In : *Proceedings of the Royal Society B : Biological Sciences* 278.1725 (2011), pp. 3662–3669 (citée pp. 65, 98).
- [254] LEE, Sean et HASEGAWA, Toshikazu. « Evolution of the Ainu Language in Space and Time ». In : *PLOS ONE* 8.4 (2013), e62243 (citée pp. 65, 98).
- [255] LEGENDRE, Pierre et FORTIN, Marie-Josée. « Comparison of the Mantel Test and Alternative Approaches for Detecting Complex Multivariate Relationships in the Spatial Analysis of Genetic Data ». In : *Molecular Ecology Resources* 10.5 (2010), pp. 831–844 (citée pp. 421–423).
- [256] LEGENDRE, Pierre et LEGENDRE, Louis. *Numerical Ecology*. 3^e éd. Elsevier, 2012 (citée p. 422).
- [258] LEMEY, Philippe, SALEMI, Marco et VANDAMME, Anne-Mieke, éd. *The Phylogenetic Handbook : A Practical Approach to Phylogenetic Analysis and Hypothesis Testing*. Cambridge University Press, 2009 (citée p. 366).
- [262] LEWIS, Oscar. « Comparisons in Cultural Anthropology ». In : *Yearbook of Anthropology*. University of Chicago Press, 1955, pp. 259–292 (citée p. 135).
- [263] LIPO, Carl P., O'BRIEN, Michael J., COLLARD, Mark et SHENNAN, Stephen J., éd. *Mapping Our Ancestors : Phylogenetic Approaches in Anthropology and Prehistory*. Transaction Publishers, 2006.
- [264] LIST, Johann-Mattis, NELSON-SATHI, Shijulal, GEISLER, Hans et MARTIN, William. « Networks of Lexical Borrowing and Lateral Gene Transfer in Language and Genome Evolution ». In : *BioEssays* 36.2 (2014), pp. 141–150 (citée pp. 63, 64, 79, 89, 98).
- [265] LOMAX, Allan. *Folk Song Style and Culture*. 4^e éd. Transaction Publishers, 2007 [1968] (citée pp. 116, 134, 135, 138, 254).
- [266] LOMOLINO, Mark V., RIDDLE, Brett R., WHITTAKER, Robert J. et BROWN, James H. *Biogeography*. Sinauer Associates, 2010 (citée p. 41).
- [267] LOVE, Harold. « The Work in Transmission and Its Recovery ». In : *Shakespeare Studies* 32 (2004), pp. 73–80 (citée p. 58).
- [268] LYCETT, Stephen J. « Why is There a Lack of Mode 3 Levallois Technologies in East Asia ? A Phylogenetic Test of the Movius–Schick Hypothesis ». In : *Journal of Anthropological Archaeology* 26.4 (2007), pp. 541–575 (citée pp. 68, 75, 98).
- [269] LYCETT, Stephen J. « Understanding Ancient Hominin Dispersals Using Artefactual Data : A Phylogeographic Analysis of Acheulean Handaxes ». In : *PLOS ONE* 4.10 (2009), e7404 (citée pp. 68, 98).
- [270] LYCETT, Stephen J. « Are Victoria West Cores “Proto-Levallois” ? A Phylogenetic Assessment ». In : *Journal of Human Evolution* 56.2 (2009), pp. 175–191 (citée pp. 68, 98).

- [271] LYCETT, Stephen J., COLLARD, Mark et MCGREW, William C. « Phylogenetic Analyses of Behavior Support Existence of Culture Among Wild Chimpanzees ». In : *Proceedings of the National Academy of Sciences (PNAS)* 104.45 (2007), pp. 17588–17592 (citée p. 79).
- [273] MAAS, Philipp A. « A Phylogenetic Approach to the Transmission of the Tibetan Kanjur — The Akṣayamatīrdeśa Revisited ». In : *Bauddhasāhityastabakāvalī : Essays and Studies on Buddhist Sanskrit Literature Dedicated to Claus Vogel by Colleagues, Students, and Friends*. T. 36. Marburg : Indica et Tibetica Verlag, 2008, pp. 229–243 (citée pp. 19, 59, 98).
- [274] MAAS, Philipp A. « Computer Aided Stemmatology – The Case of Fifty-Two Text Versions of Carakasamhitā Vimānasthāna 8.67-157 ». In : *Wiener Zeitschrift für die Kunde Südasiens* 52 (2010), pp. 67–157 (citée pp. 59, 98).
- [275] MACÉ, Caroline, DE VOS, Ilse et GEUTEN, Koen. « Comparing Stemmatological and Phylogenetic Methods to Understand the Transmission History of the *Florilegium Coislinianum* ». In : *Ars Edendi Lectures Series 2* (2012), pp. 107–129 (citée pp. 55–57, 59, 60, 90, 98).
- [276] MACÉ, Caroline, SCHMIDT, Thomas et WEILER, Jean-François. « Le classement des manuscrits par la statistique et la phylogénétique : les cas de Grégoire de Nazianze et de Basile le Minime ». In : *Revue d'histoire des textes* 31 (2001), pp. 241–273 (citée pp. 54, 55, 58, 59, 61, 98, 308, 346).
- [277] MACE, Ruth, HOLDEN, Clare J. et SHENNAN, Stephen, éd. *The Evolution of Cultural Diversity : a Phylogenetic Approach*. Left Coast Press, 2005.
- [279] MANTEL, Nathan. « The Detection of Disease Clustering and a Generalized Regression Approach ». In : *Cancer Research* 27.2 Part 1 (1967), pp. 209–220 (citée p. 421).
- [280] MARWICK, Ben. « A Cladistic Evaluation of Ancient Thai Bronze Buddha Images : Six Tests for a Phylogenetic Signal in the Griswold Collection ». In : *Connecting Empires*. Sous la dir. de BONATZ, D., REINECKE, A. et TJOA-BONATZ, M. L. Singapore : National University of Singapore Press, 2012, pp. 159–176 (citée pp. 70, 81, 88, 98, 343).
- [281] MAYAUD, Isabelle. « L'institutionnalisation de l'ethnomusicologie : les contours de la discipline en question (1950-1977) ». In : *MUSICultures* 37 (2010), pp. 32–49. URL : <https://journals.lib.unb.ca/index.php/MC/article/viewFile/20225/23327> (citée pp. 132, 133, 351).
- [282] MAYR, Ernst. « Typological Versus Population Thinking ». In : *Evolution and the Diversity of Life*. Cambridge : Harvard University Press, 1976, pp. 26–29 (citée p. 267).
- [284] MCCARTHY, Ian Paul. « Toward a Phylogenetic Reconstruction of Organizational Life ». In : *Journal of Bioeconomics* 7.3 (2005), pp. 271–307 (citée p. 75).
- [286] MCLEOD, Norma. « Ethnomusicological Research and Anthropology ». In : *Annual Review of Anthropology* 3.1 (1974), pp. 99–115 (citée p. 133).
- [287] MCMAHON, April et MCMAHON, Robert. *Language Classification by Numbers*. USA : Oxford University Press, 2005 (citée pp. 77, 87, 99).
- [290] MERRIAM, Alan P. *The Anthropology of Music*. Evanston : Northwestern University Press, 1964 (citée pp. 133, 134).
- [291] MERRIAM, Alan P. « Definitions of “Comparative Musicology” and “Ethnomusicology” : An Historical-Theoretical Perspective ». In : *Ethnomusicology* 21.2 (1977), pp. 189–204 (citée pp. 130, 131, 351).

- [292] MESOUDI, Alex. « Cultural Evolution : A Review of Theory, Findings and Controversies ». In : *Evolutionary Biology* (2015), pp. 1–17 (citée p. 251).
- [295] MIFUNE, Marie-France. « Performance et construction identitaire. Une approche interdisciplinaire du culte du bwiti chez les Fang du Gabon ». Thèse de doctorat. Paris : École des hautes études en sciences sociales (EHESS), 2012 (citée pp. 286, 378, 386).
- [296] MINELLI, Alessandro. *Biological Systematics : The State of the Art*. 2^e éd. London : Chapman & Hall, 1994 (citée pp. 363, 364).
- [298] MORA, Joaquín, GOMEZ MARTIN, Francisco, GÓMEZ, Emilia, ESCOBAR-BORREGO, Francisco J. et DÍAZ-BÁÑEZ, José Miguel. « Characterization and Melodic Similarity of a Cappella Flamenco Cantes ». In : *Proceedings of the 11th ISMIR Conference*. International Society for Music Information Retrieval (ISMIR). 2010 (citée pp. 71, 73, 99).
- [300] NAKHLEH, Luay, RINGE, Don et WARNOW, Tandy. « Perfect Phylogenetic Networks : A New Methodology for Reconstructing the Evolutionary History of Natural Languages ». In : *Language* 81.2 (2005), pp. 382–420 (citée pp. 63, 89, 99).
- [301] NAKHLEH, Luay, WARNOW, Tandy, RINGE, Don et EVANS, Steven N. « A Comparison of Phylogenetic Reconstruction Methods on an Indo-European Dataset ». In : *Transactions of the Philological Society* 103.2 (2005), pp. 171–192 (citée pp. 63, 99).
- [303] NATTIEZ, Jean-Jacques. « Arom, ou le sémiologue sans le savoir ». In : *Analyse musicale* 23 (1991), pp. 77–82 (citée pp. 123, 378, 385).
- [304] NATTIEZ, Jean-Jacques. « Inuit Throat-Games and Siberian Throat Singing : A Comparative, Historical, and Semiological Approach ». In : *Ethnomusicology* 43.3 (1999), pp. 399–418 (citée pp. 141, 269, 270).
- [305] NATTIEZ, Jean-Jacques, éd. *Musiques. Une encyclopédie pour XXI^e siècle*. Actes Sud / Cité de la Musique, 2007.
- [306] NELSON-SATHI, Shijulal, LIST, Johann-Mattis, GEISLER, Hans, FANGERAU, Heiner, GRAY, Russell D., MARTIN, William et DAGAN, Tal. « Networks Uncover Hidden Lexical Borrowing in Indo-European Language Evolution ». In : *Proceedings of the Royal Society B : Biological Sciences* 278 (2010), pp. 1794–1803 (citée pp. 63, 79, 99).
- [310] NETTL, Bruno. « Comparison and Comparative Method in Ethnomusicology ». In : *Anuario Interamericano de Investigacion Musical* 9 (1973), pp. 148–161 (citée pp. 130, 132–134, 136, 137).
- [311] NETTL, Bruno. *The Study of Ethnomusicology : Thirty-one issues and concepts*. 2^e éd. University of Illinois Press, 2005 (citée p. 134).
- [312] NETTL, Bruno. « Response to Victor Grauer : on the Concept of Evolution in the History of Ethnomusicology ». In : *The World of Music* 48.2 (2006), pp. 59–72 (citée pp. 252, 265).
- [313] NETTL, Bruno. *Nettl's Elephant. On the History of Ethnomusicology*. University of Illinois Press, 2010 (citée p. 134).
- [314] NETTL, Bruno. « Revisiting Comparison, Comparative Study and Comparative Musicology ». In : *Nettl's elephant : on the history of ethnomusicology*. University of Illinois Press, 2010. Chap. 5, pp. 70–89 (citée pp. 130–133, 137, 141, 351).
- [318] NORET, J. *Deux Vies de Saint Athanase l'Athonite*. Turnhout : Corpus Christianorum, 1982 (citée p. 346).

