

Emergence forte *versus* émergence faible en écologie linguistique : le cas du mazatec

1. Introduction

Le mazatec est une langue otomangue du sud-est du Mexique, parlée par plus de 220 000 locuteurs, dans un biotope tropical fortement diversifié, sur trois paliers écologiques auxquels correspondent trois systèmes de cultures postcoloniales : basses terres (Ixcatlán et Soyaltepec : système canne à sucre), terres moyennes (Jalapa de Diaz : élevage), et hautes terres (Huautla : système café, en crise depuis plus de 30 ans). A cette tripartition s'ajoute le Cañon tropical de Cuicatlán (avec le bourg mazatec de Chiquihuitlán), appelé le plus souvent *la Cañada*, que caractérisent sa densité de microclimats ainsi que le contact de langues et de sociétés méso-américaines (mazatec, cuicatec, mixtec, chinantec, etc.), classée réserve de la biosphère (*Reserva de la Biosfera Tehuacán-Cuicatlán*, qui s'étend sur une surface de 490 817 hectares).

A l'aide de bases de données environnementales et socioéconomiques d'une part, dialectologiques d'autre part, il s'agit de faire ici apparaître dans un premier temps les dynamiques d'émergence de la variation dialectale du mazatec au cours de l'histoire, en tant que miroir de l'interaction entre nature et culture. Dans une telle approche, l'émergence forte d'une relation adaptative culture/nature (où la *culture* s'adapte à la *nature*) conduit à l'auto-organisation d'un flux de variation dialectale des différentes communautés de peuplement coordonné entre biotopes en relation de complémentarité. Dans un deuxième temps, nous tenterons de montrer que cette dynamique de diversification interne de la langue et de la culture se trouve reléguée à une dynamique d'émergence faible, tandis que c'est la relation nature/système (où c'est la *nature* désormais qui doit s'adapter à des *systèmes productifs*) qui devient dominante, menant à l'érosion des sols et à la raréfaction ou à l'épuisement des ressources naturelles, ou à des transformations massives (barrages hydro-électriques) en fonction de la survalorisation des exigences du système productif. Or, cette « seconde vague » ou ce nouveau cycle de différenciation dialectale est non plus tant géolinguistique (ou *diatopique*), comme par le passé, que sociolinguistique (ou *diastratique*). Les enquêtes de l'ALMaz (Atlas Linguistique Mazatec) menées depuis 2010 attestent un éventail très diversifié, en granularité fine, de la variation géolinguistique. En somme, l'aire mazatèque, qui compte de nombreux centres urbains de plus de 1000 habitants¹, est parcourue d'un grand nombre de variables sociolinguistiques, au même titre que des centres urbains nord-américains (Labov, 2007)². Nous l'aborderons en ayant à l'esprit en permanence l'*archipel vertical* de John Victor Murra (1956, 1985), et une synthèse récente sur l'ethnographie et la (géo)linguistique en milieu de type alpin, par Johanna Nichols (2004).

Dans la section 2 nous présenterons les méthodes d'analyse quantitative ou taxinomique (modes de classements) retenues pour l'examen des données de Kirk de manière aussi exhaustive que possible. La section 3 regroupera l'essentiel des résultats de nos analyses, et confrontera de multiples modèles (phylogénétique, ontogénétique et épigénétiques, ou *triangle de la complexité*). La section 4 présentera les conclusions et les perspectives, montrant dans quelle mesure un vaste horizon encore inexploré de normes modernes du mazatec est observable dans une aire dont nous aurons pourtant cru faire le tour par notre triple modélisation. Les exigences de vulgarisation nous ont conduit à ne pas présenter de données, afin de ne rendre compte que des résultats, mais le lecteur pourra se référer à diverses

¹ Huautla de Jiménez compte pas moins de 30 400 habitants, Santa María Chilchotla 20 584 et San José Tenango 18 000 : ces trois centres regroupent 54% de la microrégion n°13, qui couvre les hautes terres mazatèques

² Voir, pour plus de détails, Léonard & al. 2012 concernant les critères géolinguistiques ; Léonard, 2013 et Léonard & Fulcrand, 2013, pour l'énumération des variables structurales.

publications mentionnées au cours de l'argumentation, pour connaître de plus près les faits de langue qui justifient nos analyses (cf. Léonard 2010, Léonard & al. 2012, Léonard & Kihm 2012, Léonard, 2013, Léonard & Fulcrand, 2013, etc.).

2. Complexité et modélisation : méthodes

Nous utiliserons ici principalement deux méthodes de « calcul » ou de mesure de la distance : les distances dites « patristiques » de la cladistique d'une part, l'algorithme de Levenshtein d'autre part. Les sections 2.1 et 2.2 présentent brièvement les spécificités de chacune de ces deux méthodes.

2.1. La cladistique

La cladistique est une technique de classification couramment utilisée par les généticiens et les biologistes. Le choix de cette méthode ne préjuge pas de l'isomorphie entre faits de langues et phénomènes de l'ordre du vivant, comme le voulait la linguistique de la deuxième moitié du 19^{ème} siècle, quand August Schleicher reprenait les principes glanés au fil de la lecture de *l'Origine des espèces*, de Charles Darwin, pour fonder une théorie évolutionniste en linguistique génétique et typologique. Notre perspective est en effet résolument non organiciste (la langue n'est pas un organisme vivant), et non essentialiste (la cladistique intervient ici non pas pour édifier des aires ou des « dialectes » associés de manière inhérente à des territoires, mais pour construire des objets de connaissance à géométrie variable). Notre objectif est d'utiliser les résultats du traitement automatisé des faits de langue (la complexité algorithmique, cf. O'Sullivan, 2004) afin de construire des objets de connaissance complexes, ou multiplexes, susceptibles d'interprétations vicariantes (multiples points de vue ou points de référence variables, quoiqu'empiriquement fondés).

La méthode cladistique a été développée, à partir des principes élaborés par Hennig (1950,1966), pour rendre compte de l'évolution phylogénétique des espèces et reconstruire leur parenté. Le principe fondamental consiste à tirer parti de considérations sur l'évolution des caractères, en distinguant, pour chacun d'eux, une forme ancestrale et une ou plusieurs formes dérivées. Deux espèces sont dites descendantes d'un même ancêtre commun lorsqu'elles partagent les mêmes états dérivés de ces caractères. En revanche, le partage de caractères dans leur forme ancestrale ne prouve pas nécessairement la parenté. Cette méthode est en parfaite adéquation avec la recherche des affinités linguistiques qui fonde l'apparementement entre deux langues ou entre deux formes dialectales, et leur rattachement à la même proto-langue, sur le fait qu'elles ont hérité de mêmes traits linguistiques dérivés.

Une fois ce principe posé, il s'agit de proposer une procédure qui permette de l'appliquer, le but étant d'obtenir une représentation arborescente des espèces (des langues ou des dialectes), qui rende compte de leurs liens de filiation. Dans un monde idéal, la construction d'un tel arbre, le « *perfect tree* » ou l'arbre parfait ne devrait poser aucune difficulté. Malheureusement les difficultés surgissent du fait que certains traits dérivés ont pu retourner à leur forme ancestrale au cours de l'évolution (*réversion* ; les philologues utilisent le terme de *régression*), d'autres traits dérivés ont pu être acquis par convergence de manière indépendante chez deux espèces, et non par héritage, ou provenir d'un transfert d'une espèce à l'autre (transfert horizontal de gènes, emprunt dans le cas de langues). De tels événements, dénommés homoplasies, viennent perturber la véritable information sur les parentés entre espèces.

La procédure mise en place consiste donc à rechercher un arbre qui contienne le moins d'événements homoplasiques. Cela revient, selon le principe de parcimonie, à minimiser le nombre de changements d'états de caractères dans l'arbre. Cette procédure suppose, en revanche, que les traits étudiés évoluent indépendamment les uns des autres et que les événements de réversion (ou régression) et de convergence soient rares³.

Plusieurs programmes informatiques (Felsenstein, 1989, Swofford, 2002) permettent de mettre en application ces principes et de trouver l'arbre le plus parcimonieux, tout en tenant compte des diverses hypothèses que le biologiste (ou le linguiste) a pu formuler sur les modes d'évolution et de changement des caractères.

L'analyse des données mazatèques a porté sur des séries de cognats, sur trois séries de données : 1) phonologie diachronique (cf. Léonard, 2010, Léonard & al. 2012), 2) phénomènes de joncture morphologique entre préverbes et racines, ou processus morphologiques de la flexion verbale, 3)

³ Pour plus de détails, voir Darlu et Tassy (1993) ou Felsenstein (2004).

morphologie synchronique, ou contraintes de sélection de préverbes dans les radicaux (cf. Léonard & Kihm 2012).

