

Léonard, Jean Léo (IUF & UMR 7018 CNRS-Paris 3)

**Enquêtes exploratoires pour l'ALMaz (*Atlas Lingüístico Mazateco*) :
élicitation croisée, entre typologie et codification d'une langue otomangue**

1. Introduction

1.1. Géolinguistique otomangue : un défi, un horizon exploratoire, une ressource

Les langues otomangues constituent un phylum linguistique parmi les plus diversifiés, et probablement parmi les plus anciens d'Amérique (Kaufman 2006, Longacre 1967). Bien qu'une reconstruction solide et fiable du phylum soit possible (Rensch 1976) et de ses différentes composantes (Longacre 1957, 1966 ; Mak & Longacre 1960 ; Suarez 1973 ; Swadesh 1947 ; Fernandez de Miranda 1995-[1966] ; Gudschinsky 1959 ; Kirk 1966, etc.), la distance entre les sous-familles de ce groupe génétique¹ n'en reste pas moins déconcertante², comparée aux différences structurales observables dans d'autres phylae mésoaméricains (mixe-zoque ou maya, bien plus denses sur le plan structural). *A priori*, les très grands écarts structuraux, notamment en termes d'inventaires lexicaux et de choix des racines et des éléments spécifieurs qui donnent aux lexèmes leurs formes variables, entre langues otomangues, rend très difficile la conception et la réalisation d'atlas linguistiques.

Le présent article aura pour objectif de présenter les premières avancées d'une enquête pilote, en suivant trois lignes d'action : de compilation, de collecte et de conceptualisation. Dans le cadre d'un projet quinquennal IUF de géolinguistique mésoaméricaine (2009-2013), le MaMP, les données de langues otomangues orientales (essentiellement tu'un savi ou mixtec, zapotec et mazatec) de provenance diverse (Bradley & Josserand 1982, Rendon 1995³) ont été rassemblées dans des bases de données cartographiques, ouvrant un *horizon* de traitement géolinguistique de langues d'un haut de gré de diversification interne. Des prototypes ont été créés, à l'aide de la cartographie automatisée développée au CELE⁴, dont on voit ci-dessous l'un des produits. Cette approche permet de mettre en valeur les convergences, les grandes dynamiques structurales qui fédèrent le macrosystème otomangue, du point de vue de ses caractéristiques typologiques héritées, mais ne dispense pas pour autant d'enquêtes complémentaires ou de vérification sur le terrain, bien au contraire. La phase de compilation atlantographique doit mener dans une deuxième temps à une phase d'enquêtes de terrain, de révision et d'extension des données mises à la disposition des chercheurs et des usagers de la langue. Ces atlas compilatoires, qui peuvent aussi bien, en termes d'édition des cartes, ne pas dépasser le stade de la base de donnée cartographique sous format électronique,

¹ Dans le sens de « linguistique génétique » – entendue comme *affiliation structurale* et *classement* des langues ou des variétés dialectales issues d'un système linguistique « hérité », dont la continuité est inférable par le nombre et la cohérence structurale des correspondances (racines lexicales, clitiques et affixes) et des processus évolutifs (lois phonétiques), observables. La « linguistique génétique » n'a rien à voir avec la génétique en biologie, si ce n'est pas les congruences que l'on peut rechercher entre la classification des langues et la classification du vivant, qui n'est pas ici de notre ressort. Espérons que la linguistique parviendra un jour à s'émanciper de ces métaphores de « phylum », « famille de langues », « parenté linguistique », « langues-mères » et « langues-filles » héritées des sciences naturelles et du positivisme du XIXe siècle. Elles gardent cependant, dans l'état actuel des choses, une certaine valeur heuristique afin d'opposer les *réseaux génétiques* aux *réseaux typologiques* et aux *réseaux aréaux* en linguistique.

² Cf. Kaufman 2006 pour un survol critique de la structure interne du *Stammbaum* otomangue. Voir Cazes (s.d.) pour un examen glottochronologique de la branche occidentale du *Stammbaum* (ou « *arbre généalogique* de classification des langues » – cf. note 1 supra), à comparer avec Soustelle 1937.

³ Les listes de correspondances de Josserand 1983 pour le mixtec et de Kirk 1966 pour le mazatec sont actuellement en cours de saisie pour compléter cette base de données cartographique.

⁴ V. <http://www.celeuropa.eu/home.page> et <http://www.celeuropa.eu/bibliografia.page>, pour les publications.

servent par ailleurs de maquettes ou de bancs d'essai à des atlas plus classiques dans la méthodologie : collecte, édition cartographique et analyse de données de première main, avec une campagne d'enquêtes systématiques sur le terrain.



Fond de carte pour un atlas linguistique de l'Etat de Oaxaca (Mx) et aire du mazatec (otomangue oriental), par Vittorio dell'Aquila (CELE)

Ces atlas, qu'ils soient compilatoires ou canoniques dans leur élaboration, du terrain aux cartes version électronique ou papier, doivent par la suite fournir aussi bien des données utiles aux instituteurs bilingues et aux promoteurs de langues indigènes qu'au linguiste typologue, qui ne dispose guère pour le moment que des compilations du WALs (Haspelmath & al. 2005-08) pour étudier la diversité des langues dans cette région du monde, pourtant l'une des plus riches en langues sur la planète.

1.2. Un projet pilote : enquêtes exploratoires pour l'ALMaz (*Atlas Lingüístico Mazateco*)

1.2.1. Contraintes techniques et cadre général

Le présent article esquissera dans le cadre d'un exposé nécessairement limité en extension, les défis théoriques et empiriques d'une telle approche dans un contexte comme celui des sociétés « indigènes » ou « autochtones » du Mexique. L'ensemble de ces remarques, qui décriront un fragment d'enquête pilote, tiré d'une intense activité de prospection dialectologique menée entre le 22 juillet et le 21 août 2010 dans la région mazatèque, permettra de contextualiser les conditions de réalisation d'un atlas linguistique dans une région linguistique de la Mésoamérique moderne. Ce cheminement ne sera pas seulement l'occasion de répéter l'antienne bien connue de tous comme quoi les données linguistiques ne se recueillent pas aussi facilement que les papillons dans les prés, à l'aide d'un simple filet (les appareils d'enregistrement audiovisuels, le stylo et le papier), en courant avec des jambes solides et des bottes bien arrimées aux chevilles. Les données linguistiques, pour être fiables ou valides, exigent *en amont* une théorisation, une construction, une réflexion sur la nature des catégories lexicales et fonctionnelles ; *en aval*, le linguiste de

terrain ne doit pas seulement se confronter avec les aspérités du relief, les écarts climatiques et l'hermétisme des réseaux sociaux et des conventions ou des contraintes psychosociales – dont la diglossie n'est pas la moindre – : il doit s'apprêter à accompagner ses collaborateurs *ad hoc* que sont ses « informateurs » ou « maîtres de langue » dans une introspection de leur compétence linguistique, un travail ardu sur leur langue maternelle, qui n'est en rien immédiat ni facile, d'autant plus lorsque l'écart typologique avec la langue véhiculaire ou la *lingua franca* de l'enquête est considérable, comme c'est le cas entre toute langue amérindienne et l'espagnol ou toute autre langue coloniale issue de l'Europe romane ou anglo-saxonne. Bien entendu, la meilleure solution reste l'enquête entièrement conduite dans la langue, ou n'utilisant qu'accessoirement la langue véhiculaire nationale, comme c'est le cas dans les enquêtes du projet ALTO, décrites dans le précédent numéro de *Géolinguistique* (Polian & Léonard 2009). C'est en mazatec, dans les variétés haute et basse selon les localités, que devra se dérouler la seconde étape des enquêtes pour l'ALMaz (*Atlas Lingüístico Mazateco*), avec des linguistes natifs. Mais il serait illusoire de penser que la seule mazatéquisition de la procédure d'enquête résoudrait les problèmes que nous allons examiner ici, si les enquêteurs ne sont pas au préalable formés sur le plan théorique et avertis des écueils que l'enquête réalisée à l'aide de l'espagnol comme *lingua franca* révèle avec d'autant plus d'acuité. Sur le plan phonologique, l'enquêteur doit connaître avec une lucidité théorique impeccable les principes de structuration du système tonal au-delà de ses propres impressions phonétiques, il doit comprendre les contraintes de distribution en surface des traits de glottalité *breathy* et *creaky* (voix soufflée et voix craquée), il doit être averti des degrés de relâchement vocalique en surface. La tâche à laquelle il se trouve confronté sur le plan morphologique, dans le domaine de la flexion verbale, est encore plus ardue : il lui faut se détacher résolument du système temporel de l'espagnol et de l'organisation des marques d'Aspect-Temps-Mode dans les thèmes flexionnels pour accepter de voir se reconfigurer aspect et personne d'une manière totalement étrangère à l'espagnol, sa langue d'alphabétisation et de scolarisation. Cette opération est d'autant plus difficile que la grammaire du mazatec – mais aussi de n'importe quelle autre langue otomangue dans l'Etat de Oaxaca, et ailleurs au Mexique – ne fait jamais l'objet d'une analyse en paradigmes ordonnés selon des critères structurels endogènes, dans le système de formation des maîtres d'écoles. Lorsqu'on sollicite un tel raisonnement, c'est toujours à partir de l'espagnol que cette tâche est accomplie, menant indéfectiblement à une aporie – disons-le tout net sans plus d'ambages, à un échec –, qui fait que le jeu de l'élicitation est terminé avant même d'avoir réellement commencé. Malgré ces limites, dues aux carences de l'éducation formelle en langues autochtones et sur les langues autochtones du Mexique, il n'en reste pas moins que nombre de locuteurs du mazatec et des langues originaires du Mexique ont accompli des prodiges, par la force de la volonté et l'acuité de leur intelligence : tous les maîtres d'écoles bilingues avec lesquels j'ai eu le plaisir et l'honneur de travailler durant cette enquête pilote savaient écrire leur langue maternelle d'une manière systématique et informée. Ils étaient le plus souvent parvenus à cette maîtrise de l'écrit en langue maternelle de leur propre chef, en autodidactes, en dépit des carences du système de formation des maîtres, qui ne peut faire face, avec le peu de moyens dont disposent les Ecoles Normales et les Universités Pédagogiques de l'Etat de Oaxaca et d'ailleurs, pour assurer un enseignement satisfaisant de l'écriture et de la grammaire des seize langues officiellement recensées, mais que la diversification dialectale multiplie, pour certaines, par plus de 20, comme c'est le cas pour le zapotec et le mixtec notamment. J'ai même travaillé à la fin de mon séjour dans les Hautes Terres (Mazateca Alta) avec un locuteur de San Jeronimo Tecoatl, Leonardo García Baltazar, qui avait entièrement traduit les évangiles sous la direction d'un conseiller du *Summer Institute of Linguistics*. L'élicitation de paradigmes verbaux nécessita cependant beaucoup de patience et de tact, et il s'avéra impossible d'effectuer un travail d'analyse syntaxique d'énoncés que je lui avais demandé de traduire à partir d'une liste de