- [319] NUNN, Charles L. « Comparative Methods for Studying Cultural Trait Evolution : A Simulation Study ». In : *Cross-Cultural Research* 40.2 (2006), pp. 177–209 (citée pp. 79, 81).
- [320] NUNN, Charles L. *The Comparative Approach in Evolutionary Anthropology and Biology*. University Of Chicago Press, 2011 (citée p. 251).
- [321] NUNN, Charles L., ARNOLD, Christian, MATTHEWS, Luke et BORGERHOFF MULDER, Monique. « Simulating Trait Evolution for Cross-Cultural Comparison ». In : *Philosophical Transactions of the Royal Society B : Biological Sciences* 365.1559 (2010), pp. 3807–3819 (citée pp. 79, 80, 344).
- [322] O'BRIEN, Michael J., DARWENT, John et LYMAN, R. Lee. « Cladistics Is Useful for Reconstructing Archaeological Phylogenies : Palaeoindian Points from the Southeastern United States ». In : *Journal of Archaeological Science* 28.10 (2001), pp. 1115–1136 (citée p. 99).
- [323] ODEN, Neal L. et SOKAL, Robert R. « Directional Autocorrelation : An Extension of Spatial Correlograms to Two Dimensions ». In : *Systematic Biology* 35.4 (1986), pp. 608–617 (citée p. 422).
- [325] O'HARA, Robert J. « Trees of History in Systematics and Philology ». In : *Memorie della Società Italiana di Scienze Naturali e del Museo Civico di Storia Naturale di Milano* 27.1 (1996), pp. 81–88 (citée pp. 18, 25, 61).
- [326] O'HARA, Robert J et ROBINSON, Peter. « Computer-Assisted Methods of Stemmatic Analysis ». In : *Occasional Papers of the Canterbury Tales Project* 1.1993 (1993), pp. 53–74 (citée pp. 56–58).
- [330] PAGE, Roderic D. M. et CHARLESTON, Michael A. « Trees Within Trees : Phylogeny and Historical Associations ». In : *Trends in Ecology & Evolution* 13.9 (1998), pp. 356–359 (citée p. 263).
- [331] PAGEL, Mark, ATKINSON, Quentin D. et MEADE, Andrew. « Frequency of Word-Use Predicts Rates of Lexical Evolution Throughout Indo-European History ». In : *Nature* 449.7163 (2007), pp. 717–720 (citée p. 99).
- [332] PAGEL, Mark et MEADE, Andrew. « Estimating Rates of Lexical Replacement on Phylogenetic Trees of Languages ». In : *Phylogenetic Methods and the Prehistory of Languages*. Sous la dir. de FORSTER, Peter et RENFREW, Colin. McDonald Institute for Archaeological Research, 2006. Chap. 15, pp. 173–182 (citée pp. 64, 85, 99).
- [333] PAGEL, Mark et MEADE, Andrew. *BayesTraits V2.0 (Beta)*. 2013. URL : <http://www.evolution.rdg.ac.uk/BayesTraits.html> (consultée le 09/11/2016) (citée p. 158).
- [334] PAMJAV, Horolma, JUHÁSZ, Zoltán, ZALÁN, Andrea, NÉMETH, Endre et DAMDIN, Bayarlkhagva. « A Comparative Phylogenetic Study of Genetics and Folk Music ». In : *Molecular Genetics and Genomics* 287.4 (2012), pp. 337–349 (citée pp. 252, 253, 256, 258, 259).
- [339] PLATNICK, Norman I. et DON CAMERON, H. « Cladistic Methods in Textual, Linguistic, and Phylogenetic Analysis ». In : *Systematic Biology* 26.4 (1977), pp. 380–385 (citée pp. 18, 19, 21, 25, 55, 58, 61, 362).
- [340] PRENTISS, Anna Marie, SKELTON, Randall R., ELDREDGE, Niles et QUINN, Colin. « Get Rad! The Evolution of the Skateboard Deck ». In : *Evolution : Education and Outreach* 4.3 (2011), pp. 379–389 (citée pp. 67, 99).

- [341] PRENTISS, Anna Marie, WALSH, Matthew J., SKELTON, Randall R. et MATTES, Matt. « Mosaic Evolution in Cultural Frameworks : Skateboard Decks and Projectile Points ». In : *Cultural Phylogenetics — Concepts and Applications in Archaeology*. Sous la dir. de STRAFFON, Larissa Mendoza. Interdisciplinary Evolution Research 4. Springer, 2016, pp. 113–130 (citée p. 100).
- [342] DE QUEIROZ, Kevin. « Systematics and the Darwinian Revolution ». In : *Philosophy of Science* 55.2 (1988), pp. 238–259 (citée pp. 20, 363).
- [346] REXOVÁ, Kateřina, BASTIN, Yvonne et FRYNTA, Daniel. « Cladistic Analysis of Bantu Languages : A new Tree Based on Combined Lexical and Grammatical Data ». In : *Naturwissenschaften* 93.4 (2006), pp. 189–194 (citée pp. 63, 100, 263).
- [347] REXOVÁ, Kateřina, FRYNTA, Daniel et ZRZAVÝ, Jan. « Cladistic Analysis of Languages : Indo-European Classification Based on Lexicostatistical Data ». In : *Cladistics* 19.2 (2003), pp. 120–127 (citée pp. 62, 75, 100).
- [348] REY-DEBOVE, Josette et REY, Alain, éd. *Le nouveau Petit Robert de la langue française*. Le Robert, 2007 (citée p. 359).
- [349] RINGE, Don, WARNOW, Tandy et TAYLOR, Ann. « Indo-European and Computational Cladistics ». In : *Transactions of the Philological Society* 100 (2002), pp. 59–129 (citée pp. 63, 64, 89, 100).
- [350] ROBERT-DEMONTROND, Philippe, éd. *L'importation de concepts*. Editions Apogée, 2005 (citée pp. 308, 313).
- [353] ROBINS, William. « Editing and Evolution ». In : *Literature Compass* 4.1 (2007), pp. 89–120 (citée pp. 18–20, 53, 348, 375).
- [354] ROBINSON, Peter. « A Stemmatic Analysis of the Fifteenth-Century Witnesses to the Wife of Bath's Prologue ». In : *The Canterbury Tales Project : Occasional Papers, Vol. II*. London : Office for Humanities Communication Publications, 1997, pp. 69–132 (citée pp. 55, 57, 100).
- [355] ROBINSON, Peter M. W. « Computer-Assisted Stemmatic Analysis and 'Best-Text' Historical Editing ». In : *Studies in Stemmatology*. Sous la dir. de VAN REENEN, P. et VAN MULKEN, M. Amsterdam : John Benjamins Publishing Company, 1996, pp. 71–103 (citée pp. 55, 57).
- [356] ROBINSON, Peter M. W. et O'HARA, Robert J. « Report on the Textual Criticism Challenge 1991 ». In : *Bryn Mawr Classical Review* 3.4 (1992), pp. 331–337 (citée pp. 55, 57, 59, 61).
- [357] ROBINSON, Peter M. W. et O'HARA, Robert J. « Cladistic Analysis of an Old Norse Manuscript Tradition ». In : *Research in Humanities Computing* 4 (1996), pp. 115–137 (citée pp. 55–57, 59, 61, 100).
- [358] ROGERS, Deborah S., FELDMAN, Marcus W. et EHRLICH, Paul R. « Inferring Population Histories Using Cultural Data ». In : *Proceedings of the Royal Society B : Biological Sciences* 276.1674 (2009), pp. 3835–3843 (citée pp. 69, 81, 100, 264).
- [360] RONQUIST, Fredrik. « Dispersal-Vicariance Analysis : A new Approach to the Quantification of Historical Biogeography ». In : *Systematic Biology* 46.1 (1997), pp. 195–203 (citée pp. 42, 46).
- [361] ROOS, Teemu et HEIKKILÄ, Tuomas. « Evaluating Methods for Computer-Assisted Stemmatology Using Artificial Benchmark Data Sets ». In : *Literary and Linguistic Computing* 24.4 (2009), pp. 417–433 (citée pp. 54, 59–61, 75, 87, 90, 91, 100, 368).

- [363] RZESZUTEK, Tom, SAVAGE, Patrick E. et BROWN, Steven. « The Structure of Cross-Cultural Musical Diversity ». In : *Proceedings of the Royal Society B : Biological Sciences* 279.1733 (2012), pp. 1606–1612 (citée pp. 138, 254).
- [366] SALEMANS, Ben J. « Building Stemmas with the Computer in a Cladistic, Neo-lachmannian, Way : The Case of Fourteen Text Versions of Lanseloet Van Denemerken ». Thèse de doctorat. Katholieke Universiteit Nijmegen, 2000 (citée pp. 19, 20, 25, 29–31, 55, 100, 335, 336, 348, 362).
- [367] SALEMANS, Ben J. P. « Cladistics or the Resurrection of the Method of Lachmann : On Building the Stemma of *Yvain* ». In : *Studies in Stemmatology*. Sous la dir. de REENEN, Pieter van et MULKEN, Margot J. P. van. John Benjamins Publishing Company, 1996, pp. 3–70 (citée pp. 55, 100).
- [368] SAPORTA, Gilbert. *Probabilités, analyse des données et statistiques*. TECHNIP, 2011 (citée pp. 165, 256, 257, 267).
- [369] SAVAGE, Patrick E. et BROWN, Steven. « Toward a New Comparative Musicology ». In : *Analytical Approaches to World Music* 2 (2013), pp. 148–198 (citée pp. 132, 252, 265).
- [370] SAVAGE, Patrick E., MATSUMAE, Hiromi, OOTA, Hiroki, STONEKING, Mark, CURRIE, Thomas E., TAJIMA, Atsushi, GILLAN, Matt et BROWN, Steven. « How “Circumpolar” is Ainu Music ? Musical and Genetic Perspectives on the History of the Japanese Archipelago ». In : *Ethnomusicology Forum* 24.3 (2015), pp. 443–467 (citée pp. 252, 254, 255, 257–259).
- [373] SCHILLINGER, Kerstin, MESOUDI, Alex et LYCETT, Stephen J. « Copying Error, Evolution, and Phylogenetic Signal in Artifactual Traditions : An Experimental Approach Using “Model Artifacts” ». In : *Journal of Archaeological Science* 70 (2016), pp. 23–34 (citée pp. 60, 70, 87, 100, 368).
- [377] SCHNEIDER, Albrecht. « Comparative and Systematic Musicology in Relation to Ethnomusicology : A Historical and Methodological Survey ». In : *Ethnomusicology* 50.2 (2006), pp. 236–258 (citée pp. 130, 132, 133, 136).
- [379] DE SCHRYVER, Gilles-Maurice, GROLEMUND, Rebecca, BRANFORD, Simon et BOSTOEN, Koen. « Introducing a State-of-the-Art Phylogenetic Classification of the Kikongo Language Cluster ». In : *Africana Linguistica* 21 (2015), pp. 87–162 (citée pp. 65, 94).
- [380] SHELEMAY, Kay Kaufman. « Recording Technology, the Record Industry and Ethnomusicological Scholarship ». In : *Comparative Musicology and Anthropology of Music : Essays on the History of Ethnomusicology*. Sous la dir. de NETTL, Bruno et BOHLMAN, Philip V. Essays on the History of Ethnomusicology. University Of Chicago Press, 1991, pp. 277–292 (citée p. 139).
- [382] SITUNGKIR, Hokky. *Conjectures to the Memes of Indonesian Songs*. Rapp. tech. Working Paper Series WP-VI-2008, Bandung Fe Institute, 2008 (citée p. 100).
- [383] SITUNGKIR, Hokky. *Constructing the Phylomemetic Tree : Case of Study : Indonesian Tradition-Inspired Buildings*. Rapp. tech. Bandung Fe Institute, 2008 (citée p. 100).
- [384] SKELTON, Christina. « Methods of Using Phylogenetic Systematics To Reconstruct The History Of The Linear B Script ». In : *Archaeometry* 50.1 (2008), pp. 158–176 (citée pp. 74, 101).
- [385] SMOUSE, Peter E., LONG, Jeffrey C. et SOKAL, Robert R. « Multiple Regression and Correlation Extensions of the Mantel Test of Matrix Correspondence ». In : *Systematic Zoology* 35.4 (1986), pp. 627–632 (citée pp. 367, 422).

- [390] SPENCER, Matthew, BORDALEJO, Barbara, ROBINSON, Peter et HOWE, Christopher J. « How Reliable is a Stemma ? An Analysis of Chaucer's Miller's Tale ». In : *Literary and Linguistic Computing* 18.4 (2003), pp. 407–422 (citée pp. 54, 55, 58–60, 101).
- [391] SPENCER, Matthew, BORDALEJO, Barbara, WANG, Li-San, BARBROOK, Adrian C., MOONEY, Linne R., ROBINSON, Peter, WARNOW, Tandy et HOWE, Christopher J. « Analyzing the Order of Items in Manuscripts of The Canterbury Tales ». In : *Computers and the Humanities* 37.1 (2003), pp. 97–109 (citée pp. 54, 59, 101).
- [392] SPENCER, Matthew, DAVIDSON, Elizabeth A., BARBROOK, Adrian C. et HOWE, Christopher J. « Phylogenetics of Artificial Manuscripts ». In : *Journal of Theoretical Biology* 227.4 (2004), pp. 503–511 (citée pp. 54–56, 60, 61, 82, 87, 90, 91, 101, 368).
- [393] SPENCER, Matthew, MOONEY, Linne R., BARBROOK, Adrian C., BORDALEJO, Barbara, HOWE, Christopher J. et ROBINSON, Peter. « The Effects of Weighting Kinds of Variants ». In : *Studies in Stemmatology II*. Sous la dir. de VAN REENEN, Pieter, DEN HOLLANDER, August et VAN MULKEN, Margot. Amsterdam/Philadelphia : John Benjamins Publishing Company, 2004, pp. 227–239 (citée pp. 54, 55, 57, 101).
- [394] SPENCER, Matthew, WACHTEL, Klaus et HOWE, Christopher J. « The Greek Vorlage of the Syra Harclensis : A Comparative Study on Method in Exploring Textual Genealogy ». In : *TC : A Journal of Biblical Textual Criticism* 7 (2002), pp. 8–2 (citée pp. 55, 57, 101).
- [395] SPENCER, Matthew, WACHTEL, Klaus et HOWE, Christopher J. « Representing Multiple Pathways of Textual Flow in the Greek Manuscripts of the Letter of James Using Reduced Median Networks ». In : *Computers and the Humanities* 38.1 (2004), pp. 1–14 (citée pp. 54, 101).
- [396] STEVENS, Peter F. *The Development of Biological Systematics*. Columbia University Press, 1994 (citée p. 363).
- [397] STOCK, Jonathan P. J. « Clues from Our Present Peers ? : A Response to Victor Grauer ». In : *The World of Music* 48.2 (2006), pp. 73–91 (citée pp. 143, 252).
- [398] STONE, Ruth M. *Theory for Ethnomusicology*. Routledge, 2007. 242 Seiten (citée p. 133).
- [399] STRAFFON, Larissa Mendoza, éd. *Cultural Phylogenetics. Concepts and Applications in Archaeology*. 4. Springer, 2016.
- [400] STRAFFON, Larissa Mendoza. « The Applications and Challenges of Cultural Phylogenetics in Archaeology : An Introduction ». In : *Cultural Phylogenetics — Concepts and Applications in Archaeology*. Sous la dir. de STRAFFON, Larissa Mendoza. Interdisciplinary Evolution Research 4. Springer, 2016, pp. 1–15 (citée p. 356).
- [401] STUESSY, Tod F. « Phenetic Approach ». In : *Plant Taxonomy : The Systematic Evaluation of Comparative Data*. New York : Columbia University Press, 2009. Chap. 7, pp. 51–72 (citée p. 21).
- [408] TEHRANI, J. et COLLARD, M. « Investigating Cultural Evolution Through Biological Phylogenetic Analyses of Turkmen Textiles ». In : *Journal of Anthropological Archaeology* 21.4 (2002), pp. 443–463 (citée pp. 67, 79, 81, 101, 366, 367).
- [409] TEHRANI, Jamshid J. « The Phylogeny of Little Red Riding Hood ». In : *PLOS ONE* 8.11 (2013), e78871 (citée pp. 74, 80, 101, 352).