2.2. Distance d'édition (algorithme de Levenshtein)

L'algorithme de Levenshtein est désormais tellement connu et utilisé en taxinomie quantitative des langues du monde qu'il suffira ici de se référer aux contributions de Beijering & al. 2008 ; Bolognesi & Heeringa, 2002 sur le plan du procédé et, dans une certaine mesure, de Brown & al. 2008 ; Wichmann & al. 2010, sur le plan des applications. Il importe de souligner que nous sommes convaincus qu'il est hasardeux d'utiliser cet algorithme à l'échelle inter-groupes (comme le font Brown & al., *op. cit.*), car les résultats du traitement de nos données mazatèques de phonologie comparative avec les données de Paul L. Kirk indiquent clairement que la valeur heuristique des résultats se lit en fonction du niveau le plus superficiel de variation linguistique (la phonologie, ou systèmes de sons, catégories sonores allophoniques *versus* phonémiques). A ce titre, la présente contribution apportera des éléments afin d'améliorer la grille de lecture des résultats de l'analyse de données dialectales ou linguistiques à l'aide de cet artefact. La distance d'édition, que mesure cet algorithme, compare les caractères de la chaîne parlée (en réalité, de la chaîne transcrite) un à un dans une logique d'alignement, en calculant la différence entre les caractères (*ts* contre *s*, ou *b* contre *v*, etc. correspondant implicitement à une désaffrication ou à une spirantisation pour le phonologue, et correspondant à une distance manifeste pour l'analyseur automatique qu'est l'algorithme). Il n'est pas nécessaire d'indexer identité contre différence cellule par cellule dans la matrice, comme le faisait la dialectométrie classique (cf. Goebel, 1981, 1982, 2002) : l'algorithme calcule les différences automatiquement, pour arriver pratiquement au même résultat, à la différence que, comme dans le traitement cladistique des données dans des graphes de dérivations et des matrices d'appariement, on peut guider sa lecture en fonction de la nature et du poids des variables structurales, en amont. L'algorithme de Levenshtein (tout comme la cladistique pour le traitement des « caractères ») prend toujours en considération l'ensemble du lexème, donc chaque élément de sa forme phonologique, y compris l'accent tonique ou les tons (s'ils sont marqués dans la transcription) et ce, de manière automatique. Il n'y a donc pas de taxa à préparer à l'avance, ce qui permet une rapidité exceptionnelle de la procédure d'analyse (sauf cas d'incompatibilité avec certains logiciels de classement de données, mais c'est au praticien de faire ses choix en ce domaine).

2.3. Complexité, modélisations multiplexes

Les deux méthodes – cladistique et calcul de Levenshtein – sont donc hautement complémentaires, à la différence qu'à notre avis, la méthode cladistique permet d'obtenir des résultats davantage exhaustifs sur le plan de la complexité des structures, dans une perspective intercomposante (phonologie diachronique, morphophonologie et morphologie/lexique, en termes de construction des radicaux verbaux), alors que l'algorithme de Levenshtein, dans cette étape du moins du traitement des mêmes données (Kirk, *op. cit.*), donne des résultats certes intéressants, mais principalement heuristiques en ce qui concerne les structures de variation sur le plan de la phonologie.

Notre objectif méthodologique (mais également épistémologique) est d'associer ces méthodes de calcul dans un esprit complexionniste (sciences de la complexité, systèmes complexes, cf. Gribbin 2004, Massip-Bonet & Bastardas-Boada, 2013). De ce point de vue, nous distinguerons trois ordres de complexité, comme le suggère O'Sullivan (*op. cit.*) : la complexité *déterministe*, la complexité des *agrégats communaux* et la complexité *algorithmique*. Nous venons d'illustrer cette dernière dans les sections 2.1 et 2.2. La *complexité déterministe* pour nous traite des faisceaux de causalité externe à la structuration ou à l'émergence de complexes observables, comme les aires dialectales. De ce point de vue, nous sommes bien conscients que les structures émergentes ne dérivent pas d'une seule cause (par exemple climatique, orographique, hydrographique), mais de la convergence de déterminismes multiples, dont les effets sont plus ou moins intenses (émergence forte si un faisceau de causes est susceptible d'expliquer un phénomène émergent, émergence faible si les observables ne sont qu'un épiphénomène, au terme d'une turbulence ou de transitions de phases). La *complexité des agrégats communaux* selon nous correspond à la diversité interne et aux hiérarchies de ressources (échanges, réciprocité, complémentarité), ou de pouvoir (hégémonie, dépendance, etc.) entre groupes humains, sur des bases communales (les communautés ou sociétés locales, organisées selon des modes divers en termes d'alliances, de systèmes de parenté et d'endo/exogamie, etc.).

3. Systèmes, indices, interactions, aires de résonance ou de réfraction

Depuis soixante ans les populations autochtones de la région mazatèque affrontent une crise profonde et multiforme, à tous les étages du biotope complexe que nous venons de décrire. La structure agraire de *microfundios* (petite propriété, polyculture) et les réseaux familiaux et de solidarité (les « loyautés primordiales » selon les anthropologues, cf. Boege, 1988), mais aussi l'accès à la terre et les fondements même de la territorialité des populations autochtones ont été défaits dans les basses terres par la construction d'un barrage hydroélectrique⁴ et l'essor de structures latifundiaries liées à la monoculture de la canne à sucre (grandes propriétés, agro-industrie). Dans les terres moyennes, c'est le complexe agro-industriel lié à l'élevage intensif qui a remis en cause les liens sociaux et territoriaux. Enfin dans les hautes terres, l'effondrement du *système café* (plantations de café et circuits commerciaux liés à cette production) dans les années 1980 a recentré l'économie des centres urbains vers le commerce et les services (principalement dans le secteur éducatif), dans une région désavantagée par le relief. Le *système milpa* (l'agriculture traditionnelle, fondée sur la polyculture familiale de subsistance) et le marché du travail local ne suffisent plus à maintenir la population sur place, qui tend à émigrer vers les centres urbains ou vers l'Amérique du nord.

Nous étudierons aussi bien les facteurs adaptatifs entrant en jeu dans la relation de l'homme au milieu naturel que les indices de crise environnementale dans la région mazatèque, à travers les données statistiques disponibles, que nous qualifierons en fonction des observations de terrain obtenues au cours des enquêtes menées pour la réalisation de *l'Atlas Linguistique Mazatèque* (ALMaz), dans le cadre du projet MAmP de l'IUF (2009-14). Nous montrerons en quoi les tendances menant à la dégradation de la relation de l'homme au milieu naturel dans cette région d'un haut degré de diversité biologique sont généralisables à un grand nombre de situations de « langues en danger », en Méso-Amérique et ailleurs dans le monde. Nous analyserons les étapes successives de cette dérive dans une perspective historique, en utilisant les concepts et les méthodes de la théorie des systèmes complexes. L'émergence sera ainsi caractérisée de manière dialectique et cyclique. Les conditions d'auto-organisation et la marge d'agentivité des sociétés autochtones modernes (*accommodation, résilience, vicariance, etc.*) seront évaluées en termes de renforcement et d'opposition entre émergence forte et émergence faible de modalités écoculturelles dans la relation de l'homme à son milieu naturel.

Il va de soi pour nous, conformément à l'approche en termes de systèmes complexes, que cette périodisation et ces phases ne constituent qu'un modèle possible parmi d'autres. En aucun cas nous ne souhaitons poser en 3.1 des construits essentialistes, ou des totalités. Ce que nous écrivons ici n'a rien d'une profession de foi gratuite, puisque nous montrerons bientôt, à l'aide de la modélisation par algorithme de Levenshtein, que l'on peut donner une toute autre vision ou représentation des configurations géolinguistiques du mazatec, en relation avec l'effet fort de l'adaptation au milieu (les biotopes condensant des agrégats humains différenciés).

3.1. Contextualisation : topographie et réseau dialectal

Les figures 1.1 à 1.4. présentent d'une part les localités retenues pour l'analyse, en fonction des données dialectologiques existantes (fig. 1.1-2), d'autre part une périodisation de l'expansion du mazatèque dans cette grille géographique (fig. 1.3) et les attributs écohumains observables au cours du siècle passé (fig. 1.4). Une topographie plus détaillée, fondée sur une grille géodésique, est donnée en annexe dans les cartes 1 et 2 (occupation des sols, ou *pédologie* d'une part, démographie et voies de communication d'autre part). La figure 1 énumère les douze municipalités qui ont fait l'objet de l'enquête réalisée par Paul Livingston Kirk dans la première moitié des années 1960, dans la continuité des travaux de Gudschinski (1958a & b)⁵, à l'occasion de sa thèse de doctorat (Kirk, 1966). Ces données ont été saisies et transformées en une base de donnée géolinguistique dans le cadre du projet

⁴ La Presa Miguel Alemán, mise en service en 1959, s'étend sur 47 800 hectares de terres engouties, avec une capacité de 8 millions de m³, 725 million kilowatts/an de la centrale hydroélectrique de Temascal: 22 000 paysans indigènes furent alors réinstallés dans l'Etat voisin de Veracruz. Si l'on tient compte que le barrage en question est relié par un canal à une autre structure du même type, la presa Cerro de Oro, les proportions s'amplifient pour atteindre un volume de 13,380 millions de m³.

⁵ V. aussi Rensch, 1976, en particulier p. 87-116 pour une profondeur temporelle et comparative plus ample.

MamP/MPMA de la chaire IUF 2009-14 de (morpho)phonologie méso-américaine⁶. Elles ont été abondamment complétées depuis 2010 par des enquêtes de terrain de vérification, dans les mêmes localités, ainsi que par des enquêtes complémentaires dans une dizaine d'autres localités (Santa Ana Ateixtlahuaca, San Francisco Huehuetlán, Santa Maria Chilchotla, San Antonio Eloxochitlán, San Lucas Zoquiapam, San Mateo Yoloxochitlán, Santa Cruz Acatepec, San Antonio Nopalera, et divers hameaux périphériques de centres comme Huautla et Mazatlán, comme San Andrés Hidalgo).