phrases en espagnol conçues initialement en mazatec des Basses Terres (Mazateca Baja) par un autre informateur ou maître de langue (l'écrivain Apolonio Bartolo Ronquillo, 49 ans, originaire de San Miguel Soyaltepec), mais il n'en reste pas moins que le travail et les connaissances de cet homme sur sa langue maternelle imposent le respect. Leonardo García Baltazar est par ailleurs auteur de divers opuscules en mazatec publiés à compte d'auteur, d'une grande utilité pour la diffusion écrite de sa langue à échelle locale. Mais ces compétences, bien que hautement estimables, restent bien en-deçà des aptitudes requises pour envisager un travail systématique de collecte de données.

En admettant qu'un projet tel que l'ALMaz dispose d'un financement permettant la réalisation des enquêtes – et il se trouve qu'un tel financement existe potentiellement, à travers une dotation financière de l'IUF (*Institut Universitaire de France*), qui a permis la réalisation de cette enquête pilote et a également soutenu le projet ALTO durant le printemps et l'été 2010 –, il serait, dans l'état actuel des choses, prématuré d'envoyer une équipe d'enquêteurs sur le terrain, tant les prérequis sont élevés, en termes de compétence technique en linguistique théorique et descriptive. Autrement dit, le coût théorique d'un atlas linguistique est bien plus élevé qu'on ne songe ordinairement, notamment dans le contexte européen, où il est moins difficile de trouver des compétences en termes de ressources humaines, que d'obtenir des crédits de recherches suffisants pour la réalisation d'enquêtes, l'édition et l'impression d'un atlas linguistique, selon les standards européens (coûts salariaux, frais de fonctionnement et d'équipement, frais de transports et d'hébergement). On pourra toujours se donner bonne conscience d'une manière ou d'une autre, mais il n'en reste pas moins que, alors que des sommes colossales sont engagées dans des projets pharaoniques, aussi bien au Mexique qu'en Europe, des populations entières, comme les Mazatecs et la plupart des peuples originaires (« amérindiens ») du Mexique, ne disposent ni d'infrastructures pour l'enseignement et l'étude de leurs langues, malgré des dispositions juridiques nationales (*Ley General de Derechos Lingüísticos de los Pueblos Indígenas 2003*) et internationales (dont la *convention C 169 de 1989 de l'Organisation Internationale du Travail*), ni d'outils pour la connaissance et la mise en valeur des structures de leurs langues. Tant que de tels instruments n'existeront pas, ou ne seront pas diffusés – et nous verrons bientôt les enjeux de la divulgation des données et des connaissances, lorsque je relaterai l'atelier de linguistique diachronique mazatèque à huautla –, tant que chaque communauté ne disposera pas d'un groupe d'experts techniques (autrement dit, de *linguistes*) formés et disposés à former de nouvelles générations de locuteurs natifs impliqués dans l'enseignement, la promotion et l'aménagement de leur langue, ces dispositions favorables aux langues resteront lettre morte, et ne serviront qu'à donner bonne conscience aux décideurs politiques et aux observateurs. Le terme qualifiant les linguistes mazatèques que j'entendais souvent mentionner à Huautla était « ethno-linguiste » (*etnolingüista*). Il me semble réducteur de donner une spécification *ethnique* au linguiste, à travers ce terme : sans une vision généraliste, *extra-ethnique*, et une solide formation en linguistique générale et descriptive, la ghettoïsation de la langue, son enfermement dans une logique exogène unique qui s'avère être celle de l'espagnol en tant que vulgate nationale de la grammaire grécolatine, continueront de bloquer l'aménagement linguistique des langues indigènes du Mexique.

Parallèlement, de manière aussi inexorable qu'insidieuse, l'assimilation suit son cours : le mazatec, langue réputée hors de danger avec sa population dépassant les 200 000 locuteurs, densément rassemblés dans leur région historique au sein de la Cuenca du Papalopapan, se trouve actuellement sur la pente glissante de l'assimilation et de la substitution (ou *Language Shift*, ou *desplazamiento lingüístico*) : en milieu urbain, les enfants de moins de dix ans sont de plus en plus nombreux à ne plus parler que l'espagnol, ou à ne maîtriser que de manière très imparfaite ou approximative le mazatec, comme langue seconde. Leurs familles n'en sont pas mieux loties sur le plan économique, et rien ne prouve

que les perspectives d'avenir de ces enfants en sont pour autant meilleures que s'ils avaient atteint un degré satisfaisant de bilinguisme équilibré, davantage propice à une insertion future aussi bien locale que hors de Jalapa de Díaz ou de Huautla de Jiménez. Dans un grand centre multiethnique et à dominante métisse comme Tuxtepec, la transmission du mazatec est encore bien plus affaiblie que dans les centres urbains de la Mazateca. En milieu rural, le monolinguisme continue d'accompagner la misère des petits paysans vivant en autosuffisance, ou dépendant de subsides gouvernementaux et d'apports financiers de migrants, dans des zones sous-équipées en services éducatifs et en services de santé, mal reliées aux rares voies de communications qui sillonnent sur de grands axes tortueux la Cuenca del Papaloapán.

La Cuenca del Papaloapán : un vaste bassin entièrement fait de montagnes, de fleuves, de ravins, de cañons tropicaux, qui a davantage connu dans son histoire moderne l'effet dévastateur de l'aménagement positiviste et technocratique du territoire qu'une politique attentive à sa diversité écologique et culturelle, comme nous le verrons plus loin, lorsque nous aborderons l'espace mazatec par les Basses Terres, où le lac artificiel Miguel Alemán, issu d'un projet pharaonique de barrage hydroélectrique réalisé entre 1947 et 1954 forme une gigantesque tache, un monde englouti, au-milieu d'un des plus anciens versants du monde mazatec – les Basses Terres et la vallée de Soyaltepec. A perte de vue s'alignent les anciens sommets des collines immergées, ultimes vestiges d'un espace de 51 000 hectares de terres fertiles, dont 21 000 ha éjidales (ou terres communales) et 18 000 ha de petites propriétés familiales de moins de 4 ha⁵. Bien que mal distribuée et en mal d'une réforme agraire, une terre particulièrement fertile, aujourd'hui engloutie, alimentait une population paysanne implantée de longue date, relativement moins pauvre et moins asservie que dans d'autres zones de la Cuenca et de l'Etat de Oaxaca. Ses réseaux sociaux et son tissu socioculturel furent entièrement démembrés, disloqués, déplacés et les paysans mazatèques de la région de San Miguel Soyaltepec furent condamnés à l'exode rural ou à la relocalisation dans des villages artificiels lors de la mise en place du complexe hydroélectrique. La géolinguistique devra tenir compte de cet événement majeur, en tant que facteur déstructurant et restructurant de la diversité dialectale de la langue. Ironie du sort, dans les nouveaux villages, tels que Pescadito de Abajo Segundo, où j'ai mené des enquêtes dialectales durant une semaine, j'ai pu constater que les habitants, qui résident au bord du lac artificiel, ne bénéficient ni de l'eau courante, ni de services d'eau potable. La société locale, suite à cette reconfiguration agraire imposée il y a plus d'un demi-siècle à une population de paysans pauvres, petits propriétaires ou exploitants de terres éjidales (communales), s'est réorientée vers les grands centres urbains des alentours : Tuxtepec, en premier lieu, la ville de Oaxaca et bien entendu, la capitale du pays, le District Fédéral, qui draine un nombre considérable de migrants issus de l'exode rural. La seule enquête dans les Basses Terres exigerait donc une diversification fine des critères sociolinguistiques motivant le choix des informateurs. On peut s'attendre à une diversité d'idiolectes mixte, à une certaine labilité des idiolectes et des sociolectes, au-delà des isoglosses tracées avant, pendant ou après la construction du barrage, à l'époque où Sarah Gudschinsky menait ses enquêtes dialectales pour sa monographie comparative popolocan-mixtecan (Gudschinsky 1959).