- [410] TEHRANI, Jamshid J. et COLLARD, Mark. « On the Relationship Between Interindividual Cultural Transmission and Population-Level Cultural Diversity : A Case Study of Weaving in Iranian Tribal Populations ». In : *Evolution and Human Behavior* 30.4 (2009), pp. 286–300 (citée pp. 67, 69, 101).
- [411] TEHRANI, Jamshid J., COLLARD, Mark et SHENNAN, Stephen J. « The Cophylogeny of Populations and Cultures : Reconstructing the Evolution of Iranian Tribal Craft Traditions Using Trees and Jungles ». In : *Philosophical Transactions of the Royal Society B : Biological Sciences* 365.1559 (2010), pp. 3865–3874 (citée pp. 69, 79, 101).
- [412] TEMKIN, Ilya. « The Evolution of the Baltic Psaltery : A Case for Phyloorganology ». In : *The Galpin Society Journal* 57 (2004), pp. 219–230 (citée pp. 69, 70, 101, 264, 266, 283).
- [413] TEMKIN, Ilya et ELDREDGE, Niles. « Phylogenetics and Material Cultural Evolution ». In : *Current anthropology* 48.1 (2007), pp. 146–153 (citée pp. 69, 71, 79, 101, 262, 273, 286, 287).
- [414] THUL, Eric et TOUSSAINT, Godfried T. « A Comparative Phylogenetic-Tree Analysis of African Timelines and North Indian Talas ». In : *Bridges Leeward : Mathematics, Music, Art, Architecture, Culture*. Sous la dir. de SARHANGI, Reza et SÉQUIN, Carlo H. Winfield, Kansas : Southwestern College, 2008, pp. 187–194. URL : <http://archive.bridgesmathart.org/2008/bridges2008-187.pdf> (citée pp. 71, 72, 90, 101).
- [417] TOIVIAINEN, Petri et EEROLA, Tuomas. « A Method for Comparative Analysis of Folk Music Based on Musical Feature Extraction and Neural Networks ». In : *Proceedings of the VII International Symposium of Systematic and Comparative Musicology and the III International Conference on Cognitive Musicology*. Sous la dir. de LAPPALAINEN, H. Jyväskylä : University of Jyväskylä, 2001, pp. 41–45 (citée p. 123).
- [419] TOUSSAINT, Godfried. « A Mathematical Analysis of African, Brazilian, and Cuban Clave Rhythms ». In : *Proceedings of Bridges : Mathematical Connections in Art, Music, and Science*. Towson, 2002, pp. 157–168 (citée pp. 71, 84, 101).
- [420] TOUSSAINT, Godfried. *Curriculum Vitae : Highlights*. URL : <http://www-cgr1.cs.mcgill.ca/~godfried/cv/highlights.html> (consultée le 13/04/2017) (citée p. 303).
- [421] TOUSSAINT, Godfried T. « Classification and Phylogenetic Analysis of African Ternary Rhythm Timelines ». In : *Proceedings of Bridges : Mathematical Connections in Art, Music and Science*. 2003, pp. 25–36 (citée pp. 71, 73, 102).
- [422] TOUSSAINT, Godfried T. « A Comparison of Rhythmic Similarity Measures ». In : *Proceedings of the 5th ISMIR Conference*. International Society for Music Information Retrieval (ISMIR). 2004 (citée p. 290).
- [423] TRIPP, Allison. « A Cladistics Analysis Exploring Regional Patterning of the Anthropomorphic Figurines from the Gravettian ». In : *Cultural Phylogenetics — Concepts and Applications in Archaeology*. Sous la dir. de STRAFFON, Larissa Mendoza. Interdisciplinary Evolution Research 4. Springer, 2016, pp. 179–202 (citée pp. 68, 102).
- [424] TUCKER, Aviezer. « Sciences of Historical Tokens and Theoretical Types : History and the Social Sciences ». In : *The Oxford Handbook of Philosophy of Social Science*. Sous la dir. de KINCAID, Harold. 1^{re} éd. Oxford Handbooks. Oxford University Press, 2012, pp. 274–297 (citée pp. 183, 249, 250, 342).
- [425] URBAN, D. L. *Spatial Analysis in Ecology : Mantel's Test*. URL : <http://www.nceas.ucsb.edu/files/scicomp/doc/SpatialEcologyMantelTest.pdf> (consultée le 30/04/2014) (citée p. 421).

- [426] VAN DER VEEN, Lolke, QUINTANA-MURCI, Lluís et COMAS, David. « Linguistic, Cultural and Genetic Perspectives on Human Diversity in West-Central Africa ». In : *Becoming Eloquent : Advances in the Emergence of Language, Human Cognition, and Modern Cultures*. Sous la dir. de D'ERRICO, Francesco et HOMBERT, Jean-Marie. John Benjamins Publishing Company, 2009, pp. 93–121 (citée p. 183).
- [427] VAN DRIEM, George. « Sino-Austronesian vs. Sino-Caucasian, Sino-Bodic vs. Sino-Tibetan, and Tibeto-Burman as Default Theory ». In : *Contemporary Issues in Nepalese Linguistics*. Sous la dir. d'YADAVA, Yogendra Prasada, BHATTARAI, Govinda, LOHANI, Ram Raj, PRASAIN, Balaram et PARAJULI, Krishna. Kathmandu : Linguistic Society of Nepal, 2005, pp. 285–338 (citée pp. 22, 23).
- [428] VECCHIONE, Bernard. « Musique et modèles. Vers une typologie des modèles en musique ». In : *Analyse Musicale* 22 (1991), pp. 13–29 (citée p. 124).
- [430] VENDITTI, Chris, MEADE, Andrew et PAGEL, Mark. « Multiple Routes to Mammalian Diversity ». In : *Nature* 479.7373 (2011), pp. 393–396 (citée pp. 157–159, 161, 164, 391).
- [434] WALKER, Robert S. et RIBEIRO, Lincoln A. « Bayesian Phylogeography of the Arawak Expansion in Lowland South America ». In : *Proceedings of the Royal Society B : Biological Sciences* 278.1718 (2011), pp. 2562–2567 (citée pp. 65, 102).
- [437] WALLIN, Nils Lennart, MERKER, Björn et BROWN, Steven, éd. *The Origins Of Music*. MIT Press, 2000 (citée p. 341).
- [438] WIKTIONNAIRE. *fibule*. online. URL : <https://fr.wiktionary.org/w/index.php?title=fibule&oldid=21616029> (consultée le 06/12/2016) (citée p. 66).
- [439] WIKTIONNAIRE. *inférer*. URL : <https://fr.wiktionary.org/w/index.php?title=inf%C3%A9rer&oldid=20093018> (consultée le 09/08/2016) (citée p. 344).
- [441] WILLIAMS, D. M. et EBACH, M. C. « What, Exactly, is Cladistics? Re-writing the History of Systematics and Biogeography ». In : *Acta Biotheoretica* 57.1-2 (2009), pp. 249–268 (citée p. 21).
- [444] WINDRAM, Heather F., SHAW, Prue, ROBINSON, Peter et HOWE, Christopher J. « Dante's Monarchia as a Test Case for the Use of Phylogenetic Methods in Stemmatic Analysis ». In : *Literary and Linguistic Computing* 23.4 (2008), pp. 443–463 (citée pp. 54, 55, 59, 60, 92, 102).
- [447] WOERTHER, Frédérique et KHONSARI, Hossein. « L'application des programmes de reconstruction phylogénétique sur ordinateur à l'étude de la tradition manuscrite d'un texte : l'exemple du chapitre XI de l'Ars Rhetorica du Pseudo-Denys d'Halicarnasse ». In : *Revue d'histoire des textes* 31 (2001), pp. 227–240 (citée pp. 55, 57, 60, 102).

C.2 Autres publications citées

- [A1] ADLER, Guido. « Umfang, Methode und Ziel der Musikwissenschaft ». In : *Vierteljahrsschrift für Musikwissenschaft* 1 (1885), pp. 5–20. URL : <https://ia800307.us.archive.org/6/items/vierteljahrssch23unkngoog/vierteljahrssch23unkngoog.pdf> (citée pp. 130, 351).
- [A2] ALAGO, E. J. « Le chant comme source d'information historique ». In : *La tradition orale : problématique et méthodologie des sources de l'histoire africaine*. Sous la dir. de LAYA, Diouldé. Niamey : Centre régional de documentation pour la tradition orale, 1972 (citée p. 139).

- [A4] AMIOT, Joseph-Marie. *Mémoire sur la musique des Chinois, tant anciens que modernes*. Paris : Nyon l'aîné, 1779. URL : https://ia600209.us.archive.org/7/items/bub_gb_OI1YAAAAcAAJ/bub_gb_OI1YAAAAcAAJ.pdf (citée p. 130).
- [A5] AMSELLE, Jean-Loup et M'BOKOLO, Elikia. *Au cœur de l'ethnie : ethnie, tribalisme et État en Afrique*. Paris : La Découverte, 1999 (citée p. 267).
- [A6] ANDERSEN, Esben Sloth. « The Evolving Tree of Industrial Life : An Approach to the Transformation of European Industry ». In : *Second Workshop on the Economic Transformation of Europe, Torino 31 Jan–2 Feb. 2003* (citée p. 75).
- [A8] ARCHIE, James W. « A Randomization Test for Phylogenetic Information in Systematic Data ». In : *Systematic Biology* 38.3 (1989), pp. 239–252 (citée pp. 344, 366).
- [A10] ARCHIE, James W. « Homoplasy Excess Statistics and Retention Indices : A Reply to Farris ». In : *Systematic Zoology* 39.2 (1990), pp. 169–174 (citée p. 344).
- [A29] AUROUX, Sylvain. « Representation and the Place of Linguistic Change before Comparative Grammar ». In : *Leibniz, Humboldt, and the Origins of Comparativism*. Sous la dir. de DE MAURO, Tullio et FORMIGARI, Lia. T. 49. Studies in the History of the Language Sciences. Amsterdam : John Benjamins Publishing Company, 1990, pp. 213–238 (citée p. 18).
- [A30] BALDWIN, James S., ALLEN, Peter M., WINDER, Belinda et RIDGWAY, Keith. « Simulating the Cladistic Evolution of Manufacturing. » In : *Innovation : Management, Policy and Practice* 5 (2003), pp. 144–156 (citée p. 75).
- [A31] BANDELT, Hans-Jürgen. « Phylogenetic Networks ». In : *Verhandlungen des naturwissenschaftlichen Vereins Hamburg* 34 (1994), pp. 51–71 (citée pp. 62, 89).
- [A32] BANDELT, Hans-Jürgen, FORSTER, Peter, SYKES, Bryan C. et RICHARDS, Martin B. « Mitochondrial Portraits of Human Populations Using Median Networks ». In : *Genetics* 141 (1995), pp. 743–753 (citée p. 91).
- [A35] BARTH, Frederik. « Les groupes ethniques et leurs frontières ». In : *Théories de l'ethnicité*. Sous la dir. de POUTIGNAT, P. et STREIFF-FENART, J. Presses universitaires de France (PUF), 1995 [1969], pp. 203–249 (citée p. 267).
- [A36] BATEMAN, Richard, GODDARD, Ives, O'GRADY, Richard, FUNK, Vicki A., MOOI, Rich, KRESS, W.. John et CANNELL, Peter. « Speaking of Forked Tongues : The Feasibility of Reconciling Human Phylogeny and the History of Language ». In : *Current Anthropology* 31.1 (1990), pp. 1–24 (citée p. 251).
- [A37] BAYARD, Samuel P. « Prolegomena to a Study of the Principal Melodic Families of British-American Folk Song ». In : *The Journal of American Folklore* 63.247 (1950), pp. 1–44 (citée pp. 137, 141).
- [A38] BEC, Pierre. *Les instruments de musique d'origine arabe. Sens et histoire de leurs désignations*. Toulouse-Montpellier : Conservatoire occitan-musique et danse en Languedoc-Roussillon, 2004 (citée p. 140).
- [A39] BÉDIER, Joseph. *La tradition manuscrite du Lai de l'ombre : réflexions sur l'art d'éditer les anciens textes*. Paris : Honoré Champion, 1928 (citée p. 19).
- [A42] BEN HAMED, Mahé, DARLU, Pierre et VALLÉE, Nathalie. « On Cladistic Reconstruction of Linguistic Trees Through Vowel Data ». In : *Journal of Quantitative Linguistics* 12.1 (2005), pp. 79–109 (citée pp. 63, 93).
- [A44] BLACKING, John. « Review of "The Anthropology of Music" by Alan P. Merriam ». In : *Current Anthropology* 7.2 (1966), p. 218 (citée p. 131).