San Lorenzo Cuaunecuiltitla	Huautla de Jiménez			San Pedro Ixcatlán	San Miguel Soyaltepec
San Jerónimo Tecoaatl					
Mazatlán Villa de Flores	SM Jiotes	San Miguel Huauteppec	Ayautla	San Felipe Jalapa De Diaz	
			Chiquihuitlán	Santo Domingo	

Figure 1.1. Localités (et lococolectes) représentées dans les données de Kirk (1966)

La disposition dans les cellules de la série de figures 1.1-4 reprend de manière schématique la disposition des localités dans un espace écohumain qui s'étage sur quatre paliers: les hautes terres (fig. 1.2 Lo, Te à l'ouest, Hu au centre, Ji et Mg sur le versant centre-oriental), les terres moyennes, ou piémont et plaines (Ay, Ja, Do), les terres basses (Ix, So) et deux contreforts, constituant deux versants, occidental et oriental, du Canyon tropical de la vallée de Cuicatlán, ou *Cañada* – respectivement, Mazatlán et Chiquihuitlán (Mz et Cq).

Lo	Hu			Ix	So
Te					
Mz	Ji	Mg	Ay	Ja	
			Cq	Do	

Figure 1.2. Abréviations des lococolectes de Kirk (1966), correspondant aux localités de la figure 1.1.

5	2c			4	2d
1c					
1b	2b	2a	3a	3b	
			1a	3c	

Figure 1.3. Divisions dialectales et scénario de diversification du diasystème mazatec

Lo Versant <i>poblano</i>	Hu		Ix	
--	-----------	--	-----------	--

⁶ V. la page en cours de construction depuis octobre 2012, sur le lien <http://jll.smallcodes.com/home.page>.

(café)	Hautes terres centrales (café)			<i>Cuenca</i> , versant Occidental (microfundio > canne à sucre)	So <i>Cuenca</i> , versant Oriental (microfundio > canne à sucre)
Te Versant hautes terres Occidentales (café)		Mg Versant centre-oriental 2 (milpa)	Ay Piémont (café, milpa)		
Mz <i>Cañada</i> Occidentale (milpa > café)	Ji Versant centre-oriental 1 (milpa)		Cq <i>Cañada</i> orientale (milpa)	Do Rivière SD (élevage, milpa)	

Figure 1.4. Divisions écologiques et spécialisations.

L'état initial (ou du moins, les conditions préalables à la période que nous analysons ici, qui couvre environ un millier d'années, depuis le début de la fragmentation dialectale), sur le plan agraire, à l'époque « précolombienne » ou précoloniale, est caractérisé par la prépondérance de la *milpa* et des vergers, des terres hautes aux terres basses et, de manière générale, par une constante interaction et des échanges de longue durée en ingénierie agraire, entre les différents niveaux de ce que J. V. Murra appelle l'*archipel vertical* (cf. Blanton & al. 1993). A partir de l'indépendance du Mexique (1821) commence l'implantation et l'expansion de cultures commerciales nouvelles dans la région, telle que la culture du café. La polyculture vivrière de type *milpa* doit céder devant la monoculture d'exportation. Une monoculture organisée sur la base de la grande propriété foncière – le *latifundio*. Les basses terres seront plus longtemps préservées, et pourront maintenir le régime agraire de la petite propriété individuelle : le minifundio. La carte 1 infra signale les aires de peuplement mazatec et les voies de communication. La carte 2 rend compte de la diversité du relief et des sols.

3.2. Relation culture/nature : phases d'expansion et de diversification du mazatec

Revenons à la phase d'*effet fort* de la relation homme/nature ou culture/nature : celle qui se matérialise, pour le linguiste et l'ethnohistorien cherchant des traces de l'histoire des mouvements de population à travers les indices linguistiques, dans la mesure où tout se passe comme si l'adaptation de communautés humaines à des niches écologiques était corrélée à une diversification de traits culturels, tels que la langue, envisagée dans ses trois composantes formelles : phonologie, morphologie et morphosyntaxe et lexicale. L'analyse cladistique des données de Paul L. Kirk (1966), comprenant plus de 750 *cognats* (ou correspondances lexicales permettant la comparaison, comme *lune*, *luna*, *lua* du français, de l'espagnol ou de l'italien et du portugais, respectivement, sont autant de *cognats* romans issus de l'étymon LUNAM) partagés par les douze lococlectes (ou variétés dialectales locales, dont la liste est donnée dans les figures 1.1 et 1.2 supra) retenus par Kirk pour son analyse visant à reconstruire les *proto-formes* (ou étymons abstraits issus de la comparaison interdialectale, faute d'attestations écrites anciennes)⁷. La figure 1.3 propose une synthèse des résultats de l'analyse cladistique⁸ des données en

⁷ Voir aussi Gudschinsky, 1948a, b pour une approche de la variation dialectale du mazatec. L'article de 1948 paru dans *Language* pour une approche ethnohistorique en miniature de la variation dialectale mazatèque reste un jalon des études ethnohistoriques appuyées sur des données dialectales, dans l'histoire de la géolinguistique. Article peu connu des dialectologues européens, mais qui confirme, si besoin était, que le décloisonnement géographique est de rigueur en dialectologie, au plus grand bénéfice des sciences de l'homme.

⁸ Les moments cladistiques ou phases de la figure 1.3 sont le résultat d'un traitement cladistique des données de Kirk (1966) réalisés par Jean Léo Léonard et Antonella Gaillard-Corvaglia (postdoctorante à l'Inalco en 2011) et Pierre Darlu (INserm, CNRS), based on Paul L. Kirk's cognate sets for Mazatec varieties (1966). On trouvera les détails sur la méthode mise en oeuvre dans Léonard, Dell'Aquila & Gaillard-Corvaglia, 2012. Le traitement

question, sous forme de périodisation par *phases* ou *moments* (les moments *cladistiques*) du processus de fragmentation dialectale du mazatec. Cette figure 1.3 est conçue comme une sorte de scénario chronologique, en cinq phases :

Phase 1 (1a-c) : intrusion des populations mazatèques dans le bassin du Papaloapam, à travers la *Cañada* de Cuicatlán. Les dialectes correspondant aux segments 1a et 1b sont respectivement le parler de Chiquihuitlán et de Mazatlán Villa de Flores. Leur relation au reste de l'aire mazatèque est complexe : Chiquihuitlán (Cq) est typiquement une aire latérale, rétentrice, et très différenciée, tandis que Mazatlán, bien au contraire, est une aire qui s'avère, d'après les résultats alternatifs que fournit l'analyse par algorithme de Levenshtein, très consensuelle avec l'ensemble du réseau dialectal (cf. carte 3 infra). Cette propriété du segment 1b est due, à notre avis, à un autre effet fort : celui des interactions avec les agrégats communaux (ou sociétés, communautés) des hautes terres, contrairement à la situation d'enclavement du segment 1a, dont le tropisme suivait plutôt les réseaux de la *Cañada* en interaction intense avec d'autres langues et agrégats communaux (aire cuicatèque, notamment).

Phase 2 (2a-d) : Il s'agit du principal *moment* ou clade de diversification du mazatec. Deux grands dialectes se constituent, à travers l'expansion sur deux versants : d'une part, les hautes terres centrales et leur périphérie méridionale (segments 2a-c), d'autre part les basses-terres, représentées ici par leur périphérie orientale (segment 2d). Aujourd'hui, ces deux secteurs sont connus de tous les locuteurs de mazatec comme correspondant aux dialectes de Huautla et de San Miguel Soyaltepec.

Phase 3 (3a-c) : un agrégat communal (autrement dit, des groupes d'agriculteurs méso-américains) se forme autour de la rivière Santo Domingo, au sud-est de l'aire mazatèque. Ce groupe descend du piémont oriental des hautes terres, que représente la zone d'Ayautla. Son centre est actuellement San Felipe Jalapa de Diaz, qui s'avère hautement consensuel avec le reste du réseau dialectal, du point de vue de l'algorithme de Levenshtein (cf. carte 4 infra). Tout comme pour le segment 1b, ce consensus s'explique à la fois par le caractère précoce de la séparation, et par l'effet fort des interactions ultérieures avec le reste du réseau dialectal, notamment avec le centre des hautes terres (Huautla). Le consensus est encore plus patent avec le petit bourg d'Ayautla, en plein piémont mazatec qui converge avec tous les autres parlars, sauf avec Huautla et le parler le plus réfractaire à l'ensemble de l'aire mazatèque : San Lorenzo. Cette différence de profil de consensus entre Jalapa et Ayautla vis-à-vis du parler de Huautla peut s'expliquer par l'effet des forces de démarcation (Séguy, 1973) ou de distinction de proximité entre variétés en situation de concurrence plutôt que de coopération, en ce qui concerne le binôme Ayautla-Huautla, tandis que la force des traits consensuels entre Jalapa et Huautla peut s'expliquer par la continuité entre deux centres massifs formés au centre de la plaine méridionale d'une part (Jalapa) et au centre des hautes terres d'autre part (Huautla) à époque précoce (force d'inertie des rétentions, dans un contexte où l'autonomie et le puissance régionale des deux centres est garantie).

Phase 4 : dans la plaine centrale et dans la partie occidentale du bassin aujourd'hui submergée par la réserve d'eau Miguel Alemán se constitue une aire dialectale innovante autour du bourg de San Pedro Ixcatlán. Les résultats de l'analyse par algorithme de Levenshtein suggèrent une forte concurrence entre les hégémonies locales des hautes terres centrales (Huautla, Jiotes) et ce centre. Là encore, la séparation a dû se faire à époque précoce, étant donné le haut degré de consensus avec le reste du réseau dialectal, hormis les centres réfractaires ponctuels que nous venons de citer (Hu, Ji) et le centre réfractaire pour tous les dialectes (Lo).