1.2.2. Etat des lieux de l'enquête dialectale : comparatisme popolocan de Gudschinsky (1959) et monographie dialectale de Kirk (1966)⁶

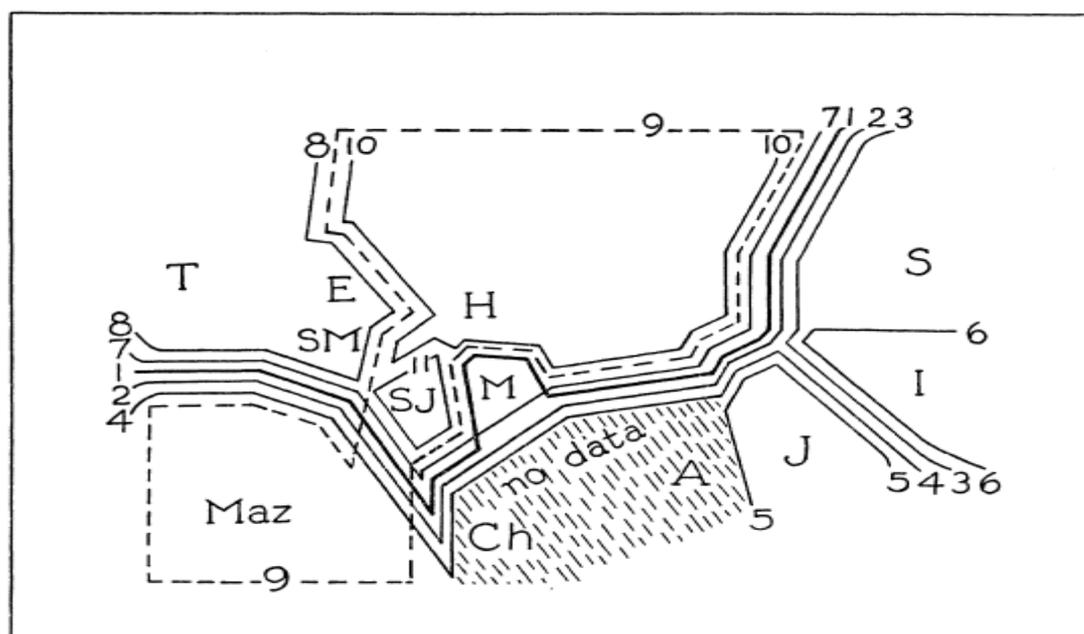
⁵ D'après Benítez 1993-[1970] : 15-16. Cf. aussi McMahon 1971.

⁶ La modélisation des données de Gudschinsky présentée ici a fait l'objet d'une communication préalable au colloque du réseau français de phonologie (Léonard 2010b). Je remercie chaleureusement Michael Swanton pour son aide documentaire *in extremis*, lors de la phase préparatoire du terrain, à Oaxaca, qui fut décisive.

Il existe deux grandes enquêtes dialectales réalisées au milieu du 20^{ème} siècle sur le mazatec : d'une part, celle issue des recherches comparatistes de grande ampleur de Sarah Gudschinsky dans les années 1950, publiées à la fin de cette décennie, dont nous venons de voir l'importance pour le monde socioculturel mazatec (Gudschinsky 1958, 1959), d'autre part celle de Paul Livingston Kirk (Kirk 1966), qui revisite et complète considérablement le travail de Gudschinsky, en resserrant le domaine comparatif au seul champ du mazatec, en affinant les étymologies et en diversifiant le réseau de points d'enquêtes. La dimension cartographique de ces deux études reste avant tout implicite. Hormis quelques cartes sous forme de croquis isoglottiques dans Gudschinsky 1958, l'approche comparative et analytique sous forme linéaire ou tabulaire domine chez les deux auteurs, qui n'aboutissent qu'à une vision incomplète ou fragmentée de la dynamique des aires phonologiques du mazatec. Il n'existe d'atlas linguistique pour aucune langue otomangue. On trouve des schémas, des croquis cartographiques, des représentations de faisceaux d'isoglosse, des tableaux de cognats à foison, des monographies comparatives, mais aucun atlas linguistique.

Comme on peut le voir dans la carte de la figure 1, reprise d'un article pionnier – tant sur le plan empirique, pour la qualité et la richesse de ses données mazatèques que pour ses objectifs interdisciplinaires entre dialectologie et ethnohistoire –, de Sarah Gudschinsky, publié en 1958 dans *Language*, des pans entiers du réseau dialectal demeurent inexplorés à cette date. Chiquihuitlán, qui figure en grisé en bas de la carte, dans une région signalée comme encore non documentée (cf. la mention « no data » dans le croquis), sera par la suite la variété la plus étudiée de mazatec par le SIL (*Summer Institute of Linguistics*, en espagnol ILV : *Instituto Lingüístico de Verano*), à travers d'importants articles ou monographies de Carole Jamieson (Jamieson 1982, 1996, 1998)⁷.

⁷ Le diagnostic de Paul Kirk en termes d'intercompréhension entre la variété de Chiquihuitlán de Juárez et les autres variétés de mazatec est impressionnant : l'index oscille entre 39 (avec Huautla) et 10 (avec Jalapa). Il est de 30 avec SM Soyaltepec. Comparé aux deux grands blocs de variétés en relation d'intercompréhension sur des seuils de 81 à 99 (autour de Huautla d'une part, de SM Soyaltepec d'autre part) et même à Jalapa (index 77 avec Huautla et 53 avec SM Soyaltepec), cette valeur est extrêmement basse. Pourquoi le SIL a-t-il donc tant investi dans la description d'une variété aussi périphérique de mazatec ? L'historiographie de la linguistique mazatec reste à faire, et serait d'ailleurs d'un grand intérêt pour l'histoire de la linguistique descriptive. Cependant, on peut se risquer à suggérer que ce choix s'explique probablement par les conditions locales privilégiées d'accès à la langue grâce à la coopération d'informateurs-clés, comme Ernesto Tejeda Salvador (décédé en 1986), à qui Carole VandenHoek Jamieson a d'ailleurs dédié sa grammaire du mazatec, basée sur cette variante (Jamieson 1988).



MAP. 2. PHONOLOGICAL ISOGLOSSES
The numbers of the isoglosses correspond to the numbering
of successive splits in Figure 1.

Fig. 1. Faisceau d'isoglosses phonologiques dans le réseau dialectal mazatèque, selon Gudschinsky 1958 : 472⁸.

Il serait légitime d'appeler le faisceau d'isoglosses de la figure 1 les *aires de Gudschinsky*, et de le considérer comme l'un des objets de connaissance fondateurs de la géolinguistique du mazatec. Ce faisceau reprend la division écologique et ethnoculturelle fondamentale en deux régions, haute et basse, explicitée par un stemma ou une arborescence *ad hoc* repris dans la figure 2 : la Mazateca Alta et la Mazateca Baja. La première se subdivise en Hautes Terres occidentales (San Antonio Eloxochitlán, San Jerónimo Tecoatl et San Mateo Huautla), *versus* Hautes Terres orientales (Huautla et Santa María Jiotes) sur la base de l'isoglosse 8 ($*tk > sk$) et 9 ($*tk > hk$) respectivement : assibilation dentale préconsonantique dorsale *versus* spirantisation préconsonantique dorsale. La seconde région est caractérisée par l'isoglosse 1, qui correspond à la dépalatalisation de l'occlusive coronale devant voyelles antérieures ($t^vi/e > t$). Les Basses Terres se divisent ensuite entre San Miguel ($t^vk > šk$) et la vallée de Soyaltepec, qui enserme dans l'isoglosse 2 ($Vhna/u > Vhña/u$) le sud (isoglosse 4 : $t^vk > tk$) et le nord, qui associe San Miguel Soyaltepec et Ixcatlán par la rétroflexion de l'affriquée palatale aspirée (en fait, *breathy*, à moins que ce ne soit, comme nous le verrons, la voyelle à être *breathy* devant voyelle basse), avec l'isoglosse 6, qui branche dans l'organigramme de Gudschinsky (fig. 2) une vocalisation du trait palatal ou vélaire des occlusives coronales prédorsales ($t^vk > tik$, $tk > tuk$), attestée dans la variété de San Pedro Ixcatlán⁹.

⁸ Les initiales des localités sont les suivantes : **A** = San Bartolomé Ayautla, **C** = San Juan Coatzospan, **E** = San Antonio Eloxochitlán, **H** = Huautla de Jiménez, **I** = San Pedro Ixcatlán, **J** = Jalapa de Díaz, **M** = San Miguel Huautla, **Maz** = Mazatlán de Flores, **S** = San Miguel Soyaltepec, **SJ** = Santa María Jiotes, **SM** = San Mateo Huautla, **T** = San Jerónimo Tecoatl.