- [A48] BOOKSTEIN, Abraham, KLEIN, Shmuel Tomi et RAITA, Timo. « Fuzzy Hamming Distance : A new Dissimilarity Measure ». In : *Proceedings of the 12th Annual Symposium on Combinatorial Pattern Matching*. Berlin : Springer-Verlag, 2001, pp. 86–97 (citée p. 290).
- [A50] BORGATTI, S., EVERETT, M. et FREEMAN, L. *UCINET 6 for Windows*. Lexington, Kentucky : Analytic Technologies, 2010 (citée p. 91).
- [A58] VAN BOXHORN, Marcus Zuerius. *Bediedinge van de tot noch toe onbekende Afgodinne Nehalennia, over de dusent ende ettelicke hondert Jaren onder het Sandt begraven, dan onlanca ontdeckt op het Strandt van VValcheren in Zeelandt*. Leyden : Willem Christiaens van der Boxe, 1647 (citée p. 22).
- [A59] VAN BOXHORN, Marcus Zuerius. *Vraagen voorghestelt ende Oghedraaghen aan de Heer Marcus Zuerius van Boxhorn over de Bediedinge van de tot noch toe onbekende Afgodinne Nehalennia, onlangs by Hem uytgegevee*. Leyden : Willem Christiaens van der Boxe, 1647 (citée p. 22).
- [A60] VAN BOXHORN, Marcus Zuerius. *Antwoord van Marcus Zuerius van Boxhorn op de Vraaghen, hem voorgestelt over de Bediedinge van de tot noch toe onbekende Afgodinne Nehalennia, onlanca uytgegeven. In welke de ghemeine herkomste van der Griecken, Romeinen, ende Duytschen Tale uyt den Scythen duydelyck bewesen, ende verscheiden Oudheden van dese Volckeren grondelijck ontdeckt ende verklaert*. Leyden : Willem Christiaens van der Boxe., 1647 (citée p. 22).
- [A62] BRĂILOIU, Constantin. « La vie antérieure ». In : *Histoire de la musique*. Sous la dir. de ROLAND-MANUEL. T. 1. La Pléiade. Paris : Gallimard, 2009, pp. 119–127 (citée p. 140).
- [A63] BREMER, K. « Ancestral Areas : A Cladistic Reinterpretation of the Center of Origin Concept ». In : *Systematic Biology* 41.4 (1992), pp. 436–445 (citée pp. 40, 42).
- [A64] BREMER, Kare. « The Limits of Amino Acid Sequence Data in Angiosperm Phylogenetic Reconstruction ». In : *Evolution* 42.4 (1988), pp. 795–803 (citée p. 343).
- [A74] CAMIN, Joseph H. et SOKAL, Robert R. « A Method for Deducing Branching Sequences in Phylogeny ». In : *Evolution* 19.3 (1965), pp. 311–326 (citée p. 21).
- [A81] CAVALLI-SFORZA, Luigi Luca et EDWARDS, A. W. F. « Phylogenetic Analysis : Models and Estimation Procedures ». In : *American Journal of Human Genetics* 19 (1967), pp. 233–257 (citée p. 90).
- [A82] CAVALLI-SFORZA, Luigi Luca et EDWARDS, A. W. F. « Phylogenetic Analysis : Models and Estimation Procedures ». In : *Evolution* 21.3 (1967), pp. 550–570 (citée p. 21).
- [A84] CHANG, Will, CATHCART, Chundra, HALL, David et GARRETT, Andrew. « Ancestry-Constrained Phylogenetic Analysis Supports the Indo-European Steppe Hypothesis ». In : *Language* 91.1 (2015), pp. 194–244 (citée p. 65).
- [A85] CHAPA, Teresa. « La escultura zoomorfa ibérica en piedra ». Thèse de doctorat. Universidad Complutense de Madrid., 1980 (citée p. 66).
- [A86] CHAPA, Teresa. « Aspectos metodológicos de la tipología arqueológica : Un ejemplo referido a las fíbulas de La Tene ». In : *Primeras jornadas de metodología de investigación prehistórica, Soria 1981*. Madrid : Instituto de Conservación y Restauración de Bienes Culturales, 1984, pp. 253–268 (citée p. 66).
- [A88] CHARLESTON, M. A. « Jungles : A new Solution to the Host/Parasite Phylogeny Reconciliation Problem ». In : *Mathematical Biosciences* 149.2 (1998), pp. 191–223 (citée pp. 69, 79).

- [A90] CHEN, Yu-Sheng, OLCKERS, Antonel, SCHURR, Theodore G., KOGELNIK, Andreas M., HUOPONEN, Kirsi et WALLACE, Douglas C. « Mitochondrial DNA Variation in the South African Kung and Khwe—and Their Genetic Relationships to Other African Populations ». In : *American Journal of Human Genetics* 66 (2000), pp. 1362–1383 (citée p. 252).
- [A91] CHENOWETH, Vida. *Melodic Analysis and Perception*. Papua-New Guinea : Summer Institute of Linguistics, 1972 (citée p. 109).
- [A95] COLLIN, H. S. et SCHLYTER, C. J. *Corpus Iuris Sueo-Gotorum Antiqui*. Stockholm : Z. Haeggstrom, 1827. URL : <https://ia801403.us.archive.org/12/items/corpusiurissueo11swedgoog/corpusiurissueo11swedgoog.pdf> (citée p. 18).
- [A96] COPAS, J. B. « Regression, Prediction and Shrinkage ». In : *Journal of the Royal Statistical Society Series B (Methodological)* 45.3 (1983), pp. 311–354 (citée p. 256).
- [A99] COWDERY, James R. *The Melodic Tradition of Ireland*. Kent State University Press, 1990 (citée p. 137).
- [A104] DAGAN, Tal, YAEL, Artzy-Randrup et WILLIAM, Martin. « Modular Networks and Cumulative Impact of Lateral Transfer in Prokaryote Genome Evolution ». In : *Proceedings of the National Academy of Sciences (PNAS)* 105 (2008), pp. 10039–10044 (citée pp. 63, 79, 89).
- [A105] DANCKERT, Werner. *Das europäische Volkslied*. J. Bard, 1939 (citée p. 133).
- [A109] DARWIN, Charles. *Charles Darwin's Notebooks, 1836-1844 : Geology, Transmutation of Species, Metaphysical Enquiries*. Sous la dir. de BARRETT, Paul H., GAUTREY, Peter J., HERBERT, Sandra, KOHN, David et SMITH, Sydney. Cambridge University Press, 1987 (citée p. 18).
- [A110] DARWIN, Charles Robert. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London : John Murray, 1859 (citée p. 20).
- [A111] DAY, William H. E. et SANKOFF, David. « Computational Complexity of Inferring Phylogenies by Compatibility ». In : *Systematic Biology* 35.2 (1986), pp. 224–229 (citée p. 32).
- [A113] DEARING, V. A. *Principles and Practice of Textual Analysis*. Berkeley : University of California Press, 1974 (citée pp. 19, 29, 30).
- [A118] DENSMORE, Frances. *Chippewa Music II*. Washington D.C. : Government Printing Office, 1913 (citée p. 137).
- [A119] DENSMORE, Frances. *Teton Sioux Music*. Washington D.C. : Government Printing Office, 1918 (citée p. 137).
- [A125] DON CAMERON, H. « The Upside-down Cladogram : Problems in Manuscript Affiliation ». In : *Biological Metaphor and Cladistic Classification : An Interdisciplinary Perspective*. Sous la dir. d'HOENIGSWALD, Henry M. et WIENER, Linda F. Philadelphia : University of Pennsylvania Press, 1987, pp. 227–242 (citée pp. 18, 58).
- [A126] DRIVER, Harold E. *Indians of North America*. Chicago : University of Chicago Press, 1961 (citée p. 134).
- [A130] EDWARDS, Anthony W. F. et CAVALLI-SFORZA, Luigi Luca. « Reconstruction of Evolutionary Trees ». In : *Phenetic and Phylogenetic Classification*. Sous la dir. d'HEYWOOD, V. H. et MCNEILL, J. T. 6. London : Systematics Association, 1964, pp. 67–76 (citée p. 22).

- [A133] FARRIS, James S. « The Retention Index and the Rescaled Consistency Index ». In : *Cladistics* 5.4 (1989), pp. 417–419 (citée pp. 38, 344).
- [A139] FISHER, Daniel C. « Stratocladistics : Integrating Temporal Data and Character Data in Phylogenetic Inference ». In : *Annual Review of Ecology, Evolution, and Systematics* 39 (2008), pp. 365–385 (citée p. 90).
- [A140] FISLER, Marie. « Histoire de la métaphore de l’arbre dans les sciences naturelles ». Thèse de doctorat. Paris : Muséum national d’histoire naturelle (MNHN), 2014 (citée p. 328).
- [A142] FITCH, Walter M. et MARGOLIASH, Emanuel. « Construction of Phylogenetic Trees ». In : *Science* 155 (1967), pp. 279–284 (citée pp. 21, 90, 375).
- [A148] FRISBIE, Charlotte. « Anthropological and Ethnomusicological Implications of a Comparative Analysis of Bushmen and Pygmy Music ». In : *Ethnology* X.3 (1971), pp. 265–290 (citée p. 115).
- [A149] FROBENIUS, Leo. « The Origin of African Civilizations ». In : *Annual Report of the Board of Regents of the Smithsonian Institution showing the Operations, Expenditures, and Condition of the Institution for the Year ending June 30, 1898*. T. 1. Washington : Government Printing Office, 1899, pp. 637–650. URL : <https://archive.org/download/annualreportofb1898smit/annualreportofb1898smit.pdf> (citée p. 133).
- [A155] GAILLARD-CORVAGLIA, Antonella, LÉONARD, Jean-Léo et DARLU, Pierre. « Testing Cladistics on Dialect Networks and Phyla (Gallo-Romance and Southern Italo-Romance), » in : *Proceedings of Ninth Meeting of the ACL Special Interest Group in Computational Morphology and Phonology*. Prague : Association for Computational Linguistic (ACL), 2007, pp. 23–30. URL : <http://aclweb.org/anthology/W/W07/W07-1304.pdf> (citée pp. 63, 284).
- [A158] GASCUEL, Olivier. « BIONJ : An Improved Version of the NJ Algorithm Based on a Simple Model of Sequence Data ». In : *Molecular Biology and Evolution* 14.7 (1997), pp. 685–695 (citée p. 90).
- [A161] GOMBERG, David. « “Bayesian” Postdiction in an Evolution Process ». Istituto di Genetica, University of Pavia, Italy. 1968 (citée p. 22).
- [A170] GREENBERG, Joseph H. *Studies in African Linguistic Classification*. New Haven : Branford, 1955 (citée p. 23).
- [A171] GREENHILL, Simon J., ATKINSON, Quentin D., MEADE, Andrew et GRAY, Russell D. « The Shape and Tempo of Language Evolution ». In : *Proceedings of the Royal Society B : Biological Sciences* 277.1693 (2010), pp. 2443–2450 (citée p. 263).
- [A173] GREG, W.W. *The Calculus of Variants. An Essay on Textual Criticism*. Oxford. : Clarendon Press, 1927 (citée pp. 19, 29, 30).
- [A177] GRIMAUD, Yvette. « Note sur la musique vocale des Bochimani Kung et des pygmées Babinga ». In : *Les colloques de Wégimont III*. T. II. Ethnomusicologie. Paris, 1960, pp. 105–126 (citée p. 115).
- [A178] GRIMM, Jacob. *Deutsche grammatik. Part I : Zweite ausgabe*. Gottingen : Dieterich, 1822 (citée p. 23).
- [A183] HARRIS, Marvin. *The Rise of Anthropological Theory. A History of Theories of Culture*. New-York : Thomas Y. Crowell, 1968 (citée p. 134).
- [A185] HAUSDORF, Bernhard. « Weighted Ancestral Area Analysis and a Solution of the Redundant Distribution Problem ». In : *Systematic Biology* 47.3 (1998), pp. 445–456 (citée p. 42).