Phase 5 : le dernier moment cladistique serait représenté par le dialecte de San Lorenzo, autrement dit, le dialecte poblano, ainsi appelé car il déborde sur la Sierra negra, dans l'Etat voisin de Puebla. Nous avons consacré un article entier à la structure de ce dialecte ainsi qu'aux déterminismes qui ont pu présider à son émergence (Léonard, 2013). Nous avons également réalisé de nombreuses enquêtes dans le cadre du projet MAMP, surtout en 2011 et 2013, dans trois localités (Santa Ana, San Lorenzo, Huehuetlán), étonnés à chaque fois d'en mesurer les puissantes tendances à l'innovation, mais aussi à la diversification, par effet fort de déterminismes socioéconomiques récents – notamment l'expansion, puis l'effondrement du système café. L'intensité de ses changements vocaliques (vowel

automatisé s'est fait à l'aide du logiciel PAUP 4.0 et FACTOR (voir à ce sujet Gaillard-Corvaglia, 2012; Gaillard-Corvaglia & al. 2007, 2008, Brun-Trigaud & al. 2014). Les résultats de la figure 1.3 représentent la synthèse de six phylogrammes optimaux, à partir de données pondérées et non pondérées. Un article présentant le détail des données et de leur traitement est actuellement en préparation.

shifts, ou chaînes de propulsion et de traction vocalique) ainsi que la complexification des chaînes préfixales (préverbes de personne sujet et de temps, aspect, mode, voix) expliquent cette idiosyncrasie d'une petite aire périphérique (la population totale du mazateco poblano avoisine les 3500 locuteurs).

3.3. Falsification par la complexité algorithmique

Nous allons maintenant falsifier ou controuver le modèle présenté dans la figure 1.3, issue de l'analyse cladistique sur la base de trois ordres de faits (phonologie diachronique, morphologie verbale et construction préverbale des radicaux flexionnels) par les résultats qu'apporte le traitement des mêmes données de Kirk (1966) à l'aide de l'algorithme de Levenshtein, ainsi que par un autre ordre de données (et de complexité) : les tests d'intercompréhension réalisés par Paul L. Kirk dans 23 localités mazatèques, quelques années après son enquête dialectale (cf. Kirk, 1970).

Là encore, il faut distinguer trois types de complexité, selon O'Sullivan : la complexité *phylogénétique*, *ontogénétique* et *épigénétique*. La première nous est donnée par les phylogrammes de l'approche cladistique – et nous avons vu combien ceux-là sont eux-mêmes intriqués, car la figure 1.3 n'est pas un construit définitif, mais le produit de la synthèse d'au moins six arbres phylogénétiques optimaux). La seconde est donnée par un calcul trivial de distances, comme par exemple à l'aide de l'algorithme de Levenshtein, qui fournit pour chaque locolecte un gradient typologique. Nous avons vu combien les résultats obtenus par cet algorithme sont massifs, et semblent rendre principalement compte des différences phonologiques. Ils n'en sont pas moins précieux, car on sait combien l'intercompréhension et donc, la distance *épigénétique* (autrement dit, par interprétation des observables, en situation d'interaction subjective, ici intercommunale) en dépend.

Si les données de l'enquête sur l'intercompréhension par P. L. Kirk fournit la dimension épigénétique, nous disposons alors des trois volets, pour compléter et nuancer les objets de connaissance présentés dans les tableaux 1.3 et 1.4.

La carte 5 infra montre la relation ontogénétique, principalement du point de vue phonologique, entre San Pedro Ixcatlán (Ix) et le reste du réseau dialectal mazatec. On voit émerger une structure en relais avec un effet fort associant ce parler directeur, au centre de la plaine des basses terres, avec un parler périphérique de la plaine sud (Do, ou Santo Domingo, au bord du fleuve éponyme), qui laisse supposer que le segment 4 de la figure 1.3 est bien issue secondairement du segment 3. Autrement dit, Ix est bien un parler de la plaine, qui s'est individué, tout en connaissant une expansion à l'ouest des terres basses aujourd'hui immergées. Son aire de consensus est très vaste : elle comprend toute la périphérie du réseau dialectal. Ses centres de réfraction (les aires qui lui sont non pas consensuelles, mais en quelque sorte adverses, opposées) sont au centre des hautes terres : Huautla (Hu) et Santa Maria Jiotes (2c & 2b dans la figure 1.3), ce qui suggère, comme nous l'avons mentionné précédemment, une compétition entre les deux versants : centre des hautes terres versus plaine centrale. La carte 6 présente une structure frappante, qui milite en faveur d'une vision décroisée, ouverte, des configurations géolinguistiques du mazatec : l'aire de consensus, ou aire de résonance structurale de la variété la plus orientale – et donc périphérique – des basses terres, San Miguel Soyaltepec, est quasiment pan-mazatèque. Seul un locolecte s'avère réfractaire, et encore, à un degré moindre que dans ses relations avec la plupart des autres variétés : le parler poblano (Lo), au nord-ouest des hautes terres. Enfin, la carte 7 déploie la structure intriquée des relations de réfraction et consensus pour le parler central des hautes terres, celui de Huautla, rendu célèbre par les travaux de Kenneth Pike⁹. Il s'agit d'une variété dialectale puissante, expansive, comprise pratiquement partout dans les hautes terres – souvent estimée relever du « grand parler », notamment dans les pratiques cérémonielles et les commémorations et déclarations officielles, sur la place publique. Or, son profil de distance est complexe : il se laisse davantage interpréter en termes de dynamique de démarcation qu'en termes de consensus. Sa résonance

⁹ Cf. Pike, 1948 : Pike & Pike, 1947 et voir la modélisation alternative par Golston & Kehrein 1998 de la phonologie du mazatec, tout aussi décisive pour la phonologie générale, à partir du dialecte de Huautla. Les Mazatecs souvent ignorent que leur langue a eu un impact décisif sur les théories linguistiques au 20^e siècle – tout comme l'article de Gudschinsky, 1958b peut être considéré, comme il a été mentionné plus haut, comme une contribution empirique de premier plan aux études ethnohistoriques « ancrées » géographiquement par les données dialectales.

structurale avec Mz, Ji, Mg, Ja et So, en périphérie, peut aisément s'expliquer par la rétention dans les aires périphériques. En revanche, ses versants occidentaux (Te et, bien entendu, avec effet fort, Lo) s'en démarquent nettement, de même qu'une ceinture ou amphizone qui va de la plaine centrale au Canyon oriental (Cq) en passant par le piémont (Ay) et la plaine sud (Do). Cette différenciation pourrait être récente et s'être accentuée avec le passage au système agraire commercial – on peut même risquer l'hypothèse que ce sont davantage les versants mentionnés, à l'ouest et à l'est, qui ont préféré se différencier activement du parler central que l'inverse. Surtout lorsqu'on connaît la relation des populations des versants proches avec les gens de Huautla : il est fréquent d'entendre que l'on préfère se tenir à distance de « ceux de Huautla », au comportement jugé un peu trop hégémonique ou dominateur, sur le plan politique et social.

Enfin, dernier ordre de données, les résultats des tests d'intercompréhension dialectale (Kirk, 1970, Casad, 1974 : 47), qui fournissent des indices de lecture épigénétique de la différenciation dialectale du mazatec, dans son contexte écohumain. Conformément aux résultats de nos analyses phylogénétiques (cladistique) et ontogénétiques (dialectométrie à l'aide de l'algorithme de Levenshtein), le regroupement des basses terres orientales (So) avec les hautes terres y est confirmé : la périphérie de la montagne (par ex. Te) peut comprendre les basses terres lointaines (So). De même, un grand ensemble réunit, quoiqu'en préservant leur autonomie, aussi bien les grands centres des basses terres occidentales ou la plaine (Ix et Ja) et le centre des hautes terres (Hu). Mais au lieu que ce soit le versant le plus occidental et le plus périphérique qui soit isolé des autres, comme précédemment par distance de Levenshtein, c'est le parler de Chiquihuitlán, à la périphérie orientale du Canyon de Cuicatlán, qui se retrouve isolé. Le versant *poblano* (Lo), quant à lui, est bel et bien éclaté, dans son grand ensemble des hautes-terres occidentales, en de multiples sous-variétés, ce que nous avons pu également constater en explorant cette zone (cf. Léonard 2013). La pratique du terrain nous permet également de qualifier de quel ordre de complexité épigénétique il est question, dans ces données d'intercompréhension dialectale : Paul Livingston Kirk mentionne dans son article qu'il a eu recours à de nombreux informateurs mazatécophones travaillant dans les affaires municipales des 23 localités visitées. Par conséquent, le niveau d'agrégat communal dont il s'agit dans ces données est à la fois sous-représentatif de la population « moyenne », et représentatif de la partie la plus interactive de la population, à l'époque des enquêtes (années 1960), au niveau intercommunal. Comme toute problématique relevant de la complexité, on pressent que le segment de population représenté par ces données relève d'un secteur bien spécifique (une niche d'acteurs ou d'agents interactifs dans un système imbriqué en hologramme) : les résultats de l'enquête d'intercompréhension de Kirk, loin d'être peu représentatifs, donnent un reflet parmi tant d'autres possibles du champ d'interaction potentiel entre locuteurs de variétés différentes ou éloignées. La distance épigénétique se recoupe en grande partie avec nos distances phylogénétiques (cladistique) et ontogénétiques (Levenshtein), tout en apportant une tonalité différente, ou un registre d'interprétation contrasté. Une fois de plus, le tout (la *totalité*) s'avère bien plus que la somme de ses parties, mais aussi bien moins que la somme de ses parties.