⁹ Ce phénomène d'épenthèse connaît apparemment une extension encore bien plus grande que la seule variation libre en un point du réseau dialectal, comme phénomène impliqué dans le système de règles ordonnées du mazatec, à en croire la modélisation selon les principes de la *phonologie cyclique* (c'est-à-dire par *règles ordonnées*) proposée par un autre linguiste du SIL, Brian Bull (Bull, 1984 : 95, règle n° 6) pour la variété des Hautes Terres occidentales de San Jerónimo Tecoatl, très éloignée géographiquement et structurellement de celle des Basses Terres qu'est San Pedro Ixcatlán, selon la taxinomie diasystémique de Gudschinsky présentée ici.

La lecture de ces figures est particulièrement ardue pour un lecteur qui ne connaît pas la région, dans la mesure où plusieurs localités portent des noms semblables (ne pas confondre Hautla de Jiménez, dans les Hautes Terres, avec San Miguel ou San Mateo Huautla, dans les Basses terres). L'auteure a pallié à ce problème à l'aide d'un jeu d'initiales qui permet d'identifier les lieux pour le néophyte. Mais pour le dialectologue, cette classification pose des problèmes de cohérence : l'innovation $tk > hk$, qui spirantise une coronale prédorsale dans un groupe consonantique, se retrouve à deux endroits de l'arborescence de la fig. 2 - comme isophone des Hautes Terres orientales (Huautla de Jiménez et Santa María Jiotes) et comme isophone distinguant Mazatlán de Flores des autres variétés de la partie méridionale de la vallée de Soyaltepec, en dessous de la branche indexée 4 dans l'arbre de la fig. 2. L'auteure a tenu compte de cette ambivalence en traçant en pointillés l'isoglosse 9 dans le croquis de la figure 1.

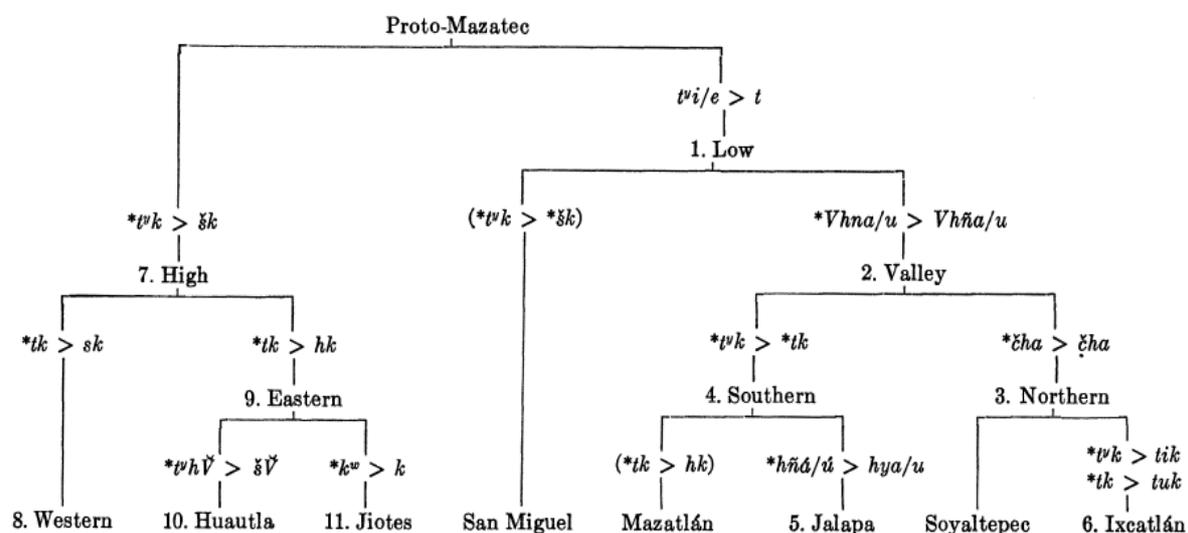


FIGURE 1. STEMMA OF THE MAZATEC DIALECTS

Numbers refer to isoglosses on Map 2

Fig. 2. Taxinomie des aires phonologiques dans le réseau dialectal mazatèque, selon Gudschinsky 1958 : 471.

Cette division fondée sur des contraintes de traitement de la palatalité des occlusives et des sonantes coronales, prévocaliques ou préconsonantiques (en ce cas, prédorsales), qui relève du consonantisme, n'est pas la seule retenue par Gudschinsky dans son *Etude historique en forme de miniature* de la variation dialectale du mazatec¹⁰. Je regrouperai ce faisceau d'isoglosses des figures 1 et 2 en un complexe (a) de variables. S'y ajoutent deux autres séries de variables fondamentales de la structure phonologique du diasystème¹¹ mazatec, listées (b) et (c) dans ce qui suit.

¹⁰ Une tentative de classification des dialectes mazatecs a été proposée par Paul Livingston Kirk dans une étude pionnière sur l'intercompréhension (Kirk 1970), dans le cadre des recherches du SIL dans ce domaine (qui lui était utile afin d'optimiser ses traductions de la Bible par le choix de variétés représentatives susceptibles d'atteindre un public aussi large que possible). Les résultats de cette étude donnent entre six et huit divisions, dont deux majeures (Alte et Baja Mazateca) regroupant chacune une dizaine de sous-variétés (op. cit. p. 209.).

¹¹ Cf. Weinreich 1954, pour une définition du *diasystème* comme métasystème qui subsume la diversité des structures d'un réseau dialectal, en termes de variables et de paramètres organisés en foncteurs.

- (b) des séquences -VhV-, soit à contour de timbre, comme *-ahu-*, soit analogues à des voyelles dites « réarticulées », selon la tradition mésoamérindianiste, comme *-i'i-* (l'apostrophe ou *saltillo* notant ici un coup de glotte) – cf. tableau 1 infra,
- (c) des alternances vocaliques **a/*u* en syllabe initiale des dissyllabes étymologiques, ou des aphérèses (Gudschinsky, 1948 : 479, carte 6).

Les trois séries de variables – (a) contraintes de palatalité pesant sur les non continues coronales, (b) gabarits CVCV de type CVhV ou CV'V et (c) variation libre **a/*u/∅* des noyaux initiaux – confirment dans les grandes lignes la configuration des aires du croquis de la figure 1 tracé par Sarah Gudschinsky : bipartition Hautes Terres occidentales *versus* orientales, et quadripartition des Basses Terres, Santa María Jiotes et San Miguel Huautla relativement autonomes au centre. Ces trois ensembles de variables résument l'essentiel de la modélisation phonologique du disystème mazatec selon Gudschinsky (il faudrait compter avec quelques autres structures, convergentes par ailleurs avec les trois modèles décrits ici). En somme, les complexes d'isophones de Gudschinsky se répartissent entre trois séries de faits qui constituent des théories, en tant que *micromodèles empiriques* décrivant les structures fondamentales du diasystème mazatèque: (a) un modèle de contraintes de stridence et d'aspiration des coronales prévocaliques ou en cluster, (b) un modèle de gabarits à hiatus glottalisé GO (*Glotte Ouverte*) de type CVhV et GF (*Glotte Fermée*) de type CV'V et (c) un modèle de réduction de la voyelle basse, par aphérèse ou par rehaussement.

Nous avons vu que le modèle (a) donne une bipartition est-ouest des Hautes terres, une quadripartition des Basses Terres, une ambivalence de Mazatlán et des îlots réfractaires au centre. Le modèle (b), comme le suggère le tableau 1, fait apparaître une division en dégradé entre l'est et l'ouest, vers lequel une variété centrale comme San Miguel Hautla tantôt converge (**dahu* « pierre », **di'i*¹² « feu » > **lahu*, **li'i*), tantôt diverge (**čahu* > **-su* « cendre »). Le modèle (c) complète la configuration suggérée par le modèle (a), et n'en diffère que par la complétude des données.

		Ouest		Est			Centre
Orientation	Proto-OM oriental	Sud-Ouest	Nord-Ouest	Centre-Est	Nord-Est	Sud-Est	
Localisation		Maz Mazatlán de Flores	T San Jerónimo Tecoahtl	I San Pedro Ixcatlán	S San Miguel Soyaltepec	A San Bartolomé Ayautlá	M San Miguel Hautla
'cendre' *-AhU-	<i>*čahu</i> >	<i>*čahu</i>				<i>*-su</i>	
'pierre' *-AhU-	<i>*dahu</i> >	<i>*lahu</i>	<i>*ntyuhu</i>				<i>*lahu</i>
'feu' *-I'I-	<i>*di'i</i> >	<i>*li'i</i>	<i>*ni'i</i>	<i>*nti'i</i>			<i>*lihi</i>

Tableau 1. Correspondances protodialectales dans le diasystème mazatèque : gabarits de Gudschinsky (Gudschinsky 1958 : 476, carte 3 et 477, carte 4. Graphie de l'auteur inchangée).