- [A188] HENDY, Michael D. et PENNY, David. « Branch and Bound Algorithms to Determine Minimal Evolutionary Trees ». In : *Mathematical Biosciences* 59 (1982), pp. 277–290 (citée p. 88).
- [A189] HENNIG, Willi. *Grundzüge einer Theorie der phylogenetischen Systematik*. Deutscher Zentralverlag, 1950 (citée pp. 21, 334).
- [A190] HENNIG, Willi. *Phylogenetic Systematics*. Trad. par DAVIS, D. Dwight et ZANGERL, Rainer. University of Illinois Press, 1966 (citée p. 34).
- [A191] HERZOG, George. « The Yuman Musical Style ». In : *Journal of American Folklore* 41.160 (1928), pp. 183–231 (citée p. 134).
- [A197] HOLDEN, Clare Janaki et MACE, Ruth. « Spread of Cattle Led to the Loss of Matrilineal Descent in Africa : A Coevolutionary Analysis ». In : *Proceedings Biological Sciences* 270.1532 (2003), pp. 2425–2433 (citée pp. 260, 348).
- [A198] HOLDEN, Clare et MACE, Ruth. « Phylogenetic Analysis of the Evolution of Lactose Digestion in Adults ». In : *Human Biology* 69.5 (1997), pp. 605–628 (citée p. 260).
- [A199] HOLLAND, Barbara R., HUBER, Katharina T., DRESS, Andreas et MOULTON, Vincent. « δ plots : A Tool for Analyzing Phylogenetic Distance Data ». In : *Molecular Biology and Evolution* 19.12 (2002), pp. 2051–2059 (citée p. 80).
- [A200] HOOD, Mantel. « Ethnomusicology ». In : *Harvard Dictionary of Music*. Sous la dir. d'APPEL, Willi. Cambridge : Harvard University Press, 1969, pp. 298–300 (citée p. 131).
- [A202] VON HORNBOSTEL, Erich Moritz. « Die Probleme der vergleichenden Musikwissenschaft ». In : *Zeitschrift der Internationalen Musikgesellschaft* 7.3 (1905), pp. 85–97. URL : <https://archive.org/download/ZeitschriftDerInternationalenMusikgesellschaft071905-06/ZeitschriftDerInternationalenMusikgesellschaft071905-06.pdf> (citée p. 130).
- [A206] HUSON, Daniel H. « SplitsTree : Analyzing and Visualizing Evolutionary Data ». In : *Bioinformatics* 14.1 (1998), pp. 68–73 (citée pp. 90, 91).
- [A208] HUSON, Daniel H., DEZULIAN, Tobias, KLÖPPER, Tobias et STEEL, Mike A. « Phylogenetic Super-Networks from Partial Trees ». In : *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 1.4 (2004), pp. 151–158 (citée p. 92).
- [A211] JEFFERSON, Thomas. *Reprint of Notes on the State of Virginia 1781–1782*. Peterson, Merrill D. New York : Library of America, 1984 (citée p. 23).
- [A212] JORDAN, Fiona M., GRAY, Russel D., GREENHILL, Simon J. et MACE, Ruth. « Matrilocal Residence is Ancestral in Austronesian Societies ». In : *Proceedings of the Royal Society B : Biological Sciences* 276.1664 (2009), pp. 1957–1964 (citée p. 260).
- [A213] JORDAN, Fiona M. et MACE, Ruth. *Changes in Post-Marital Residence Precede Changes In Descent Systems in Austronesian Societies*. Paper for the European Human Behaviour and Evolution (EHBE) Conference. 2007 (citée pp. 64, 97, 260).
- [A218] JUKES, Thomas H. et CANTOR, Charles R. « Evolution of protein molecules ». In : *Mammalian Protein Metabolism*. Sous la dir. de MUNRO, Hamish N. T. 3. New York : Academic Press, 1969, pp. 21–132 (citée p. 21).
- [A219] KANNAN, Sampath et WARNOW, Tandy. « A Fast Algorithm for the Computation and Enumeration of Perfect Phylogenies When the Number of Character States is Fixed ». In : *SIAM Journal of Computing* 26 (1997), pp. 1749–1763 (citée p. 89).

- [A220] KASHYAP, R. L. et SUBAS, S. « Statistical Estimation of Parameters in a Phylogenetic Tree Using a Dynamic Model of the Substitutional Process ». In : *Journal of Theoretical Biology* 47 (1974), pp. 75–101 (citée p. 22).
- [A221] KATZ, Ruth. « The Singing of Baqqashôt by Aleppo Jews. A Study in Musical Acculturation ». In : *Acta Musicologica* 40.1 (1968), pp. 65–85 (citée p. 136).
- [A222] KIESEWETTER, Rafael Georg. *Die Musik der Araber*. Breitkopf und Härtel, 1842. URL : https://archive.org/download/bub_gb_YNX8AAAAMAAJ/bub_gb_YNX8AAAAMAAJ.pdf (citée p. 130).
- [A224] KISHINO, Hirohisa et HASEGAWA, Masami. « Evaluation of the Maximum Likelihood Estimate of the Evolutionary Tree Topologies from DNA Sequence Data, and the Branching Order in Hominoidea ». In : *Journal of Molecular Evolution* 29.2 (1989), pp. 170–179 (citée p. 366).
- [A226] KLUGE, Arnold G. et FARRIS, James S. « Quantitative Phyletics and the Evolution of Anurans ». In : *Systematic Zoology* 18.1 (1969), pp. 1–32 (citée p. 343).
- [A227] KOERNER, Konrad. « Preface ». In : *August Schleicher : Die Sprachen Europas in systematischer Übersicht*. Sous la dir. de KOERNER, K. Amsterdam : John Benjamins Publishing Company, 1983 (citée pp. 23, 375).
- [A228] KROEBER, Alfred Louis. *Cultural and Natural Areas of Native North America*. Berkeley : University of California Press, 1939 (citée p. 134).
- [A230] KUNST, Jaap. *Musicologica : A Study of the Nature of Ethno-music, Its Problems, Methods and Representative Personalities*. Amsterdam : Koninklijke Vereeniging Indisch Instituut, 1950 (citée p. 131).
- [A231] LACHMANN, Karl. *Lucretii de rerum natura libri*. Berlin, 1850 (citée pp. 18, 348, 375).
- [A246] LE QUESNE, Walter J. « A Method of Selection of Characters in Numerical Taxonomy ». In : *Systematic Zoology* 18 (1969), pp. 201–205 (citée p. 21).
- [A247] LEASK, Graham. « Evolutionary Change Within the UK Pharmaceutical Industry : A Cladistic Approach ». In : *DRUID Academy Winter 2002 PhD Conference, Aalborg, Denmark, January 17–19*. Danish Research Unit for Industrial Dynamics (DRUID). 2002 (citée p. 75).
- [A257] LEIBNIZ, Gottfried Wilhelm. *Miscellanea berolinensia in Berlin memoirs*. Berlin : Berlin Academy, 1710 (citée p. 23).
- [A259] LEROI, Armand M. et SWIRE, Jonathan. « The Recovery of the Past ». In : *The World of Music* 48.3 (2006), pp. 43–54 (citée p. 252).
- [A260] LESEURE, Michel. « Manufacturing Strategies in the Hand Tool Industry ». In : *International Journal of Operations and Production Management* 20 (2000), pp. 1475–1487 (citée p. 75).
- [A261] LEWIS, Oscar. *Life in a Mexican Village : Tepotzlan Restudied*. University of Illinois Press, 1951 (citée p. 136).
- [A272] MAAS, Paul. *Textual Criticism*. Clarendon Press, 1958 (citée pp. 19, 21, 25, 348, 362).
- [A278] MADDISON, David R., SWOFFORD, David L. et MADDISON, Wayne P. « NEXUS : An Extensible File Format for Systematic Information ». In : *Systematic Biology* 46 (4 1997), pp. 590–621 (citée p. 389).

- [A283] MCCARTHY, Ian P., LESEURE, Michel, RIDGWAY, Keith et NICK, Fieller. « Building a Manufacturing Cladogram ». In : *International Journal of Technology Management*. 13.3 (1997), pp. 269–296 (citée p. 75).
- [A285] MCCARTHY, Ian, RIDGWAY, Keith, LESEURE, Michel et FIELLER, Nick. « Organisational Diversity, Evolution and Cladistic Classifications ». In : *The International Journal of Management Science – Omega* 28.1 (2000), pp. 77–95 (citée p. 75).
- [A288] MEACHAM, Christopher A. « A Manual Method for Character Compatibility Analysis ». In : *Taxon* 30.3 (1981), pp. 591–600 (citée p. 34).
- [A289] MERRIAM, Alan P. « African Music ». In : *Continuity and Change in African Cultures*. Sous la dir. de BASCOM, William R. et HERSKOVITS, Melville J. University of Chicago Press, 1959, pp. 49–86 (citée p. 134).
- [A293] MEYER, Leonard B. « Universalism and Relativism in the Study of Ethnic Music ». In : *Ethnomusicology* 4.2 (1960), pp. 49–54 (citée p. 131).
- [A294] MICHENER, Charles D. et SOKAL, Robert R. « A Quantitative Approach to a Problem in Classification ». In : *Evolution* 11.2 (1957), pp. 130–162 (citée p. 21).
- [A297] MOLINO, Jean. « Un nouveau paradigme en ethnomusicologie ». In : *Ndroje balendro : Musiques, terrains et disciplines : textes offerts à Simha Arom*. Sous la dir. de DEHOUX, Vincent, FÜRNISS, Susanne, LE BOMIN, Sylvie, OLIVIER, Emmanuelle, RIVIÈRE, Hervé et VOISIN, Frédéric. Peeters, 1995, pp. 197–209 (citée p. 123).
- [A299] MURDOCK, George Peter. « Ethnographic Atlas ». In : *Ethnology* 1–5 (1962–1967) (citée p. 135).
- [A302] NATTIEZ, Jean-Jacques. *Fondements d'une sémiologie de la musique*. Union générale d'éditions, 1975 (citée p. 109).
- [A307] NERBONNE, John, HEERINGA, Wilbert, VAN DEN HOUT, Erik, VAN DER KOOI, Peter, OTTEN, Simone et VAN DE VIS, Willem. « Phonetic Distance between Dutch Dialects ». In : *Proceedings of CLIN'95, Antwerp*. T. 6. 1996, pp. 185–202. URL : <http://urd.let.rug.nl/nerbonne/papers/clin96.pdf> (citée p. 62).
- [A308] NETTL, Bruno. *North American Indian Musical Styles*. American Folklore Society, 1954 (citée p. 134).
- [A309] NETTL, Bruno. « Studies in Blackfoot Indian Musical Culture, parts I and II ». In : *Ethnomusicology* 11 (1967), pp. 141–160, 293–309 (citée p. 136).
- [A315] NETTL, Bruno, STONE, Ruth, PORTER, James et RICE, Timothy, éd. *The Garland Encyclopedia of World Music*. New York : Routledge, 1998 (citée p. 265).
- [A316] NEYMAN, Jerzy. « Molecular Studies of Evolution : A Source of Novel Statistical Problems ». In : *Statistical Decision Theory and Related Topics*. Sous la dir. de GUPTA, S. S. et YACKEL, J. New York : Academic Press, 1971, pp. 1–27 (citée p. 22).
- [A317] NICHOLLS, Geoff R. et GRAY, Russel D. « Quantifying Uncertainty in a Stochastic Dollo Model of Vocabulary Evolution ». In : *Phylogenetic Methods and the Prehistory of Languages*. Sous la dir. de FORSTER, Peter et RENFREW, Colin. McDonald Institute for Archaeological Research, 2006 (citée pp. 77, 78).
- [A324] O'HARA, Robert J. « Telling the Tree : Narrative Representation and the Study of Evolutionary History ». In : *Biology and Philosophy* 7.2 (1992), pp. 135–160 (citée p. 364).