4. Conclusion

L'approche en termes de *complexité* nous a permis de dépasser ou de transcender les construits essentialistes sur « l'ethnogenèse » entendue comme processus d'émergence d'une totalité culturelle. Elle a également permis de dépasser une mise en regard trop catégorielle ou hâtive entre milieu naturel (orographie, relief, climat, etc.) et constitution d'aires culturelles à travers des aires linguistiques, comme le sont les aires dialectales. Nos résultats montrent que ce qui émerge des modèles quantitatifs phylogénétiques et ontogénétiques que nous avons choisi d'appliquer ici, ce sont en effet des structures en partie corrélées, mais qui ne se laissent en rien réduire à des cloisonnements et des déterminismes simples. Par ailleurs, nous insistons sur l'importance de l'observation de terrain, voire de l'observation participante – nous avons eu accès à un grand nombre d'observables pragmatiques, organisationnels, et à des nombreux témoignages inespérés grâce à la pratique des ateliers et stages de formation d'instituteurs et de promoteurs culturels mazatecs, nahuas, cuicatecs, chinantecs et mixtecs du bassin du Papaloapam (cf. Léonard, Gragnic-McCabe & Avilés-Gonzalez, 2013), de 2010 à 2014, dans le cadre du projet MAMP de la chaire de géolinguistique et de morphologie méso-américaine de l'IUF. A l'occasion des 22 stages de deux à trois jours, organisés avec des centaines de locuteurs, nous avons pu également observer un effet de la relation systèmes/nature, en termes de systèmes et de régimes socioéconomiques de modernité (essor d'une économie de services, mobilité sociale et géographique,

bilinguisme ou plurilinguisme horizontal et vertical, élévation du niveau d'éducation, etc.), qui est le pendant de la relation culture/nature : la diversification constante du mazatec et des autres langues de cette région, qui s'avère un véritable bassin multilingue, à travers la multiplicité des normes – normes *prescriptives*, normes *descriptives*, *évaluatives*, *imaginaires* (Baggioni & Moreau, 1997 ; Simonin & Warthon, 2013). Il s'agit là d'un effet faible, mais certain, et ouvrant la boîte de Pandore d'un autre registre de complexité, des phénomènes liés à la modernité ou à la modernisation de la région, durant ces soixante dernières années. Plusieurs catastrophes – au sens de *bifurcations*, mais aussi au sens propre, du point de vue de certains effets socioéconomiques – modifiant la relation culture/nature ont frappé la région mazatèque : l'irruption des systèmes latifundiaires (ou, aujourd'hui, de type *maquiladoras*) liés aux régimes agraires commerciaux postcoloniaux (café, élevage et canne à sucre, systèmes étagés des hautes terres au basses terres), évacuation et déplacement forcé des populations du bassin de la vallée encaissée du Rio Tonto, dans les basses terres centrales, lors de la mise en œuvre du barrage Miguel Alemán (cf. Schwartz, 2014)¹⁰. D'une certaine façon, nombre d'attributs culturels qu'on aurait pu croire détruits à jamais ont survécu sous des formes nouvelles, voire réinventées – résilientes et réaménagées. Par exemple, nos enquêtes de l'été 2013 nous ont permis de documenter de nombreuses variétés dialectales issues des terres englouties sous les eaux du barrage en 1959, faisant apparaître en hologramme, à travers les idiolectes des descendants de première ou de deuxième génération des agriculteurs mazatecs autrefois installés dans cette vallée submergée, une dimension géolinguistique qu'on croirait perdue. Là encore, de multiples *normes résilientes* ou *innovantes* ont émergées, qui sont le produit d'interactions complexes, de fusions de répertoires multidialectaux et multilingues. La variation du mazatec n'est somme toute décrite que par la pointe de l'iceberg. Le monde d'observables à décrire et expliquer reste immense, à la mesure de la profondeur de ce microcosme, qui n'est qu'un complexe parmi des milliers d'autres que compte l'humanité aujourd'hui. Une humanité où la relation culture/nature est certes de plus en plus en crise, tant nous consommons les ressources de plusieurs planètes alors que nous n'en avons qu'une à disposition. Par ailleurs, la situation que nous venons de décrire, par sa complexité, laisse deviner qu'il existe un nombre indéterminé de mondes complexes en termes de relation nature/culture sur la planète à connaître et décrire – afin d'en tirer, si possible, des leçons pour mieux construire et moins détruire, en apprenant de cette grande vertu que manifestent les cultures humaines : la *résilience*.

Références

- Abse Edward, "Don Patricio's Dream: Shamanism and the Torments of Secrecy in Fieldwork among the Mazatec Indians", in Jean-Guy Goulet & Bruce Miller (eds.), *Extraordinary Anthropology: Transformations in the Field*, University of Nebraska Press, 2007, p. 258-282.
- Baggioni Daniel & Marie-Louise Moreau, 1997, "Norme. Genèse du concept", Moreau M-L. (ed.) *Sociolinguistique. Concepts de base*, Sprimont, Mardaga, p. 217-223.
- Beijering K. Gooskens C. & Heeringa W. "Predicting intelligibility and perceived linguistic distance by means of the Levenshtein algorithm", Amsterdam, *Linguistics in the Netherlands*, 2008, p. 13-24.
- Bolognesi, R. & W. Heeringa, "De invloed van dominante talen op het lexicon en de fonologie van Sardische dialecten", in D. Bakker, T. Sanders, R. Schoonen and Per van der Wijst (eds.). *Gamma/TTT, Tijdschrift voor taalwetenschap*, Nijmegen University Press, Nijmegen, 9 (1), 2002, p. 45-84.
- Blanton R., S. Kowalewski, G. Feinman, L. Finsten, *Ancient Mesoamerica. A Comparison of Change in Three Regions*, Cambridge, Cambridge University Press, [1981]-1993.
- Boege Eckart, 1988, *Los mazatecos ante la nación. Contradicciones de la identidad étnica en el México actual*, México, Siglo 2.
- Brown Cecil H., Holman Eric W., Wichmann Søren & Velupillai Viveka, "Automated classification of the world's languages: a description of the method and preliminary results". *STUF Language Typology and Universals* 61, 2008, p. 285-308.

¹⁰ Voir Abse, 2007, sur le désenchantement du monde du point de vue mazatec, dans les conditions de postmodernité postcoloniale, ainsi que Demanget, 2006.