¹² Ou encore **ndahu* et **ndi'i*, selon les reconstructions. Sarah Gudschinsky note d'ailleurs ces étymons en variation libre dans le réseau dialectal mazatec, comme on peut le constater dans le tableau 1.

Le modèle (b) pourrait en réalité s'avérer le plus pertinent des trois sur le plan (dia)systemique, dans la mesure où il n'est pas lacunaire, comme le modèle (a) où les données de Chiquihuitlán sont absentes, et parce qu'il tient compte de la spécificité d'une importante variété de la Baja : Jalapa de Díaz, qui transphonologise le gabarit à hiatus glottal en monosyllabes à voyelles *breathy* (« pierre » **dahu* > *ⁿd^jo[□]^B* (UCLA)¹³ = <ndjio>) ou *creaky* (« feu » **di'i* > *ⁿdⁱ□^M* (UCLA)¹⁴ = <nd'í>).

Dans **dahu* > *ⁿd^jo[□]^B* (UCLA) = <ndjio>, le processus de compression syllabique et de *breathiness* de l'attaque résulte d'une syncope, phénomène par ailleurs fréquent à Jalapa, qui incorpore la fricative laryngale en position médiane au deuxième noyau V² : **dahu* > d[hu] <= /do□/¹⁵ = (« pierre », et non [dh]u ou /d□o/, en termes de constituance – c'est la voyelle qui est *breathy*, non la consonne. Le hiatus glottal de la séquence CVhV, qui a pour trait [+GO] (Glotte Ouverte, ou *spread glottis*), est ainsi incorporé au noyau de la deuxième syllabe. Le changement **di'i* > *ⁿdⁱ□^M* (UCLA) = « feu » se prête à la même analyse : la *creakiness* résulte de l'incorporation du hiatus [+GF] (= Glotte Fermée) de la chaîne CV'V au noyau destre qui reste seul en lice après syncope du noyau senestre : CV'V > C[']V > CV□.

Ce survol diachronique d'un réalignement de la glottalité dans des dissyllabes étymologiques, phénomène particulièrement attesté dans une variété comme Jalapa de Díaz¹⁶, va nous permettre d'aborder la question de la dialectique de complexité et de simplicité de l'inventaire phonémique mazatec, et de montrer les enjeux d'une modélisation typologique du système de la langue pour la conceptualisation d'un questionnaire et d'une enquête systématique pour l'ALMaz. L'enquête phonologique ne saurait se limiter simplement à répertorier les phonèmes en fonction d'une liste de correspondances attendues, selon la méthode néogrammaire. Une enquête phonologique est l'occasion de revisiter les théories passées et présentes du système d'une langue, en mettant celles-ci à l'épreuve du diasystème : son objectif est tout autant comparatif que typologique.

2. De la complexité de la langue à la parcimonie typologique et graphémique

A ce titre, on ne saurait aborder les langues otomangues sans un recul théorique, sans confronter plusieurs modélisations successives, tant une langue comme le mazatec a pu faire l'objet d'interprétations diverses. L'essentiel tient dans l'opposition entre, d'une part un modèle de constituance complexe, tel que celui de Pike & Pike 1947, qui eut d'ailleurs l'heur d'être l'une des sources historiques du modèle d'analyse syllabique en termes d'attaque et de

¹³ Archives sonores en ligne du laboratoire de phonétique de l'Université de Californie à Los Angeles, Peter Ladefoged & al. 2007. The UCLA Phonetics Lab Archive. Los Angeles, CA: UCLA Department of Linguistics. <http://archive.phonetics.ucla.edu/>. Variété : Jalapa de Díaz.

¹⁴ La *breathiness* en API/IPA se note par deux points souscrits sous le segment, tandis que la *creakiness* se note par un tilde souscrit.

¹⁵ Pour simplifier, je fais ici abstraction de l'indice palatal, qui est très discuté pour cette forme : en effet, la reconstruction de Paul Lingston Kirk est **ntyahu* (Kirk 1966) – il postule cet indice de palatalité dans l'attaque en contact avec une voyelle basse dans une chaîne hétéroorganique *-ahu-*, alors que Gudschinsky postule une chaîne homorganique *-uhu-* (op. cit.). Ces détails ne sont pas anodins, puisque Silverman, Blankenship, Kirk & Ladefoged (1987 : 113) font de l'homorganicité (V¹V¹) l'une des deux conditions de l'émergence sur le plan diachronique de la *breathiness* vocalique (condition 1 : que l'attaque initiale soit voisée comme dans les deux exemples ici examinés ; condition 2 : que la mélodie vocalique soit homorganique, ce qui est également le cas ici).

¹⁶ Dans une étape préliminaire de la théorie typologique appliquée au système phonologique du mazatec, qui définit ce système comme caractérisé par la *breathiness* et la *creakiness* plutôt que par l'aspiration et la glottalisation des attaques au sein de clusters consonantiques, comme c'était le cas dans l'approche des tagmémiciens (Pike & Pike 1947, Kirk 1966), seule la variété de Jalapa de Díaz était réputée avoir développé des segments *creaky* et *breathy* (Silverman, Blankenship, Kirk & Ladefoged *ibidem*). Golston & Kehrein (1998) étendent cette analyse à d'autres variétés de mazatec, notamment celle de Huautla, voire la généralisent à la langue dans toute l'extension de sa diversité dialectale, en revisitant les données de Pike & Pike 1947 pour la variété de Huautla.

rime, répartie entre noyau et coda, et d'autre part un modèle en termes de traits secondaires constitutifs de segments unaires, esquissé dans Silverman, Blankenship, Kirk & Ladefoged (1987) développé pleinement dans Golston & Kehrein 1998. Le tableau 2 confronte les deux inventaires. En bref, chez Pike & Pike (1947), les consonnes du mazatec sont des attaques branchantes, qui configurent à gauche ou à droite du segment-tête, coronal ou dorsal, des occlusives et des fricatives glottales, tandis que, selon Golston & Kehrein (1998), les traits de constriction et de souffle glottiques (respectivement GF et GO) sont des traits secondaires qui se répartissent entre les attaques et les noyaux. Les préaspirées et les préglottalisées en clusters de Pike & Pike 1947 sont en réalité des consonnes *breathy* et *creaky* (soufflées et craquées) qui ont fait remonter l'expression laryngée à gauche du segment, tandis que les aspirées et les postglottalisées, avec expression glottique à droite du segment doté de lieu coronal et dorsal, ne sont que de pseudo-aspirées et de pseudo-éjectives, dont le trait [GO] ou [GF] appartient en réalité au noyau syllabique. En résumé, la première analyse aurait mal découpé ou mal réparti les composantes des phonèmes, ce qui menait à une aporie. Les segments complexes étaient condamnés à être qualifiés de groupes consonantiques, tout en manifestant un comportement monosegmental à bien des égards, et une forte variation dans la position, destre ou senestre, que pouvaient prendre les expressions glottiques. Si l'on applique la théorie de Golston & Kehrein, on passe de 70 segments ou clusters contoïdes à 38, voire 41 si l'on reporte le premier élément des trois diphtongues du mazatec sur trois séries de consonnes palatalisées, labialisées et vélarisées, mais dont l'expression de surface du trait secondaire s'avère vocalisée. C'est littéralement une nouvelle répartition des cartes autour de la table de jeu que proposent Golston & Kehrein, pour un résultat convaincant, du point de vue de la parcimonie du système consonantique. En revanche, le coût est élevé pour le vocalisme, puisque cette modélisation implique désormais, en plus des voyelles orales et nasales, au moins deux autres séries de voyelles, soufflées (*breathy*) et craquées (*creaky*)¹⁷. Le système le plus répandu dans le réseau dialectal mazatec, hors de San Lorenzo étant pentavocalique (*i, e, a, o, u*), cette complexification du vocalisme reste cependant acceptable.

Consonantisme du mazatec de Huautla de Jiménez, selon Pike & Pike 1947. NB : notation API.	Consonantisme du mazatec de Huautla de Jiménez, selon Golston & Kerhein, 1998 : 319. NB : notation API.
<p style="text-align: center;"><i>t, ts, tʃ, tʃ̥, k ;</i> <i>tʔ, tsʔ, tʃʔ, tʃ̥ʔ, kʔ ;</i> <i>htsʔ, htʃʔ ;</i> <i>ht, hts, htʃ, htʃ̥, hk ;</i></p> <p style="text-align: center;"><i>nt, nts, ntʃ, ntʃ̥, nk ;</i> <i>ʔnt, ʔnts, ʔntʃ, ʔntʃ̥, ʔnk ;</i> <i>hnt, hnts, hntʃ, hntʃ̥, hnk ;</i> <i>ntʔ, nts, ntʃʔ, ntʃ̥ʔ, nkʔ ;</i> <i>nth, ntsh, ntʃh, ntʃ̥h, nkh ;</i></p>	<p style="text-align: center;"><i>t, ts, tʃ, tʃ̥, k ;</i></p> <p style="text-align: center;"><i>ṭ, tṣ, ṭʃ, ṭʃ̥, k̰ ;</i></p> <p style="text-align: center;"><i>nt, nts, ntʃ, ntʃ̥, nk ;</i> <i>nṭ, nṭs, nṭʃ, nṭʃ̥, nk̰ ;</i> <i>nṭ̰, nṭ̰s, nṭ̰ʃ, nṭ̰ʃ̥, nk̰̰ ;</i></p>
<p style="text-align: center;"><i>s, ʃ ;</i> <i>sh, ʃh, sʔ ;</i></p> <p style="text-align: center;"><i>m, n, ɲ ;</i></p>	<p style="text-align: center;"><i>s, ʃ ;</i></p> <p style="text-align: center;"><i>m, n, ɲ ;</i></p>

¹⁷ V. aussi Silverman 1997.