- [A327] OPIE, Christopher, SHULTZ, Susanne, ATKINSON, Quentin D., CURRIE, Thomas et MACE, Ruth. « Phylogenetic Reconstruction of Bantu Kinship Challenges Main Sequence Theory of Human Social Evolution ». In : *Proceedings of the National Academy of Sciences (PNAS)* 111.49 (2014), pp. 17414–17419 (citée p. 260).
- [A328] OPPENHEIMER, Stephen. *The Real Eve*. New York : Carroll & Graf, 2004 (citée p. 252).
- [A329] OSTHOFF, Hermann et BRUGMANN, Karl. *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*. Leipzig : S. Hirzel., 1878 (citée p. 23).
- [A335] PEDERSEN, H. *The Discovery of Language. Linguistic Science in the Nineteenth Century*. Bloomington : Indiana University Press, 1931. URL : <https://ia800502.us.archive.org/18/items/discoveryoflangu00pede/discoveryoflangu00pede.pdf> (citée pp. 22, 23).
- [A336] PERCIVAL, W. Keith. « Biological Analogy in the Study of Language Before the Advent of Comparative Grammar ». In : *Biological Metaphor and Cladistic Classification : An Interdisciplinary Perspective*. Sous la dir. d'HOENIGSWALD, Henry M. et WIENER, Linda F. Philadelphia : University of Pennsylvania Press, 1987. Chap. 1, pp. 3–38 (citée p. 22).
- [A337] PERI, Hiram. « Une méthode expérimentale de critique des textes ». In : *Atti del VIII Congresso Internazionale di Studi Romanzi (Firenze, 3–8 aprile 1956)*. T. 2. Florence : Sansoni, 1960, pp. 719–747 (citée p. 60).
- [A338] PIKE, Kenneth L. « Language in Relation to a Unified Theory of the Structure of Human Behavior ». In : (1954) (citée p. 108).
- [A343] QUENTIN, Henri. *Essais de critique textuelle (Ecdotique)*. A. Picard, 1926 (citée p. 19).
- [A344] R CORE TEAM. *R : A Language and Environment for Statistical Computing*. version 3.1.1. R Foundation for Statistical Computing. Vienna, Austria, 2014. URL : <http://www.R-project.org> (citée p. 166).
- [A345] RENART, Jean. *Le Lai de l'Ombre*. Sous la dir. de BÉDIER, Joseph. Paris : Société des anciens textes français, 1913. URL : <https://ia802608.us.archive.org/32/items/lelaidelombre00jeanuoft/lelaidelombre00jeanuoft.pdf> (citée p. 19).
- [A351] ROBERTS, Helen H. *Musical Areas in Aboriginal North America*. New Haven : Yale University Press, 1936 (citée p. 134).
- [A352] ROBERTS, Helen H. « The Viewpoint of Comparative Musicology ». In : *Proceedings of the Music Teachers National Association for 1936*. 1937 (citée p. 130).
- [A359] RONQUIST, Fredrik. « Ancestral Areas Revisited ». In : *Systematic Biology* 44.4 (1995), pp. 572–575 (citée p. 42).
- [A362] ROUGET, Gilbert. *La musique et la transe*. Gallimard, 1980 (citée p. 115).
- [A364] SACHS, Curt. *Geist und Werden der Musikinstrumente*. Berlin : Bard, 1929 (citée p. 133).
- [A365] SAITOU, Naruya et NEI, Masatoshi. « The Neighbor-joining Method : A New Method for Reconstructing Phylogenetic Trees ». In : *Molecular Biology and Evolution* 4.4 (1987), pp. 406–425 (citée pp. 54, 90, 357).
- [A371] SAVAGE, Patrick E., MERRITT, Emily, RZESZUTEK, Tom et BROWN, Steven. « Canto-Core : A New Cross-Cultural Song Classification Scheme ». In : *Analytical Approaches to World Music* 2.1 (2012), pp. 87–137 (citée p. 253).
- [A372] SCHAEFFNER, André. « Ethnologie musicale ou musicologie comparée ? » In : *Les colloques de Wégimont*. Sous la dir. de COLLAER, Paul. Bruxelles : Elsevier, 1961, pp. 18–32 (citée p. 351).

- [A374] SCHLEGEL, Friedrich. « Über die sprache und weisheit der Indier : Ein beitrage zur begründung der alterthumskunde ». In : *Amsterdam Classics in Linguistics*. Sous la dir. de TIMPANARO, S. Amsterdam : John Benjamins Publishing Company, 1808 (citée p. 23).
- [A375] SCHLEICHER, August. *Die Darwinsche theorie und die sprachwissenschaft*. Weimar : H. Böhlau, 1863 (citée pp. 18, 23).
- [A376] SCHNEIDER, Albrecht. « Vergleichende Musikwissenschaft als Morphologie und Stilkritik : Werner Danckerts Stellung in der Volksliedforschung und Musikethnologie ». In : *Jahrbuch für Volksliedforschung* 24 (1979), pp. 11–27 (citée p. 133).
- [A378] SCHNEIDER, Marius. *Geschichte der Mehrstimmigkeit*. Berlin : Bard, 1934 (citée p. 133).
- [A381] SIMPSON, George Gaylord. *Principles of Animal Taxonomy*. New York : Columbia University Press, 1961 (citée pp. 363, 366).
- [A386] SNEATH, Peter H. A. « The Application of Computers to Taxonomy ». In : *Journal of General Microbiology* 17 (1957), pp. 201–226 (citée p. 21).
- [A387] SNEATH, Peter H. A. et SOKAL, Robert R. *Numerical Taxonomy : The Principles and Practice of Numerical Classification*. San Francisco : W. H. Freeman, 1973, p. 573 (citée p. 21).
- [A388] SOKAL, Robert R. et MICHENER, C. D. « A Statistical Method for Evaluating Systematic Relationships ». In : *University of Kansas Science Bulletin* 38 (1958), pp. 1409–1438. URL : https://archive.org/download/cbarchive_33927_astatisticalmethodforevaluatin1902/astatisticalmethodforevaluatin1902.pdf (citée pp. 91, 369).
- [A389] SOKAL, Robert R. et SNEATH, Peter H. A. *Principles of Numerical Taxonomy*. W. H. Freeman et Company, 1963 (citée pp. 53, 365, 366).
- [A402] STUMPF, Carl. *Die Anfänge der Musik*. Leipzig : Johann Ambrosius Barth, 1911. URL : <https://archive.org/download/dieanfngedermu00stum/dieanfngedermu00stum.pdf> (citée p. 133).
- [A403] SWADESH, Moris. « Diffusional Cumulation and Archaic Residue as Historical Explanations ». In : *Southwestern Journal of Anthropology* 7 (1951), pp. 339–346 (citée p. 23).
- [A404] SWADESH, Moris. « Lexico-Statistic Dating of Prehistoric Ethnic Contacts ». In : *Proceedings of the American Philosophical Society* 96 (1952), pp. 453–463 (citée pp. 24, 341, 346).
- [A405] SWADESH, Moris. « Towards Greater Accuracy in Lexicostatistic Dating ». In : *International Journal of American Linguistics* 21 (1955), pp. 121–137 (citée pp. 23, 24, 346).
- [A406] TAJIMA, Atsushi, HAYAMI, Masanori, TOKUNAGA, Katsushi, JUJI, Takeo, MATSUO, Masafumi, MARZUKI, Sangkot, OMOTO, Keiichi et HORAI, Satoshi. « Genetic Origins of the Ainu Inferred from Combined DNA Analyses of Maternal and Paternal Lineages ». In : *Journal of Human Genetics* 49.4 (2004), pp. 187–193 (citée p. 255).
- [A407] TAPPERT, Wilhelm. *Wandernde Melodien*. 2^e éd. Leipzig : List und Francke, 1890 (citée p. 137).
- [A415] TIMPANARO, Sebastiano. *La genesi del metodo del Lachmann*. Turin : UTET, 2003 (citée pp. 18, 348, 375).
- [A416] TIMPANARO, Sebastiano. *The Genesis of Lachmann's Method*. Trad. par MOST, Glenn W. University of Chicago Press, 2005 (citée pp. 18, 19).

- [A418] TOMBEUR, P., BOULANGER, J.-C. et SCHUMACHER, J. « La génération automatique d'un stemma codicum : expériences de laboratoire ». In : *La pratique des ordinateurs dans la critique des textes. Colloque international du CNRS, Paris, 29-31 mars 1978*. Paris : Publications du Centre National de la Recherche Scientifique (CNRS), 1979, pp. 163-183 (citée p. 60).
- [A429] VENABLES, W. N. et RIPLEY, B. D. *Modern Applied Statistics with S*. 4^e éd. Springer-Verlag, 2002 (citée p. 166).
- [A431] VILLOTEAU, Guillaume André. « De l'état actuel de l'art musical en Egypte ». In : *Description de l'Égypte, ou Recueil des observations et des recherches qui ont été faites en Égypte pendant l'expédition de l'armée française*. Paris : Commission des Monuments d'Égypte, 1809, pp. 607-846. URL : <https://archive.org/download/DescriptiondeleE2Fran/DescriptiondeleE2Fran.pdf> (citée p. 130).
- [A432] VOISIN, Frédéric et CLOAREC-HEISS, France. « Échelles musicales et données linguistique : Vers une histoire des sociétés oubangiennes ». In : *Ndroje balendro : Musiques, terrains et disciplines : textes offerts à Simha Arom*. Sous la dir. de DEHOUX, Vincent, FÜRNISS, Susanne, LE BOMIN, Sylvie, OLIVIER, Emmanuelle, RIVIÈRE, Hervé et VOISIN, Frédéric. Peeters, 1995, pp. 119-140 (citée p. 142).
- [A433] WALKER, Robert S. et HAMILTON, Marcus J. « Social Complexity and Linguistic Diversity in the Austronesian and Bantu Population Expansions ». In : *Proceedings of the Royal Society B : Biological Sciences* 278 (2010), pp. 1399-1404 (citée p. 260).
- [A435] WALLASCHEK, Richard. *Primitive Music*. Longmans, Green, et Co., 1893. URL : <https://ia601408.us.archive.org/10/items/primitivemusicin00wall/primitivemusicin00wall.pdf> (citée p. 132).
- [A436] WALLASCHEK, Richard. *Anfänge der Tonkunst*. Leipzig : Johann Ambrosius Barth, 1903. URL : <https://ia800806.us.archive.org/14/items/anfngedertonkuns00wall/anfngedertonkuns00wall.pdf> (citée p. 132).
- [A440] WILL, Udo. « Et quand ils n'en disent rien?... ». In : *Cahiers de musiques traditionnelles* 11 (1998), pp. 175-185 (citée p. 123).
- [A442] WILSON, Edward O. « A Consistency Test for Phylogenies Based on Contemporaneous Species ». In : *Systematic Biology* 14.3 (1965), pp. 214-220 (citée pp. 21, 34).
- [A443] WINDELBAND, Wilhelm. « Rectorial Address, Strasbourg, 1894 ». In : *History and Theory* 19.2 (1980) (citée p. 249).
- [A445] WIORA, Walter. « Jodeln ». In : *Musik in Geschichte und Gegenwart*. Sous la dir. de BLUME, F. T. 7. Kassel/Basel : Bärenreiter, 1958, pp. 74-79 (citée p. 345).
- [A446] WISSLER, Clark. *The American Indian : An Introduction of the Anthropology of the New World*. New York : Douglas C. McMurtrie, 1917. URL : <https://ia802605.us.archive.org/18/items/americanindianin00wissuoft/americanindianin00wissuoft.pdf> (citée p. 134).
- [A448] ZUCKERKANDL, Emile et PAULING, Linus B. « Molecular Disease, Evolution, and Genetic Heterogeneity ». In : *Horizons in Biochemistry*. Sous la dir. de KASHA, M. et PULLMAN, B. New York : Academic Press, 1962, pp. 189-225 (citée p. 24).

Annexe D

Listes des figures et des tableaux

Sauf mention contraire, toutes les figures et tableaux présentés dans cette thèse sont de ma conception. Certaines images utilisées, sous license libre, proviennent de *Wikimedia Commons*. Les références précises de l'origine de ces images sont données dans le tableau ci-dessous. On trouvera dans les pages suivantes la liste des figures et la liste des tableaux de l'ensemble de la thèse.