- Brun-Trigaud Guylaine, Pierre Darlu, Antonella Gaillard-Corvaglia, Jean Léo Léonard & Patric Sauzet, « Exploration cladistique de l'ALLoc » au 10ème CIEO (Congrès de l'Associacion internacionala d'Estudis Occitans), 12-19 juin 2011, Béziers, in Alén Garabato Carmen, Claire Torrelles & Marie-Jeanne Verny (eds.), 2014 : *Actes du Xe Congrès de l'AIEO, Association Internationale d'Etudes Occitanes*, Limoges, Lambert Lucas, vol. 2, 2014, p. 489-500.
- Casad Eugene, *Dialect intelligibility testing*, University of Oklahoma, the Summer Institute of Linguistics, 1974.
- Demanget Magali, *La chair des Dieux est-elle à vendre ? Chamanisme, tourisme et ethnicité en terre mazatèque (Mexique)*, Thèse NR, Université Paris X, 2006.
- Darlu Pierre & Tassy Pascal, *La reconstruction phylogénétique. Concepts et méthodes*, Paris, Masson, (http://sfs.snv.jussieu.fr/pdf/Darlu_Tassy_online.pdf), 1993.
- Felsenstein Joseph, *PHYLIP Phylogeny Inference Package* (Version 3.2), *Cladistics* 5, 1989, p. 164-166.
- Felsenstein Joseph, *PHYLI. Phylogeny Inference Package* (version 3.6b). Distributed by the author. Department of Genome Sciences, Seattle, University of Washington, 2004.
- Gaillard-Corvaglia Antonella, *De la cladistique à la linguistique. Une étude appliquée aux dialectes italo-romans méridionaux et salentins*, Alessandria, Edizioni dell'Orso, 2012.
- Gaillard-Corvaglia, Antonella ; Pierre Darlu, Jean Léo Léonard, “Approche cladistique des vocalismes d'oïl à partir des données de l'A.L.F.”, *Bollettino dell'Atlante Linguistico Italiano III Serie*, 32, 2008, p. 36-78.
- Golston Chris & Kehrein Wolfgang, “Mazatec onsets and nuclei”, *International Journal of American Linguistics*, 64.4, 1998, p. 311-337.
- Gaillard-Corvaglia, Antonella, Léonard Jean Léo, Pierre Darlu, “Testing cladistics on dialect networks and phyla (gallo-romance and southern italo-romance)”, *Proceedings of Ninth Meeting of the ACL Special Interest Group in Computational Morphology and Phonology*, pages 23–30, Prague, Juin 2007. Association for Computational Linguistics: <http://www.aclweb.org/anthology-new/W/W07/W07-1304.pdf>
- Goebel Hans, *Dialektometrie. Prinzipien und Methoden des Einsatzes der numerischen Taxonomie im Bereich der Dialektgeographie*, Vienne, Verlag der Öst. Akademie der Wissenschaften, 1982.
- Goebel Hans, “Eléments d'analyse dialectométrique (avec application à l' AIS)”, *Revue de Linguistique romane*, 45, 1981, p. 349-420.
- Goebel Hans, “Analyse dialectométrique des structures de profondeur de l'ALF”, *Revue de Linguistique Romane*, 66, 2002, p. 5-63.
- Golston, Chris & Kehrein, Wolfgang, “Mazatec onsets and nuclei”, *International Journal of American Linguistics*, 64.4, 1998, p. 311-337.
- Gudschinsky, Sarah, *Proto-Popotecan. A Comparative Study of Popolocan and Mixtecan*, *IJAL*, 25(2), 1958a.
- Gudschinsky Sarah, “Mazatec dialect history”, *Language* 34, 1958b, p. 469-481.
- Gribbin, John, *Deep Simplicity, Chaos, Complexity and the Emergence of Life*, London, Penguin, 2004.
- Hennig Willi, *Grundzüge einer theorie der Phylogenetischen systematik*, Berlin, Deutscher Zentralverlag, 1950.
- Hennig Willi, *Phylogenetic Systematics*. Urbana, University of Illinois Press, 1966.
- Kirk Paul L., *Proto-Mazatec phonology*. PhD dissertation, University of Washington, 1966.
- Kirk Paul L., “Dialect Intelligibility Testing: The Mazatec Study”, *International Journal of American Linguistics*, 36-3, 1970, p. 205-211.
- Labov William, “Transmission and Diffusion”, *Language*, 83, 2, 2007, p. 344-387
- Léonard Jean Léo, “Enquêtes exploratoires pour l'ALMaz (*Atlas Lingüístico Mazateco*). Elicitation croisée, entre typologie et codification d'une langue otomangue”, Grenoble, *Géolinguistique* 11, 2010, 59-109.
- Léonard Jean Léo, “An xo'boo (mazatec de Puebla) : un fil d'Ariane écolinguistique dans le labyrinthe dialectal mazatec”, à paraître in Djordjevic Léonard, *Les minorités invisibles : diversité et complexité (ethno)sociolinguistiques*, Paris, Michel Houdiard Editeur, 2013, p. 290-304.
- Léonard Jean Léo ; dell'Aquila, Vittorio & Gaillard-Corvaglia, Antonella “The ALMaz (*Atlas Lingüístico Mazateco*): from geolinguistic data processing to typological traits”, *STUF*, Akademie Verlag, 6, 2012, p. 78 -94.

Léonard Jean Léo & Fulcrand, Julien, “Disentangling the inflectional role of tone in Mazatec: Tone Root Class, OCP and 1 sg/Incomplete Marking”, à paraître in Palancar & Léonard, *Disentangling the inflectional role of tone*, Berlin/NY, Mouton de Gruyter, 2013.

Léonard Jean Léo & Alain Kihm, “Classes flexionnelles du mazatec et diasystème. Empirisme critique et formalisation”, *Bulletin de la Société de Linguistique de Paris* 56, 2012, p. 379-446.

Léonard Jean Léo, Avilés González Karla & Gragnic, Julie, 2013. “Multilingual Policies put into Practice: *Co-participative Educational Workshops in Mexico*”, in Hogan-Brun, Gabrielle (éd.), *Applied Linguistics, Sociolinguistics/Current Issues in Language Planning: Language planning as a complex practice*, 14, 3-04, 2013, p. 419-435.

Massip-Bonet Àngels & Albert Bastardas-Boada (eds.), *Complexity Perspectives on Language, Communication and Society*, Berlin & Heidelberg, Springer Verlag, 2013.

Swofford David, *PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods)*, 2002.

Murra John Victor, *The Economic Organization of the Inca State*, Ph.D. diss., Department of Anthropology, University of Chicago, 1956.

Murra John Victor, "El Archipiélago Vertical" Revisited, in *Andean Ecology and Civilization*, edited by Shozo Masuda, Izumi Shimada & Craig Morris, 3-13, Tokyo, University of Tokyo Press, 1985.

Nichols Johanna, “The Origin of the Chechen and Ingush: A Study in Alpine Linguistic and Ethnic Geography”, *Anthropological Linguistics*, 46-2, 2004, p. 129-155.

O’Sullivan David, « Complexity Science and Human Geography », *Transactions of the Institute of British Geography*, 2004, p. 282-295.

Pike Kenneth, *Tone Languages. A Technique for Determining the Number and Types of Pitch Contrasts in a Language, with Studies in Tonemic Substitution and Fusion*, Ann Arbor, University of Michigan Press, 1948.

Pike Kenneth & Pike Eunice, “Immediate constituents of Mazatec Syllables”, *IJAL*, n°13, 1947: 78-91.

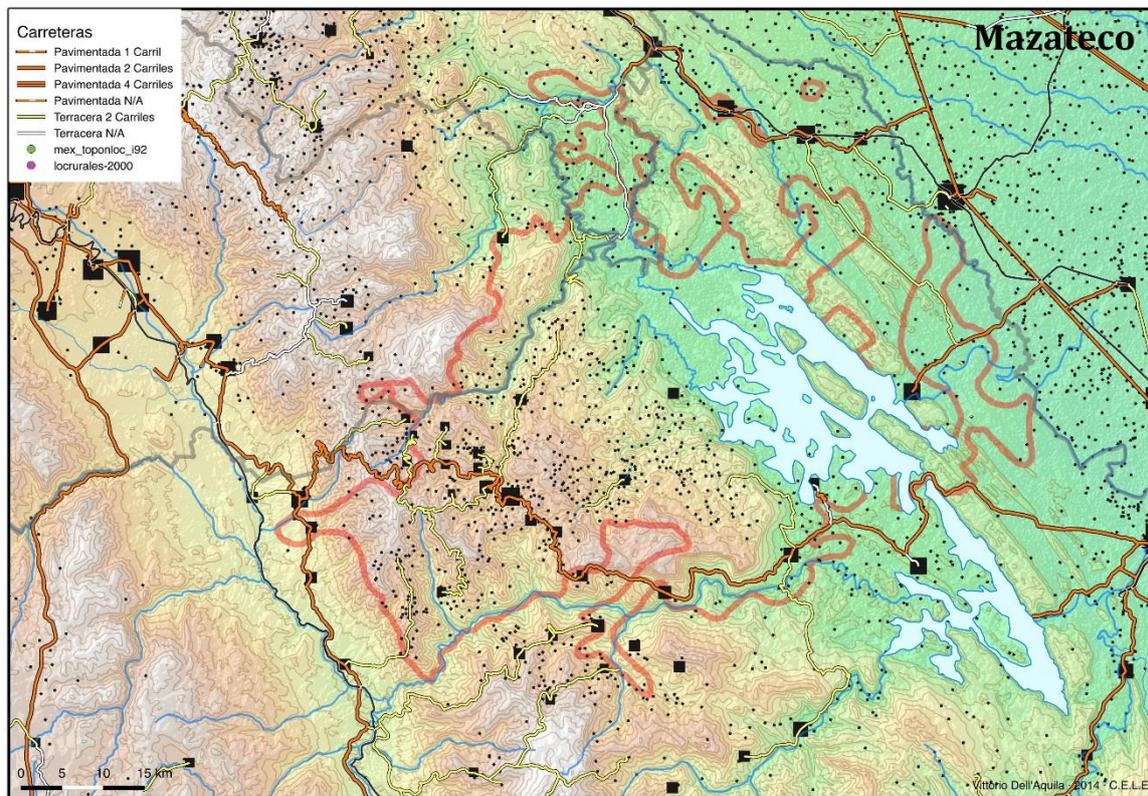
Rensch, Calvin, *Comparative Otomanguean Phonology*, Ph.D. Dissertation, University of Pennsylvania, Indiana University Publications, Language Science Monographs 14, Bloomington: Indiana University Press, [1966]-1976.

Schwartz Diana L., “Resettlement as Planned Utopia: Mexican Anthropologists and Rural Development in the Papaloapan, Mexico”, American Historical Association, symposium *Envisioning Capitalist Development in the Countryside: Perspectives from Latin America, Asia, and the United States*, 2014 (manuscrit, à paraître).