$\gamma m, \gamma n, \gamma p ;$ $m\gamma, n\gamma, p\gamma ;$ $mh, nh ;$ $l, l\gamma ; \beta, j ;$ $\beta\gamma, \gamma j ;$ $h\beta, hj ;$ $\beta h ; \gamma\beta, j\gamma$	$m, n, j ;$ $\underline{m}, \underline{n}, \underline{j} ;$ $\beta, j ; \beta, j ;$ β, j $Cj, Cw, C\gamma$ <i>(élimine les diphtongues)</i>
---	--

Tableau 2. Deux théories du consonantisme mazatec : Pike & Pike 1947 et Golston & Kerhein, 1998 : 319 ainsi que Silverman 2003

Le modèle de Pike & Pike 1947 joua un rôle fondamental pour la codification de la langue à des fins de littérature en langue mazatèque. Nous allons voir maintenant en quoi ce modèle induit un système de codification dont les ultimes conséquences dépassent les simples choix ergonomiques. Le choix d'un écrivain et phonologue – ou ethnolinguiste, selon la terminologie en vigueur au Mexique – qui a eu un impact décisif, dans les années 1990, sur la codification du mazatec¹⁸, consiste à réduire les phonèmes consonantiques de la langue à des segments simples, combinables en clusters aussi bien aux sonantes nasales qu'aux expressions glottiques continue (notée par la jota de la graphie espagnole <j>) et discontinue (notée par le « saltillo » <'>¹⁹).

	labial	labio-dental	dental	alvéolaire	alvéo-palatal	palatal	vélaire	Glottal
Occlusives	<i>p</i>		<i>t</i>				<i>k</i>	'
	<i>b</i>		<i>d</i>				<i>g</i>	
fricatives		<i>f</i>		<i>s</i>		<i>x</i>	<i>j</i>	
Affriquées					<i>ts</i>	<i>ch</i>		
					<i>ds</i>	<i>ˈch,</i> <i>(z)</i> ²⁰		

¹⁸ Pour une contextualisation sociopolitique de la codification et de l'élaboration linguistique mazatèque en cours depuis une trentaine d'années, cf. Duke (sans date), document de travail accessible sur Internet. Cf. aussi Benton 1999 pour une présentation des principes qui ont présidé à l'élaboration des graphies de langues indigènes du Mexique par le SIL. Ces graphies ont posé les jalons pour les formes de codification endogènes récentes, en suivant d'ailleurs de près les quatre principes énoncés par Benton (1) motivation maximale en fonction des conventions d'écriture la langue-toit, 2) optimisation des contrastes phonémiques, 3) mise en transparence des faits de langue au-delà des effets de compactage de la parole, 4) césures et attachement des mots composés selon la conscience lexicale des locuteurs. Généralement, les codifications endogènes ont tendu à atténuer le premier principe, afin de systématiser l'univocité des notations, en s'inspirant de la notation API (<k> constant pour noter une occlusive dorsale au lieu des allographèmes <c>, <qu>, etc.). Pour la notation des tons en mazatec, le deuxième principe a été appliqué en alliant pertinence maximale et parcimonie.

¹⁹ Pour la constriction glottique, ou trait [GF], notée <'> désormais dans la graphie, Carole Jamieson utilisait dans sa grammaire (Jamieson 1988) et dans son dictionnaire (Jamieson 1996) le graphème <h>, à distinguer de <j>, qui note dans son système de codification le trait de *breathiness* [GO]. Les implications typologiques d'un tel choix (qui n'a d'ailleurs pas été retenu par les locuteurs) sont grandes, au détriment d'une conception typologique en système de traits.

²⁰ Le graphème <z> est retenu, dans la Mazateca Alta, notamment à huautla, pour noter l'affriquée rétroflexe voisée, contre les graphèmes concurrents que sont, ailleurs, <ˈch>, avec trémas suscrits (à ne pas confondre avec la notation API de la *breathiness*) ou <chr>, dont la liquide graphique <r> note – quelque peu abusivement *a priori*, mais tel est le choix de ces scripteurs – la résonance sourde, quasiment rhotique, de la rétroflexion linguale.

Nasales	<i>m</i>			<i>n</i>		<i>ñ</i>		
Latérales				<i>l</i>				
Vibrantes					<i>r</i>			
Semiconsonnes						<i>y</i>		

Tableau 3. Inventaire consonantique et graphèmes pour la codification du mazatec, selon Juan Gregorio Rufino (1993 : 27)

Il s'agit, comme l'indique d'ailleurs le titre du tableau dans le manuel d'où est tiré ce tableau, d'un « cadre pratique du consonantisme », autrement dit d'une grille, qui a le mérite de faciliter l'acquisition d'une littéracie bilingue, en prenant pour base les principes d'organisation segmentale de la langue-toit qu'est l'espagnol. Cette approche se situe dans la continuité de la constituance des contoïdes complexes du mazatec en termes de groupes consonantiques, telle que la préconisaient les linguistes du SIL/ILV. Du point de vue de la typologie des langues, qui n'a finalement que peu à voir avec les exigences et les contraintes pratiques de la littéracie, cette théorie est pauvre : plus de traits de souffle ou de craquement glottique des coronales et des dorsales, plus d'occlusives pré-nasales²¹. Elle introduit même une corrélation de voisement dans les occlusives, qui est bien éloignée des caractérisations typologiques que l'on peut faire par ailleurs du consonantisme mazatec. C'est une théorie pauvre en idiosyncrasie typologique, mais riche en possibilités de combinatoire entre segments simples, efficace sur le plan ergonomique. Je qualifierai le modèle de Golston & Kehrein de « modèle riche en traits », et le modèle implicite de Juan Gregorio Regino de « modèle monosegmental ». Le tableau 4 ci-dessous reprend le détail des variables consonantiques des aires de Gudschinsky présenté dans la figure 2 supra, en déclinant, pour chaque processus décrit par l'auteure dans le diasystème mazatec, les conséquences du point de vue de ces deux théories. Le modèle riche en traits incite à une analyse des données en termes de spécification ou de neutralisation de traits, tandis que le modèle monosegmental ne pose guère de questions d'ordre typologique, et peut se contenter de qualifier de manière catégorielle, sans entrer dans la systémique des interactions entre segments ou traits infrasegmentaux, les principaux changements. Il est important ici d'opposer ces deux visions des choses car, lorsque je décrirai le déroulement de l'atelier de protomazatec à Huautla, nous verrons que la seule démarche pertinente sur le plan pédagogique aura été à cette occasion celle du modèle monosegmental. Une analyse en termes de modèle de traits aurait été extrêmement onéreuse en temps et en explications. L'intérêt de la linguistique réside entre autres dans l'éventail d'explications que la maîtrise des règles de l'art, en tant que *techné* descriptive, permet de projeter sur des phénomènes de langue : la modélisation revêt une géométrie variable selon les objectifs et les publics. La finalité de l'argumentation en typologie linguistique et en recherche fondamentale n'est pas la même qu'en linguistique appliquée. Or, même si le (géo)linguiste a pour but de décrire la variation de la langue à travers cet outil qu'est un atlas linguistique, il ne peut se passer d'une réflexion à multiples vecteurs sur le fonctionnement de la langue et les défis empiriques que pose la variation à la caractérisation typologique de la langue. Les colonnes *Spec.Trait* et *Modèle riche en traits* du tableau 4 sont, de ce point de vue, stratégiques. Elles ne présentent que peu d'intérêt pour le praticien de la langue, engagé dans le développement de la littéracie, mais elles posent des questions au typologue sur la relation entre *stridence* et *breathiness* dans une langue dotée de traits secondaires glottiques comme le mazatec. Et quel est le statut des expressions [GO] issues de la spirantisation, dans les séries (4) et (9) du tableau ? Comment catégoriser la phase [GO] pour la variable (5) du même tableau, dans un modèle riche en trait ? La *breathiness* s'aligne-t-elle désormais sur un noyau *breathy*, après que la sonante nasale s'est amuïe, ou

²¹ Bien qu'il soit probable que les pré-nasales otomangues ne sont autres que des sonantes nasales postoralisées, cf. Marlett 1992, pour des arguments convaincants en ce sens.

doit-on la considérer comme seule attaque, et en ce cas, que faire des attaques remplies uniquement par un trait [GO], désassocié de l'attaque sonantique coronale ?