Nom	Utilisation	Auteur-e	license
<i>Bromus erectus</i> ^a	Fig. 1.4 page 37	F. W. van Eeden	Domaine public
<i>Calluna vulgaris</i> ^b	Fig. 1.4 page 37	J. D. Reitter et G. F. Abel	Domaine public
<i>Pinus halepensis</i> ^c	Fig. 1.4 page 37	F. Bauer, D. Don et A. B. Lambert	CC-By ^d
<i>Crocus nudiflorus</i> ^e	Fig. 1.4 page 37	James Sowerby	Domaine public
<i>Robot metaphor</i> ^f	Fig. 5.3 page 163	Mecanopsis	CC-By-SA ^g

a. ↑ [https://commons.wikimedia.org/wiki/File:Bromus erectus — Flora Batava — Volume v20.jpg](https://commons.wikimedia.org/wiki/File:Bromus_erectus—Flora_Batava—Volume_v20.jpg)

b. ↑ [https://commons.wikimedia.org/wiki/File:Reitter-Abel Holzarten T089.jpg](https://commons.wikimedia.org/wiki/File:Reitter-Abel_Holzarten_T089.jpg)

c. ↑ [https://commons.wikimedia.org/wiki/File:A description of the genus Pinus \(Tab. XI\) \(7797055556\).jpg](https://commons.wikimedia.org/wiki/File:A_description_of_the_genus_Pinus_(Tab._XI_(7797055556)).jpg)

d. ↑ <https://creativecommons.org/licenses/by/2.0/deed.en>

e. ↑ [https://commons.wikimedia.org/wiki/File:Crocus nudiflorus \(Sowerby\).jpg](https://commons.wikimedia.org/wiki/File:Crocus_nudiflorus_(Sowerby).jpg)

f. ↑ [https://commons.wikimedia.org/wiki/File:Robot metaphor.png](https://commons.wikimedia.org/wiki/File:Robot_metaphor.png)

g. ↑ <https://creativecommons.org/licenses/by-sa/4.0/deed.en>

D.1 Liste des figures

1.1	Déroulement illustré d'une analyse comparative en linguistique historique	28
1.2	Déroulement illustré d'une analyse stématique	30
1.3	Déroulement d'une analyse de compatibilité des caractères.	33
1.4	Exemple d'analyse cladistique	37
1.5	Déroulement d'une analyse cladistique	38
1.6	Classification de méthodes de biogéographie des taxons	41
1.7	Déroulement illustré de l'inférence d'aire ancestrale par optimisation de Camin–Sokal	43
1.8	Déroulement illustré de l'inférence d'aire ancestrale par optimisation de Fitch	44
1.9	Déroulement illustré de l'inférence d'aire ancestrale par optimisation pondérée	45
1.10	Déroulement illustré de l'analyse de dispersion–vicariance (DIVA)	46
3.1	Règles de regroupement des unités étiques en unités émiques	109
3.2	Données contextuelles susceptibles d'attester la pertinence de données musicales	110
4.1	Comparaison des types d'échelles et des métricités de quatre pièces présentes dans plusieurs populations	138
5.1	Arbre des patrimoines musicaux de 58 populations du Gabon	160
5.2	Exemple de chaînes d'états	163
5.3	Illustration du principe de la marche guidée	163
5.4	Représentation de la vraisemblance de la chaîne pour chaque itération échantillonnée	166
5.5	Autocorrélation de la vraisemblance de chaque itération échantillonnée	166
5.6	Longitude et latitude de trois nœuds à chaque itération échantillonnée	167
5.7	Histogramme des valeurs de longitude et de latitude de trois nœuds	169
5.8	Localisation des positions inférées des 20 nœuds étudiés	170
5.9	Localisations les plus fréquemment échantillonnées pour le nœud <i>O1</i>	172
5.10	Localisation des différentes positions vraisemblables pour les 6 nœuds dont la position est incertaine	174
5.11	Diagramme de correspondance pour le nœud <i>O1</i>	178
6	<i>Fig 1. Geographical distribution of beat subdivision</i>	203
6	<i>Fig 2. Geographical distribution of rhythmic figures</i>	206
6	<i>Fig 3. Repartition of the pièces analyzed according to their scalar organization</i>	209
6	<i>S1 Fig. Proceedings of a recording session with the CRIP technique</i>	223
6	<i>S2 Fig. Location of the populations studied on the map of Gabon</i>	225
6	<i>S6 Fig. The three most frequents canonical interval codes sequences</i>	241

8.1	Exemple de passage entre deux cellules rythmiques	291
8.2	Influence de la subdivision de la pulsation sur le codage des formules rythmiques	293
A.1	Exemples d'application des conventions utilisées pour la notation des références	375
B.1	Distribution des localisations échantillonnées pour le nœud Bko	394
B.2	Distribution des localisations échantillonnées pour le nœud Bo	395
B.3	Distribution des localisations échantillonnées pour le nœud Bt	396
B.4	Distribution des localisations échantillonnées pour le nœud F	397
B.5	Distribution des localisations échantillonnées pour le nœud Ho	398
B.6	Distribution des localisations échantillonnées pour le nœud La	399
B.7	Distribution des localisations échantillonnées pour le nœud Ma	400
B.8	Distribution des localisations échantillonnées pour le nœud My	401
B.9	Distribution des localisations échantillonnées pour le nœud N	402
B.10	Distribution des localisations échantillonnées pour le nœud Ng	403
B.11	Distribution des localisations échantillonnées pour le nœud Ngo	404
B.12	Distribution des localisations échantillonnées pour le nœud Oba	405
B.13	Distribution des localisations échantillonnées pour le nœud Oi	406
B.14	Distribution des localisations échantillonnées pour le nœud Ol	407
B.15	Distribution des localisations échantillonnées pour le nœud Pa	408
B.16	Distribution des localisations échantillonnées pour le nœud Pu	409
B.17	Distribution des localisations échantillonnées pour le nœud Root	410
B.18	Distribution des localisations échantillonnées pour le nœud Tek	411
B.19	Distribution des localisations échantillonnées pour le nœud Ts	412
B.20	Distribution des localisations échantillonnées pour le nœud Wn	413
B.21	Dépendance à la position des autres nœuds pour le nœud Ma	415
B.22	Dépendance à la position des autres nœuds pour le nœud N	416
B.23	Dépendance à la position des autres nœuds pour le nœud Ol	417
B.24	Dépendance à la position des autres nœuds pour le nœud Pa	418
B.25	Dépendance à la position des autres nœuds pour le nœud Root	419
B.26	Dépendance à la position des autres nœuds pour le nœud Wn	420

D.2 Liste des tableaux

2.1	Publications utilisant des méthodes issues de l'inférence phylogénétique pour analyser des données culturelles	93
3.1	Comparaison des concepts de cognition musicale utilisés par différent-e-s ethnomusicologues africanistes français-e-s	127
5.1	Valeurs médianes des longitudes et latitudes inférées lors de l'analyse bayésienne	168
5.2	Localisations possibles des nœuds pour lesquels il existe une indétermination . .	175
5.3	Répartition croisée des itérations échantillonnées entre les positions alternatives de différents nœuds	179
6	<i>Table 1. Distinctive parameters values in a repertoire of a population</i>	204
6	<i>Table 2. Rhythmic figures shared by various populations</i>	207
6	<i>Table 3. Rhythmic figures found in the same repertoire from various musical patrimonies</i>	212
6	<i>S2 Table. Description of the populations</i>	224
6	<i>S3 Table. Elementary rhythmic cells</i>	226
6	<i>S4 Table. Dataset analysed</i>	228
6	<i>S5 Table. Result of the bootstrap test of low variability of a parameter inside a repertoire</i>	240
7.1	Métriques et échelles musicales de trois pièces dans différentes populations	271
7.2	Formules rythmiques caractéristiques de trois répertoires dans trois populations Myènè	272
8.1	Similarité des formules rythmiques en fonction des cellules rythmiques contenues	289
8.2	Similarité des formules rythmiques en fonction de l'ordre des cellules	290
9.1	Terminologies concernant la systématisation et la catégorisation musicale	333
A.1	Conventions utilisées dans ce travail pour la notation des références	375
B.1	Index des citations autour de la notion de modèle	378
B.2	Référence des figures de synthèse pour chaque nœud.	393

Annexe E

Table exhaustive des matières

Volume I : Introduction	7
Remerciements	7
Préambule	9
Introduction	13
1 L'inférence historique par comparaison des structures formelles	15
1.1 De quoi s'agit-il ?	17
1.2 Historique disciplinaire	18
1.2.1 En critique textuelle	18
Naissance de la stématisation	18
Formalisation des méthodes	19
Raffinements méthodologiques	19
1.2.2 En biologie	20
La conception évolutionniste et la filiation entre les espèces	20
Naissance de la systématique biologique	20
Premières méthodes d'inférence historique	20
Le développement des méthodes automatisées d'inférence	21
1.2.3 En linguistique	22
L'évolution et la filiation des langues	22
Premières propositions comparatistes	22
Le développement de méthodes rigoureuses	23
Le développement marginal d'autres méthodes	24
1.3 Description de méthodes d'inférence historique	24
1.3.1 Objectifs	24
Un travail préparatoire à l'analyse	24
Un outil pédagogique au service de l'interdisciplinarité	24
Une base de comparaison	25
1.3.2 Méthodes et limites	25

	Diversité des méthodes	25
	Choix des méthodes présentées	26
1.3.3	La méthode comparative (linguistique historique)	26
	Principe général	26
	Identification des cognats	27
	Établissement de correspondances phonologiques	27
	Reconstruction de la proto-phonologie	27
	Reconstruction des proto-morphèmes	29
1.3.4	Les méthodes stématisques	29
	Principe général	29
	Classification des lieux variants	29
	Construction de la chaîne reliant les manuscrits	31
	Nœuds intermédiaires et orientation de la chaîne	31
1.3.5	La compatibilité des caractères	31
	1.3.5.1 Principe	32
	1.3.5.2 Méthodes de recherche	32
	Cas de caractères binaires	32
	Cas des caractères à états multiples	34
1.3.6	Les méthodes cladistiques	34
	Principe général	34
	Choix des caractères	35
	Identification des états de caractères	35
	Polarisation des états de caractères	35
	Recherche de l'arbre le plus parcimonieux	36
1.3.7	Les méthodes probabilistes	36
	1.3.7.1 Les méthodes de maximum de vraisemblance	39
	1.3.7.2 L'inférence bayésienne	39
	Principe général	39
	Analyse bayésienne empirique	39
	Analyse bayésienne hiérarchique	40
1.3.8	La biogéographie historique des taxons	40
	1.3.8.1 Optimisation de Camin–Sokal	40
	1.3.8.2 Optimisation de Fitch	42
	1.3.8.3 Optimisation pondérée	42
	1.3.8.4 Analyse de dispersion–vicariance (DIVA)	42
1.4	Conclusion du chapitre	47
2	Les méthodes issues de l'inférence phylogénétique appliquées aux données culturelles	49
2.1	État de l'art disciplinaire	53
	2.1.1 Limites	53
	2.1.2 Critique textuelle	53
	2.1.2.1 État des lieux	53
	Les méthodes de distances	53
	Les distances utilisées	54
	La cladistique	55
	Autres méthodes	56
	2.1.2.2 Controverses et difficultés	56
	Différences dans les mécanismes d'évolution	56

	Les ancêtres observés	56
	Les multifurcations	56
	Le traitement des homoplasie	57
	La contamination	57
	La co-incidence	57
	L'identification et le traitement des caractères	57
	Le choix des caractères	57
	La pondération des caractères	57
2.1.2.3	Évaluer la pertinence des méthodes	58
	Justifications théoriques	58
	Justifications empiriques indirectes	59
	Comparaison avec les résultats de la stémato­logie	59
	Robustesse de l'analyse	59
	Utilisation d'indices statistiques	60
	Validation empirique directe	60
	Validation sur données artificielles	60
	Utilisation de critères externes de validation	61
2.1.3	Linguistique	61
2.1.3.1	Premières approches	61
	Divergence et convergence de la biologie et de la linguistique	61
	Méthode en réseau de Bandelt	62
2.1.3.2	Les méthodes de caractères	62
	Méthodes cladistiques	62
	Méthodes de compatibilité	63
	Réseau par minimisation des transferts horizontaux (MLN)	63
2.1.3.3	Méthodes probabilistes.	64
	L'inférence bayésienne	64
	Les migrations des populations	64
2.1.4	Anthropologie de la culture matérielle	65
2.1.4.1	Premières approches	65
	L'inférence historique en anthropologie	65
	Des travaux précurseurs, mais sans suite	66
2.1.4.2	Tester des hypothèses historiques	66
	L'étude de Collard et Shennan	66
	Ethnogenèse ou phylogenèse ?	67
	Autres hypothèses testées	68
2.1.4.3	Élargissement des approches	68
	Inférer les relations	68
	Évaluer ces outils d'inférence	70
2.1.5	Musicologie	70
2.1.5.1	Les travaux de Temkin	70
2.1.5.2	Les travaux de Toussaint et coll.	71
	Présentation des analyses	71
	En visant la classification	72
	En visant la reconstruction évolutive	73
2.1.5.3	Les travaux de Le Bomin et coll.	73
2.1.6	Autres applications	73
	Concepts scientifiques	74
	Traditions orales	74

	Économie	75
2.2	Considérations générales sur ces transpositions interdisciplinaires	75
2.2.1	Spécificités disciplinaires	75
	Diversité des utilisations	75
	Particularismes	75
2.2.2	La question de la transmission horizontale	76
2.2.2.1	Influence de la transmission horizontale	76
	Sur la topologie inférée	76
	Sur l'estimation des dates de divergence	78
	Dans les méthodes comparatives phylogénétiquement informées	78
2.2.2.2	Détection et mesure de la transmission horizontale	79
	À partir des arbres phylogénétiques inférés	79
	À partir d'indices statistiques	79
	Indices de cohérence et de rétention	79
	Score δ et score Q-résiduel	80
	Par visualisation des données	80
	Autres méthodes	81
2.3	Tableau comparatif de quelques études	81
2.3.1	Objectifs et limites de ce tableau	81
2.3.2	Présentation générale	82
2.3.2.1	Contenu du tableau	82
	Les lignes : des <i>analyses</i>	82
	Les colonnes : typologie des analyses	83
2.3.2.2	Choix des analyses	83
	1. Des méthodes issues de l'inférence phylogénétique	83
	2. Des données culturelles	84
	3. Des relations comme résultat	84
2.3.3	Description détaillée des champs du tableau	84
2.3.3.1	Objectifs des analyses	84
2.3.3.2	Données analysées	86
2.3.3.3	Typologie des méthodes utilisées	87
2.3.3.4	Éléments de caractérisation de ces méthodes	88
	Méthodes cladistiques	88
	Méthodes phénétiques	90
	Méthodes probabilistes	91
	Autres	91
2.3.3.5	Logiciels utilisés	92
2.3.4	Tableau comparatif	93
2.4	Conclusion du chapitre	103
3	Systématique musicale	105
3.1	Outils théoriques et méthodologiques	107
3.1.1	Les patrimoines musicaux des musiques du Gabon	107
	Pièces	108
	Répertoires	108
3.1.2	Le structuralisme : de la linguistique à l'ethnomusicologie	108
3.1.2.1	Phases de l'analyse	108
3.1.2.2	Identification des unités émiques	109
3.1.3	La modélisation de la systématique musicale vernaculaire	109