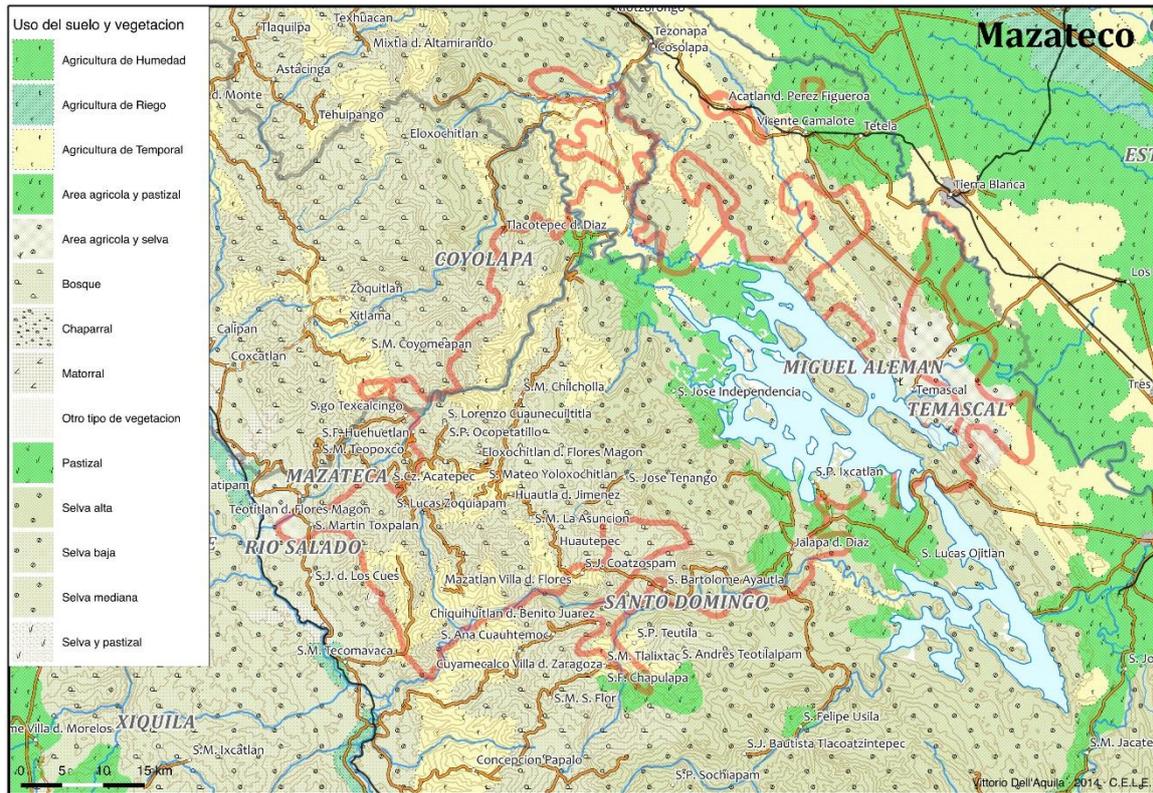
Séguy Jean, “La fonction minimale du dialecte”, in *Les dialectes français à la lumière des atlas linguistiques régionaux*, éditions du CNRS, Paris, 1971.

Simonin Jacky & Wharton Sylvie (eds.), *Sociolinguistique du contact. Dictionnaire des termes et des concepts*, Lyon, ENS éditions, 2013.

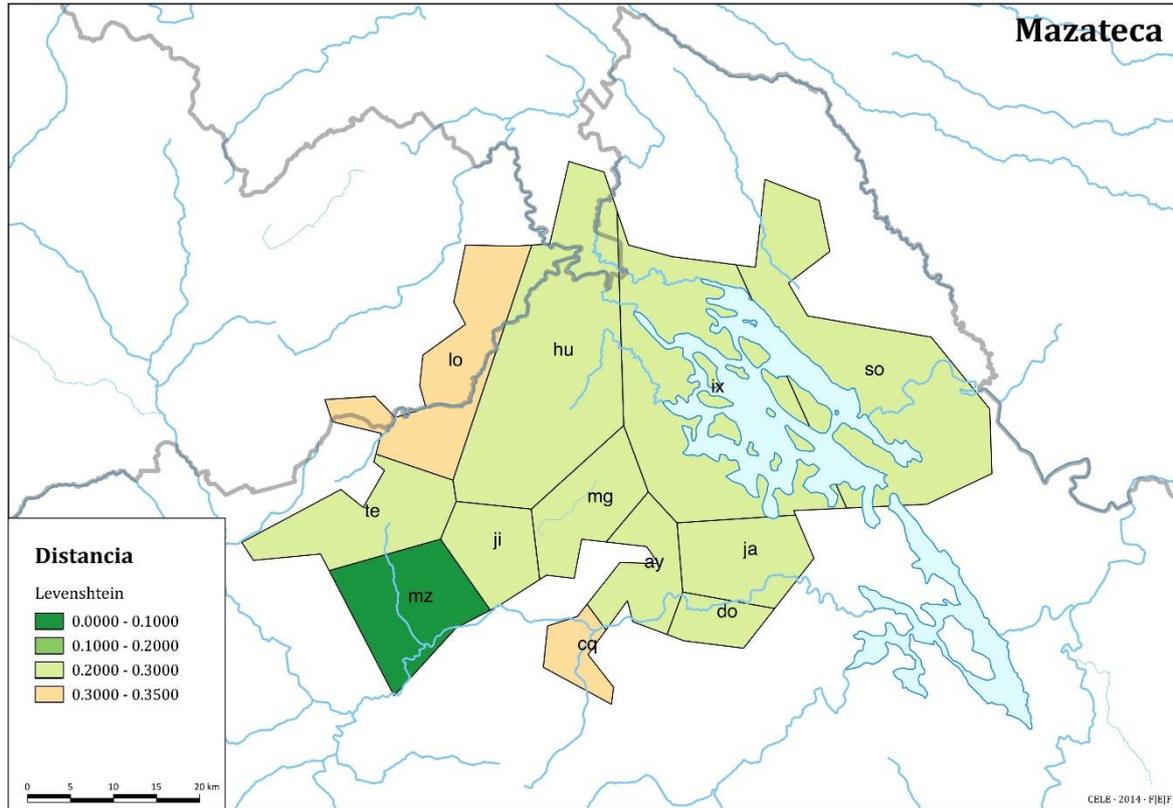
Wichmann Søren, Müller André, Velupillai Viveka, Brown Cecil H., Holman Eric W., Brown Pamela Urban Matthias, Sauppe Sebastian, Belyaev Oleg, Molochieva Zarina, Wett Annkathrin, Bakker Dik, List Johann-Mattis, Egorov Dmitry, Mailhammer Robert & Geyer Helen. *The ASJP Database* (version 12), 2010.



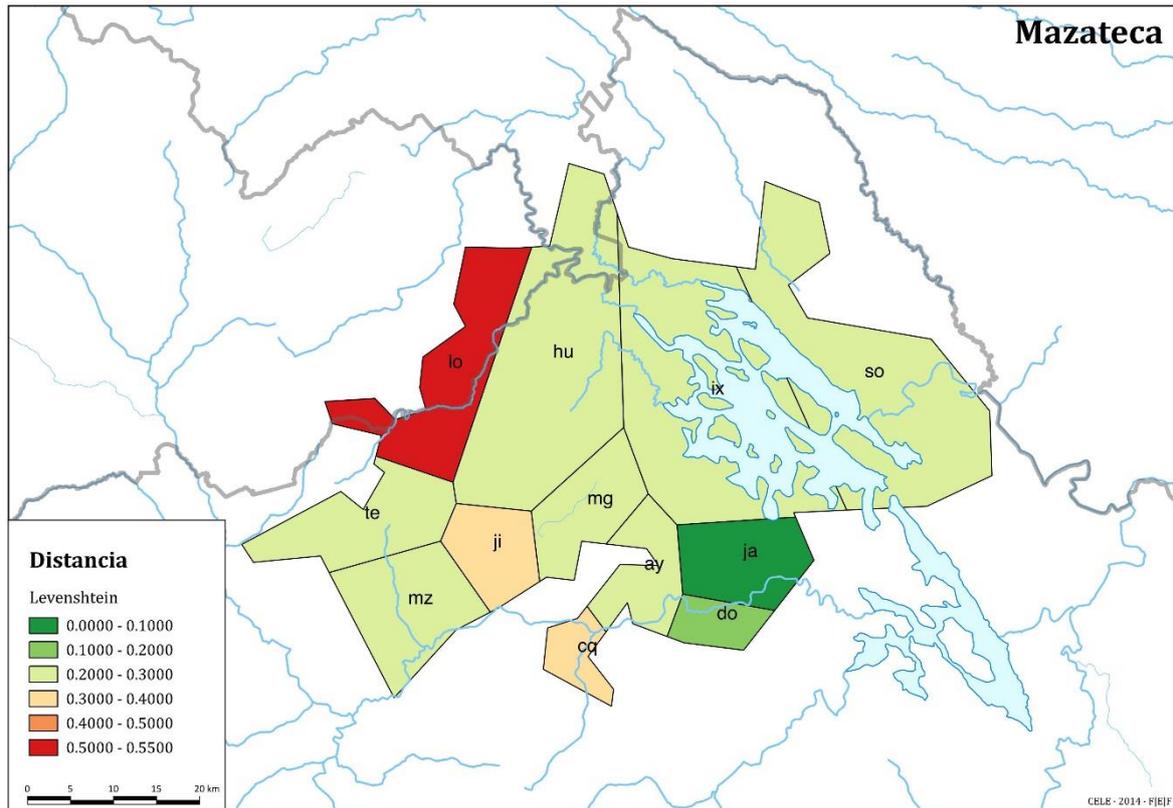
Carte 1. Peuplement, centres urbains et voies de communication dans la région mazatèque
 NB : référence pour éditeur, titre de fichier joint : Population_4