Changement phonétique (notation API)	Processus phonologique	Spec.Trait	Modèle riche en traits ²²	Modèle mono-segmental
1) $t^j i/e > t$	dépalatalisation coronale	-	Neutralisation de stridence	Simplification coronale
2) $Vn \square a/u > Vn \square^j$	palatalisation sonantique breathy	[+strident]	Stridence associée à GO	Palatalisation de nasale
3) $t \square a \square > t \blacksquare a \square$	rétroflexion affriquée préconsonantique breathy	[-réparti]	Dissimilation de stridence et de GO	Rétroflexion
4) $t^j k > hk$ (= $k \square$)	spirantisation dépalatalisée préconsonantique dorsale	[+continu]	Neutralisation de stridence en faveur de GO	Cluster -jk- < -tyk- préaspiration
5) $\square \square á/ú > j \square a/u$	oralisation d'une sonante breathy	[approx.]	Dissimilation de nasalité et de GO	$j\tilde{n} > jy$ Dénasalisation
6) $t^j k > tik$ $tk > tuk$	épenthèse vocalique haute préconsonantique dorsale	[+vocoïde]	Syllabation d'acuité et de gravité	Vocalisation
7) $t^j k > \square k$	assibilation palatale préconsonantique dorsale	[+continu]	Contour de plosion régi par une attaque dorsale	-tyk- > -xk- -tk- > -sk- Assibilation
8) $tk > sk$	assibilation dentale préconsonantique dorsale			Contour de lieu de plosion réduit à une spécification GO
9) $tk > hk$ (= $k \square$)	spirantisation préconsonantique dorsale		Neutralisation GO par stridence continue	
10) $t^j V \square > \square V$	assibilation palatale préconsonantique breathy			
11) $k^w > k$	délabialisation dorsale	-	Neutralisation de labialité	Délabialisation

Tableau 4. Processus et spécification de traits des variables fondant la taxinomie des aires phonologiques dans le réseau dialectal mazatèque (Gudschinsky 1958 : 471). NB : notation API.

Nous allons bientôt constater, en examinant les données élicitées par écrit lors du premier exercice, lors de l'atelier de protomazatec, que de telles questions ne sont pas oiseuses, et qu'elles sont implicitement posées par les choix graphiques des scripteurs, praticiens de leur langue. En réalité, le raisonnement phonologique doit s'articuler entre ces deux pôles théoriques : riche en trait ou monosegmental, pour comprendre le fonctionnement de la variation dans une langue – et pas seulement une langue otomange, car somme toute, ce

²² Le cadre théorique d'analyse en traits distinctifs sous-jacent à l'ensemble de l'argumentation relevant de ce modèle est celui de Clements & Hume 1996.

regard, toutes choses égales par ailleurs, est transposable même sur l'évolution des langues romanes (cf. Léonard & dell'Aquila 2010²³).

3. Atelier de dialectologie mazatèque à Huautla

3.1. Méthodologie de « l'atelier de protomazatec »

L'expérience que je vais décrire ici nécessite de bousculer quelque peu le plan qui eût été canonique pour cet article, car ce n'est que maintenant que je vais présenter les rudiments de comparatisme otomangue et popolocan, puisque c'est ainsi qu'a commencé le séminaire que je vais décrire. J'aurai aussi recours à une visualisation par le biais de la photographie (séries de clichés 1a-c et 2a-c infra) de quelques documents réalisés à cette occasion. Les 18 et 19 août 2010, j'ai eu l'occasion d'organiser un atelier ou *workshop* de dialectologie et de phonologie historique mazatèque à Huautla de Jimenez, dans l'Etat de Oaxaca, avec 25 instituteurs bilingues, animateurs culturels, inspecteurs d'académie et autres acteurs de l'éducation bilingue et interculturelle dans la région des Hautes Terres mazatèques. Ce séminaire intitulé « atelier de protomazatec » a eu lieu dans la Casa de Cultura Maria Sabina, au centre de Huautla. Il s'est déroulé durant deux pleines journées, pour une durée totale de 16 heures d'intense travail collectif. L'objectif de ce stage était de former les enseignants des écoles bilingues ou tout intervenant culturel intéressé par l'élaboration du corpus ou par l'histoire de sa langue, ainsi que de profiter de cet espace de travail pour débattre de questions relatives à l'aménagement linguistique du mazatec – pour le moment dispersé entre une grande diversité d'initiatives qui ne semblent pas encore suffisamment coordonnées en un projet global.

Il n'est guère possible de rendre compte du deuxième aspect – le débat –, si ce n'est que la travail systématique sur la variation dialectale que je vais bientôt détailler a été l'occasion d'une prise de conscience quant à l'importance d'envisager un plan stratégique à moyen et long terme pour la standardisation pluraliste de la langue et surtout, pour revaloriser son statut et tenter d'enrayer ou de renverser la tendance diglossique qui risque de conduire la langue vers sa substitution progressive par l'espagnol, y compris dans le milieu familial et le voisinage, pourtant bastions de résistance des vernaculaires. On notera simplement l'effet boule de neige de tout travail de réflexion avec un collectif de locuteurs sur la structure de la variation dialectale de la langue : tout travail un tant soit peu soutenu sur cette question mène à une prise de conscience de l'importance de concerter l'action en faveur d'une extension des domaines d'usage, des fonctions, et une revalorisation du statut de la langue. Ne serait-ce que parce qu'il est impossible de manier des faits dialectaux sans immédiatement voir surgir pleinement des structures logiques, régulières, en somme, sans voir surgir la *langue* au-delà de la *parole*, le *système* au-delà de la variation des *structures*. Le sens même de la désignation péjorative de « dialecte » – dans ce contexte culturel, « dialecto » – s'en trouve immédiatement remis en question dans son acception courante, conditionnée par les valeurs basses de la diglossie, et c'est à travers la praxis de la comparaison des formes que remonte à la surface la *valeur dialectique* de la variation dialectale.

Dans ce qui va suivre, je présenterai une sorte de rapport sur l'activité en question, dans le but de montrer non seulement comment ce genre d'activité est propice, sur le plan de la praxis, à une réflexion qui implique des membres particulièrement conscients et motivés de la communauté linguistique dans l'élaboration d'un atlas linguistique, mais aussi, sur le plan de

²³ Dans l'article en question, nous montrons comment, selon le point de vue théorique adopté, les groupes consonantiques secondaires romans -P'L-, -K'L-, -T'L- par exemple peuvent être considérés, selon les réflexes attestés dans les réseaux dialectaux sardes et rhétoromans, non comme des clusters mais comme des segments complexes ou enrichis de traits secondaires sous l'effet de la coalescence de la phase d'explosion latérale avec la phase implosive occlusive : ce faisant, nous opposons un modèle riche en traits à un modèle monosegmental.

la modélisation, en quoi cette réflexion permet de mieux comprendre la nature des catégories diasystémiques et leur variation.

Ce stage ou séminaire s'est déroulé de la manière suivante :

- 1) Conférence présentant dans ses grandes lignes le comparatisme otomangue et popolocan.
- 2) Travaux dirigés sur les méthode de reconstruction appliquées au mazatec, à travers l'observation de la variation dialectale par Sarah Gudschinsky (1958), puis par Paul Livingston Kirk (1966).
- 3) *Premier exercice d'application* : compléter en notant les réflexes d'une ou plusieurs variétés dialectales de la Alta une liste d'étymons établie par Paul Kirk (Kirk 1966) et identifier les processus phonologiques observables, entre l'étymon de Kirk et *la* ou *les* variété(s) des Hautes terres élicitée par écrit durant le travail de groupe. Je donnerai un compte rendu aussi détaillé que possible des résultats de ce premier exercice.
- 4) Microconférences des groupes de participants présentant le résultat de leurs travaux, entièrement filmées à l'aide d'un caméscope numérique et enregistrées sur ZOOM, et correction de l'exercice par le dialectologue formateur.
- 5) *Deuxième exercice d'application* : compléter une simple liste de cognats, tirée du manuel de codification du mazatec de Juan Gregorio Regino *Alfabeto mazateco* (Regino 1993), qui est bien plus qu'un abécédaire ou un simple fascicule d'introduction à l'écriture du mazatec. Une fois les correspondances dans quatre variétés dialectales des Basses terres²⁴ compilées sur tableur, cet ouvrage fournit une base de données que j'ai distribuée sous formes de fiches d'une page contenant une cinquantaine d'items chacune, tous retenus par Juan Gregorio Regino en fonction de leurs caractéristiques phonologiques. Il s'agissait encore de compléter avec les données de Huautla et d'autres variétés des Hautes Terres, mais aussi de remplir la colonne des étymons, laissée vide. Pour ce faire, il était nécessaire d'appliquer la méthodologie apprise lors de la conférence et de la correction de l'exercice précédent sur les critères et les processus structuraux permettant de reconstruire une protoforme. Nous n'aurons guère la place ici pour commenter les résultats de ce deuxième exercice.