3.1.3.1	Modélisation	111
3.1.3.2	Validation	112
3.1.3.3	Techniques d'enregistrement analytique	112
3.1.4	La catégorisation des patrimoines musicaux	113
3.2	Des outils au service de la comparaison interculturelle	114
3.2.1	Modélisation des pièces musicales	114
3.2.1.1	Rigueur de l'analyse	115
	Objectivation	115
	Répétabilité	115
	Cohérence	116
3.2.1.2	Pertinence et finesse de l'analyse	116
	Pertinence	116
	Finesse	117
	Synthèse	117
3.2.2	Catégorisation musicale	117
3.3	Réflexions autour de la notion de <i>modèle</i>	118
3.3.1	La diversification progressive des concepts de cognition musicale	118
3.3.1.1	Simha Arom et la conceptualisation du modèle	118
3.3.1.2	Le substrat : pièces pluripartites et modèle polyphonique	120
3.3.1.3	Les séries hiérarchisées de modèles d'apprentissage	121
3.3.1.4	Vers une diversité des modèles cognitifs	122
3.3.2	Problèmes et perspectives des concepts utilisés	123
3.3.2.1	Pour une ethnomusicologie cognitive	123
3.3.2.2	Modèles instauratif et scientifique	124
3.3.2.3	Perspectives : une approche scientifique, matérialiste et populationnelle	125
3.3.3	Tableau comparatif des concepts de cognition musicale	127
3.4	Conclusion du chapitre	128
4	Musique, comparatisme et histoire	129
4.1	Quelle place pour la comparaison ?	130
	Une musicologie de la comparaison	130
	La comparaison n'est pas la spécificité de la discipline	130
	Critiques et rejets de la comparaison	131
4.2	La comparaison dans la pratique de l'ethnomusicologie	132
4.2.1	Le courant évolutionniste	132
4.2.2	Le courant diffusionniste	133
4.2.3	L'étude des aires musicales	134
4.2.3.1	Études régionales	134
4.2.3.2	Le projet <i>Cantometric</i>	134
4.2.4	Diversité des comparaisons	135
4.2.4.1	La comparaison diachronique	136
4.2.4.2	Les familles de pièces	137
4.2.4.3	L'étude de la diversité culturelle	138
4.2.4.4	La comparaison des comportements musicaux	138
4.3	La comparaison musicale pour l'inférence historique	139
4.3.1	Ethnomusicologie et histoire	139
	Une discipline de la synchronie	139
	Histoire des musiques	139

La musique porteuse d'histoire	139
4.3.2 L'histoire inscrite au cœur de la musique	140
L'histoire dans les premières comparaisons ethnomusicologiques	140
Histoire de pièces	141
Traces historiques entre cultures proches	141
Méthodes issues de l'inférence phylogénétique	142
Musique et migrations	142
4.4 Conclusion du chapitre	143
Conclusion de la première partie	145
Volume II : Résultats — Discussion	153
Résultats	153
5 Essai d'inférence phylogéographique	155
5.1 Contexte et objectifs	157
5.2 Méthodes	158
5.2.1 L'arbre et les nœuds étudiés	158
5.2.2 Le modèle d'évolution	159
5.2.3 La méthode de MCMC	162
5.2.3.1 Chaîne d'états	162
5.2.3.2 Marche aléatoire guidée	162
5.2.3.3 Allumage de la chaîne	164
5.2.3.4 Échantillonnage de la chaîne	164
5.3 Résultats	165
5.3.1 Vérifications	165
5.3.1.1 Convergence	165
5.3.1.2 Autocorrélation	165
5.3.2 Localisations inférées des nœuds	168
5.3.3 Cas d'incertitudes	168
5.3.3.1 Distribution dans les itérations échantillonnées	168
5.3.3.2 Association entre longitude et latitude d'un même nœud	171
5.3.3.3 Caractérisation des positions alternatives de chaque nœud	171
5.3.3.4 Associations entre les positions alternatives de différents nœuds	176
5.4 Discussion	180
5.4.1 Méthodologie	180
5.4.1.1 Autocorrélation et échantillonnage	180
5.4.1.2 Surparamétrisation	180
5.4.1.3 Arbre des relations entre patrimoines	180
5.4.2 Limite du modèle utilisé pour l'inférence	181
5.4.2.1 Le modèle arborescent d'évolution des patrimoines	181
5.4.2.2 Le patrimoine musical comme unité évolutive	181
5.4.2.3 Patrimoines musicaux et populations	182
5.4.3 Pertinence de l'analyse	182
5.4.3.1 Tester une méthode	183
5.4.3.2 Produire de la connaissance historique	183
5.5 Perspectives	184

5.5.1	Résumé du chapitre	184
5.5.2	Mieux connaître nos données	184
5.5.3	Affiner les fondements théoriques et méthodologiques	185
6	Un aperçu de la diversité des musiques du Gabon	187
6.1	Introduction	188
6.1.1	Une thèse avec article	188
6.1.1.1	Objectif de cette introduction	188
6.1.1.2	Un format adapté aux contraintes	188
6.1.1.3	Limites	189
6.1.2	Contexte et objectifs	189
6.1.2.1	Contexte	189
6.1.2.2	Objectifs	189
6.2	<i>Diversity and distribution of musics in Gabon</i>	191
	<i>Introduction</i>	192
	<i>Material and methods</i>	196
	<i>Results</i>	201
	<i>Discussion</i>	213
	<i>Conclusion</i>	219
	<i>Supporting information</i>	221
	<i>References</i>	241
6.3	Conclusion et perspectives	245
	Principaux résultats	245
	Perspectives	245
7	Quelles données, quelle(s) histoire(s) ?	247
7.1	Positionnement épistémologique	249
	Sciences nomothétiques et idéographiques	249
	Instances et types	249
	Inférence historique et mécanismes évolutifs	250
	Une approche historique	250
7.2	Croisements d'histoires	251
7.2.1	Histoires musicale et génétique.	251
7.2.1.1	Histoires linguistiques et génétiques.	251
7.2.1.2	Revue détaillée de quelques études	252
	Les travaux de Grauer (2006)	252
	Les travaux de Pamjav et coll. (2012)	252
	Les travaux de Brown et coll. (2014)	253
	Les travaux de Savage et coll. (2015)	254
7.2.1.3	Que nous apprennent ces comparaisons?	255
	Corrélation gènes/musiques et qualité du marqueur historique	255
	Gènes et musiques comme sources historiques complémentaires	258
7.2.2	Histoires musicale et culturelle	260
7.2.2.1	Différentes histoires culturelles.	260
7.2.2.2	Musiques et cultures	262
7.2.3	Histoires musicales multiples.	262
7.2.3.1	Exemples génétiques et culturels	263
	Hétérogénéité du génome	263
	Hétérogénéité culturelle	263

	En linguistique	263
	En anthropologie culturelle	263
7.2.3.2	Hétérogénéité du musical	264
7.2.4	Histoire musicale et histoire des populations	266
	Ethnie, groupe, nation, race, identité, communauté... population ?	266
	La musique, élément de définition de la population diachronique	267
7.3	Les indicateurs de choix des marqueurs historiques	268
7.3.1	La notion de potentiel historiographique	268
7.3.2	Indicateurs <i>a priori</i>	269
7.3.2.1	Savoirs théoriques	269
7.3.2.2	Savoirs empiriques	270
	Stabilité des paramètres musicaux	270
	Corrélation avec une histoire connue	271
7.3.3	Indicateurs <i>internes</i>	272
7.3.4	Indicateurs <i>a posteriori</i>	273
7.3.5	Comparaison et limites de ces indicateurs	273
	Indicateurs <i>a posteriori</i>	273
	Indicateurs internes	273
	Indicateurs <i>a priori</i>	273
7.4	Conclusion du chapitre	274
Conclusion de la deuxième partie		275
Discussion		277
8 Pistes de recherches		279
8.1	Comprendre la transmission et la transformation des musiques	280
8.1.1	Observer le changement des pratiques musicales	281
	Enregistrements d'archives	281
	Problèmes des archives	281
	Collectage <i>ad hoc</i>	282
8.1.2	Comprendre les dynamiques de la transmission musicale	283
8.1.2.1	Nécessités de s'intéresser à la transmission	283
	Comprendre la transmission pour mieux inférer l'histoire	283
	Des transmissions différentes selon les paramètres	283
	Influence du type de transmission sur le musical	284
8.1.2.2	Deux exemples d'études de la transmission	284
	Étudier la transmission au sein d'une culture musicale	284
	L'emprunt entre différentes cultures musicales	285
8.1.3	Étudier la généralité et la portée de ces mécanismes	285
8.1.4	Relier l'observé au passé	286
8.2	Approfondir l'étude des formules rythmiques	287
8.2.1	Accroître le jeu de données	288
8.2.2	Affiner la comparaison des formules rythmiques	288
8.2.2.1	Quantifier la ressemblance	288
	Cellules rythmiques contenues	289
	Similarité entre cellules rythmiques	290
8.2.2.2	Décrire la rythmique indépendamment de la métrique	291

	Une même pulsation avec plusieurs subdivisions possibles	291
	Plusieurs pulsations possibles	292
9	Réflexions sur la pratique de la recherche	295
9.1	La relation aux données	297
9.1.1	Quand commence l'archive ?	297
	Un jeu de données homogène	297
	Des données de seconde main	298
9.1.2	De nombreux implicites	298
9.1.2.1	Implicite car évident	298
9.1.2.2	Collecter pour soi	299
9.1.2.3	Implicite car ineffable	299
9.1.3	Travailler avec l'implicite	299
9.1.3.1	De l'ignorance à la méfiance	299
9.1.3.2	Directrice et informatrice	300
9.2	Le rôle d'analyste	301
9.2.1	Analyser des données préparées par d'autres	301
	Une relation interdisciplinaire	301
	Une interdisciplinarité asymétrique	301
9.2.2	Un analyste déconnecté du terrain	302
9.2.2.1	Scinder l'étape d'analyse ?	302
	La transcription est analyse	302
	L'analyse doit être globalement cohérente	302
9.2.2.2	Séparer collecte et analyse ?	303
	Spécialisation de l'analyste	303
	L'analyse liée au terrain	303
9.3	Limites de l'interdisciplinarité	304
9.3.1	Les « points aveugles » de la coopération interdisciplinaire	304
	Des zones délaissées	304
	Occultées par les différences disciplinaires	304
9.3.2	Le besoin d'unifier	305
	De la coopération à l'intégration	305
	Des ressources pour le travail interdisciplinaire	305
9.3.3	L'individualisation de la discipline	306
9.4	La transposition de méthodes	306
9.4.1	Légitimité de la transposition	307
9.4.1.1	Des raisons	307
	L'analogie a une valeur heuristique	307
	Sur les conditions de naissances de l'analogie	308
9.4.1.2	Des justifications	309
9.4.2	Évaluation	310
9.4.2.1	De la validation rationnelle	311
9.4.2.2	Des transformations induites	311
	Conceptions portées par la méthode	311
	Compétences requises	312
9.4.2.3	De la forme	312
	Conclusion générale	315

Volume III : Glossaire — Annexes	327
Glossaire argumenté	327
Annexes	370
A Conventions	373
A.1 Citations et traductions	373
A.2 Sources et références bibliographiques	374
B Suppléments	377
B.1 Chapitre 3 : Citations autour de la notion de <i>modèle</i>	378
B.2 Chapitre 5 : Inférence phylogéographique	387
B.2.1 Fichiers informatiques	387
B.2.1.1 Coordonnées des taxons de l'arbre	388
B.2.1.2 Topologie de l'arbre	389
B.2.1.3 Fichier d'instructions pour BayesTraits	391
B.2.2 Synthèse par nœud	393
B.2.3 Dépendance des positions incertaines de différents nœuds	415
B.3 Test de Mantel	421
B.3.1 Le test de Mantel	421
La statistique t	421
Le test statistique	421
B.3.2 Dérivés du test de Mantel	422
Test partiel de Mantel	422
Corrélogramme de Mantel	422
B.3.3 Critiques	422
Généralité	422
Performances	422
Procédure de permutation	423
C Bibliographie	425
C.1 Publications consultées	425
C.2 Autres publications citées	445
D Listes des figures et des tableaux	457
D.1 Liste des figures	458
D.2 Liste des tableaux	460
E Table exhaustive des matières	461