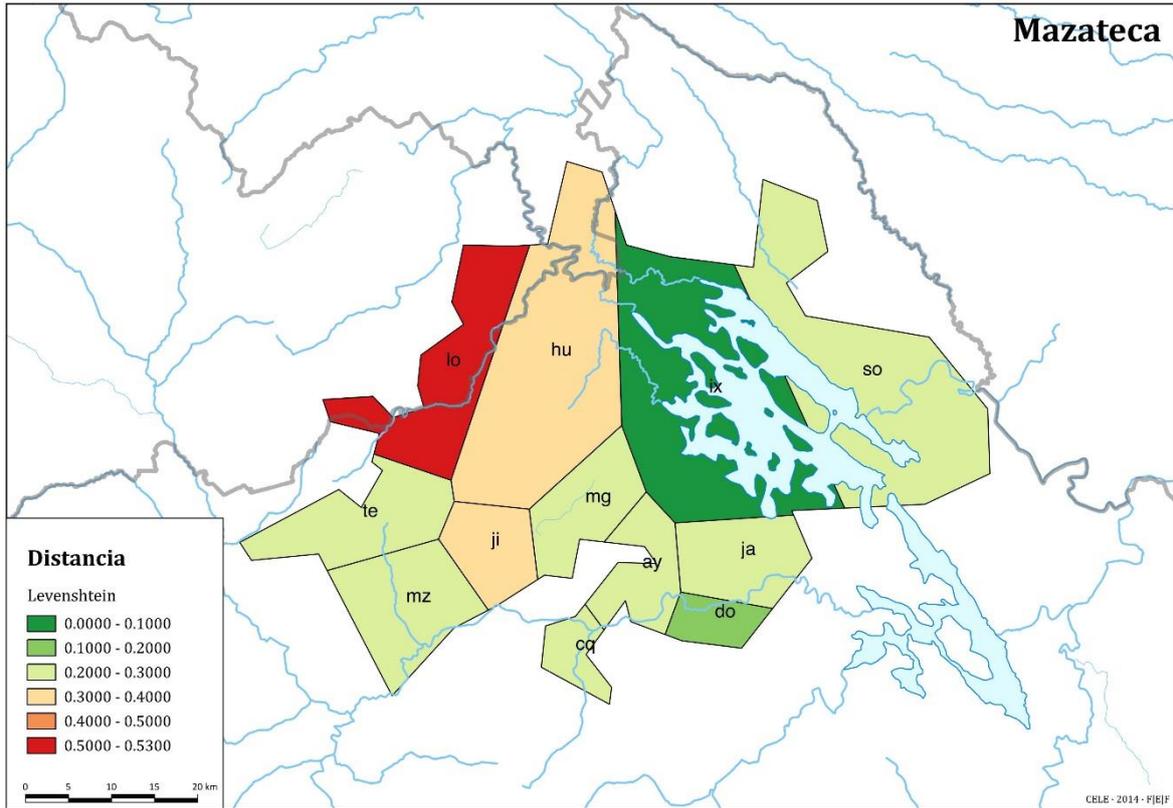
Carte 2. Types de sols dans la région mazatèque
 NB : référence pour éditeur, titre de fichier joint => Uso_del_suelo_2



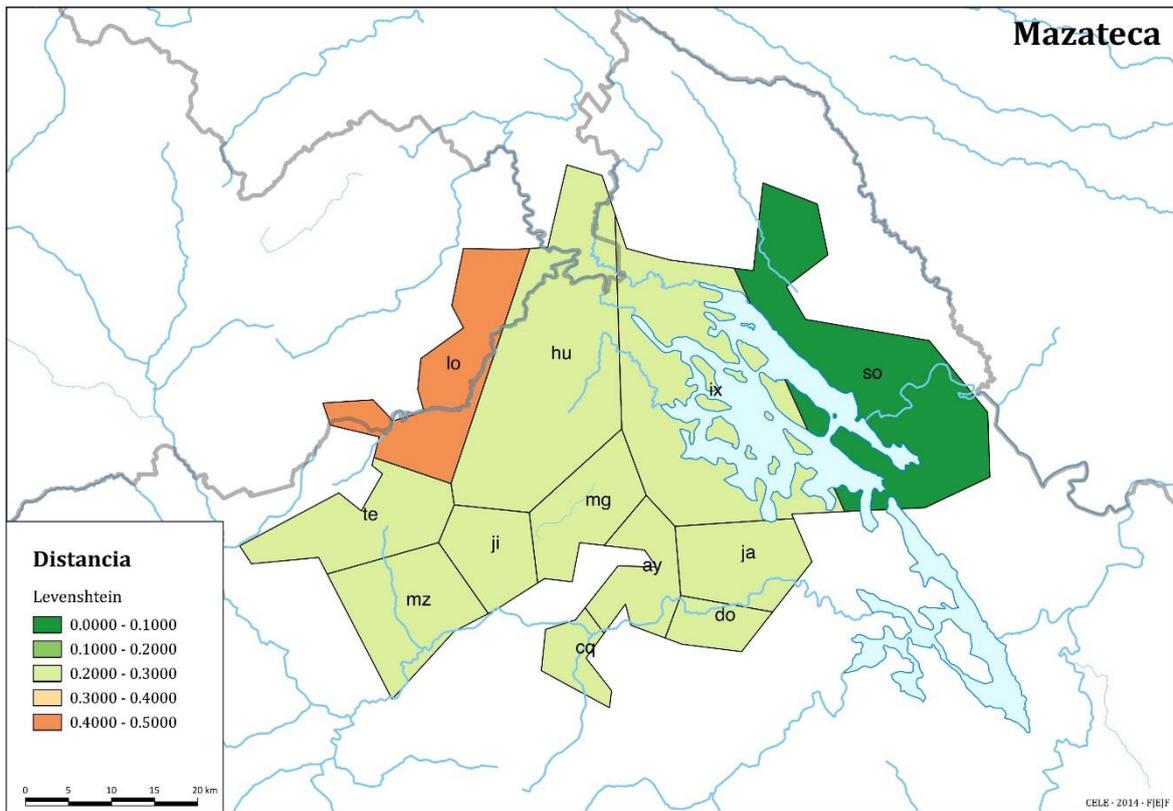
Carte 3. Application de l'algorithme de Levenshtein au locolecte Mz
NB : référence pour éditeur, titre de fichier joint => LD_MZ



Carte 4. Application de l'algorithme de Levenshtein au locolecte Ja
 NB : référence pour éditeur, titre de fichier joint => LD_JA

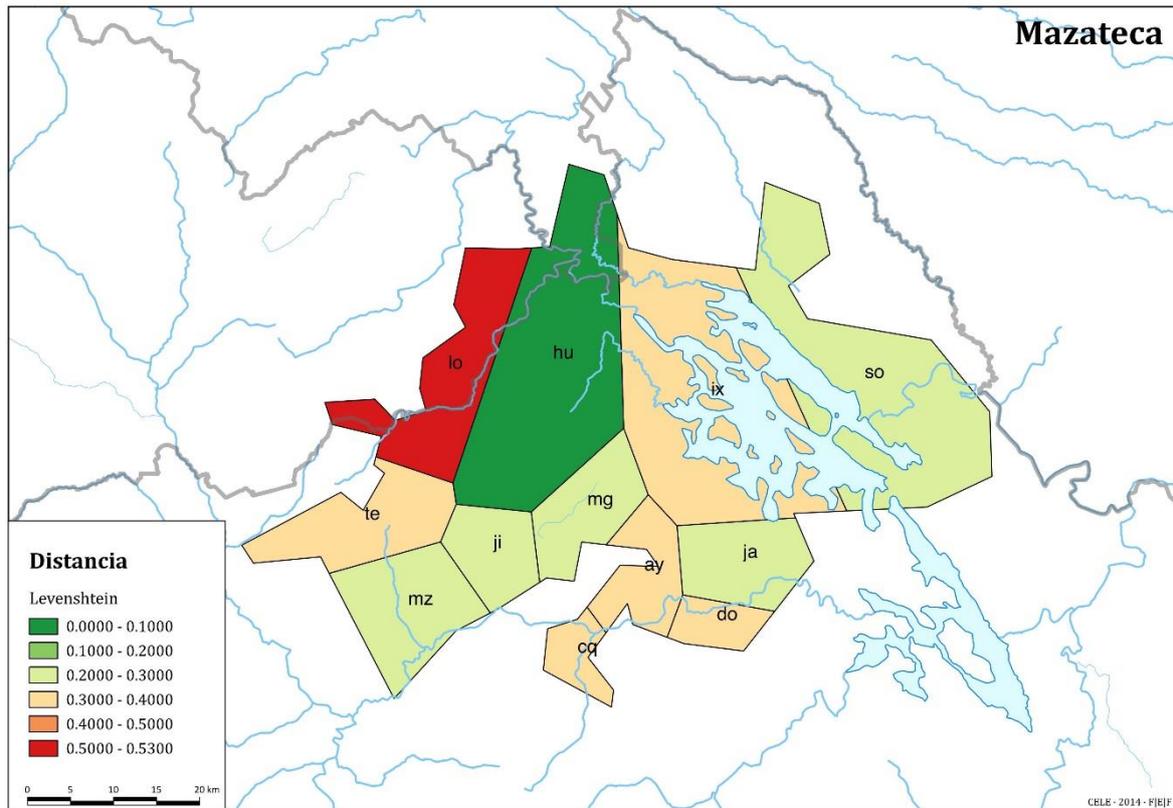


Carte 5. Application de l'algorithme de Levenshtein au locolecte Ix
NB : référence pour éditeur, titre de fichier joint => LD_IX



Carte 6. Application de l'algorithme de Levenshtein au locolecte So

NB : référence pour éditeur, titre de fichier joint => LD_SO



Carte 7. Application de l'algorithme de Levenshtein au locolecte Hu
NB : référence pour éditeur, titre de fichier joint => LD_HU