Je vais reprendre chaque étape de cette activité : en premier lieu, les rudiments de comparatisme otomangue présentés lors de la conférence initiale (section 3.2.1.), puis les éléments de méthodologie de la reconstruction du protomazatec (section 3.2.2.), les résultats du premier exercice d'application sur l'identification des processus évolutifs avec les données élicitées par écrit afin d'alimenter en données un premier microtableau comparatif (section 3.2.3.), puis les résultats du second en termes de façons de procéder des participants (section 3.2.4.). Si les analyses des participants s'écartent peu de ce qui était hautement prévisible – le campanilisme a joué une fois de plus son rôle d'écran de fumée, malgré la bonne volonté de tous d'appliquer la méthode comparatiste sur les données que je leur offrais et celles qu'ils produisaient –, nous verrons que le jeu en valait la chandelle ne serait-ce que pour la richesse des données écrites élicitées à l'aide de cette méthode. Le procédé rappelle d'ailleurs les premières heures de la dialectologie moderne, lorsque George Wenker faisait remplir à des instituteurs des questionnaires envoyés par la poste, afin de constituer son atlas linguistique de l'Allemagne (le *Deutscher Sprachatlas*) avec des données de seconde main.

Cependant, la méthode utilisée ici diffère de celle de George Wenker et l'enrichit par plusieurs aspects fondamentaux : a) le médium et le temps : la collecte de formes graphiques

²⁴ A savoir, en mazatec, suivies du toponyme officiel, les variétés *én ningotsié* (San Pedro Ixcatlán), *én naxijén* (San Miguel Soyaltepec), *én ntá xjò* (Jalapa), et *ngati'a* (San José Independencia).

et de réponses se fait en temps réel, et non pas de manière différée par voie postale, b) les supports diamésiques : aussi bien des versions écrites qu'orales sont recueillies, car durant les microconférences, les instituteurs lisent les formes qu'ils ont écrites (ou *formes autoélicitées*), c) le feedback en contexte de formation : l'ensemble de la collecte de données écrites et orales se déroule dans une relation dialogique avec le dialectologue, en développant une réflexion sur les structures dialectales de la langue, dans une finalité de dialectologie appliquée. Ces trois points font de ce type d'ateliers une innovation méthodologique, par rapport à la méthode de Wenker et de la collecte par consultation de ces « semi-experts » que sont les maîtres d'école ou les érudits locaux en Europe. Enfin, j'ajouterai que cette méthode est d'autant plus fertile que le mazatec n'est pas, comme l'allemand et les dialectes germaniques, une langue amplement décrite et abondamment attestée sur le plan philologique mais, comme nous l'avons vu dans la première partie de cet article, à la fois l'une des langues les plus complexes et les plus originales sur le plan typologique qu'on puisse décrire et l'une des langues décrites de la manière la plus fragmentaire, sous forme d'un patchwork théorique et comparatiste. Il ne s'agit pas d'une langue indo-européenne facile d'accès et dont le système phonologique serait établi une fois pour toute et cependant, il s'agit d'une des langues qui a le plus contribué au renouvellement de la phonologie moderne, depuis l'article de Pike & Pike en 1947, dont nous avons vu les implications paradoxales (bonne méthode d'analyse, novatrice par l'esquisse de modélisation de la constituance, mais application erronée sur les données de la langue). On pourrait ajouter donc une quatrième pierre de touche à la liste de points d'écart avec la méthode de Wenker : d) réflexivité et empirisme critique. En effet, les variantes dialectales recueillies dans une perspective diamésique au cours de ce type d'atelier ou de séminaire contribuent à alimenter les réflexions sur un système phonologique non seulement pertinent sur le plan typologique, mais qui pose aussi des questions de fond en phonologie générale.

3.2.1. Rudiments de comparatisme otomangue

Le récit du déroulement de ce stage/séminaire de dialectologie mazatec ne sera pas un simple compte rendu, dans la mesure où je tenterai de montrer des paradoxes ou des enseignements d'un intérêt plus général. Cet « atelier protomazatec » avait été précédé d'une réunion le lundi 9 août à San Andrés Hidalgo, dans la périphérie rurale de Huautla de Jiménez, avec le directeur de la Casa de Cultura Maria Sabina et son frère, tous deux auteurs d'un projet triennal de promotion de la culture mazatèque dans les Hautes terres. L'idée avait alors germé d'organiser un séminaire de ce type, dont j'avais acquis l'expérience lors de ma coopération bénévole avec l'ALMG (*Academia de las Lenguas Mayas de Guatemala*). J'avais réalisé auparavant trois séminaires analogues auprès de linguistes et d'enseignants de langues mayas : l'un sur la morphosyntaxe du protomaya à Cobán, au nord-est du Guatemala, avec des participants poqomchi' et q'eqchi', en 2002, et deux autres sur dialectologie et standardisation ; en domaine q'anjob'al, à Jacaltenango en 2005 et à San Miguel Acatán en 2006. Deux autres ateliers sur la standardisation de langues mayas quichéanes, le tz'utujil et l'achi', en 2003, avaient précédé ces deux tentatives de construire une dialectologie appliquée de langues mésoaméricaines en collaboration étroite avec des linguistes et des instituteurs ou d'autres praticiens de la langue. Mais dans le cas de l'atelier de protomazatec, l'objectif fixé initialement lors de la réunion du 9 août n'était pas d'aborder la standardisation de la langue²⁵,

²⁵ Je ne m'y risquerais d'ailleurs pas sans prendre de l'élan. La standardisation d'un complexe dialectal comme le mazatec présente un degré de difficulté analogue à celui qu'a dû affronter Koldo Mitxelena/Luis Michelena et ses épigones pour standardiser le basque (euskera), qui compte au moins sept variétés structurellement très différenciées. En revanche, l'achi' ou le tz'utujil sont des langues mayas quichéanes en situation de continuum dialectal à granularité fine, dont l'individuation est davantage ethnopolitique que fondée sur une distance structurale importante.

mais plutôt d'apporter des connaissances sur l'histoire de la langue et de la culture mazatèques à l'aide des données de la géolinguistique, et de familiariser les participants à l'approche comparative des variétés dialectales du mazatec, qui pose un problème de taille à l'éducation bilingue et interculturelle dans cette région, en raison de la grande distance structurale entre les variétés, notamment entre les deux grands ensembles dialectaux que sont la Alta et la Baja. Faire connaître les divisions majeures, ou les lignes de partage (les faisceaux d'isoglosses), la succession chronologique des innovations structurales et leur distribution polycentrique, faire apparaître et expliquer aussi bien la régularité de certains changements phonétiques, et la variation libre ou le polymorphisme observables pour d'autres variables, en fonction des contacts, des échanges et des relations de prestige ou d'interdépendance entre les différents centres urbains qui constituent le réseau dialectal mazatèque, tels étaient les objectifs immédiats de ce séminaire.

La plupart des participants n'avaient pas une conscience ou une connaissance nette de l'affiliation génétique de leur langue, du point de vue de la classification des langues de Mésoamérique. S'ils avaient eu vent que le mazatec était une langue otomangue, ils n'en savaient guère plus sur le détail de la classification – par exemple, le rattachement du mazatec à la sous-famille popolocane. L'extension géographique du phylum otomangue, sa configuration, la profondeur historique de ses subdivisions restaient des concepts flous, ce qui pour quiconque a travaillé dans un pays comme le Guatemala, où la carte des langues mayas avec la distribution des sous-familles est omniprésente et connue de tous les techniciens des langues et des cultures du pays, ne laisse pas de surprendre. La parenté d'une langue comme le mixtec, ou tu'un savi, pourtant en contact étroit avec le mazatec des Hautes Terres dans l'enclave multilingue de San Juan Coatzacoapan, n'était pas connue de tous les participants. J'avais pris soin de distribuer un exemplier de deux pages denses, qui incluait la figure 3 et le tableau 5 infra, afin de montrer l'extension et la configuration des langues otomangues au Mexique et de donner quelques exemples de la (relativement) proche parenté entre le mazatec et les trois autres langues popolocanes (toutes sévèrement en danger) : l'ixcatec, le chocho ou *njigua* – toutes deux au bord de l'extinction – et le popoloca à proprement parler.

La carte schématique extraite de la thèse de Kathryn Josserand sur l'histoire dialectale du mixtec/tu'un savi, reproduite dans la figure 3, rend clairement compte de l'extension territoriale des langues otomangues en Mésoamérique, et de la position du popolocan. Ce phylum divisé entre le groupe occidental (le pame-otomi, cf. Soustelle 1937), au nord de Puebla, et le groupe oriental incluant toutes les langues à l'est d'une ligne Puebla-Acapulco, est l'un des plus anciens de cette région du monde. Terrence Kaufman fait remonter le proto-otomangue à 6 000 ans et situe la dichotomie entre les deux principales branches, otomangue occidentale et orientale, à 5 000 ans (Kaufman 2006). L'intense différenciation interne de ce phylum a déjà été évoquée plus haut. Il suffit de dire que, selon cet auteur, la subdivision majeure de l'otomangue orientale opère entre d'une part l'amuzgo-mixtécane (4 000 ans) et d'autre part le popolocan-zapotécane (3500 ans), avec le mazatec constituant un rameau individué face au chocholtecan, au sein de la sous-famille popolocane. La fragmentation dialectale du mazatec se déroule, dans le scénario de Kaufman, sur un millénaire. Cette vision des choses peut être considérée comme d'autant plus acceptable que prudente.



Figure 3. La sous-famille popolocane dans le phylum otomangue
(extrait de Kathryn Josserrand, 1983. *Mixtec Dialect History*,
Ph. D. Dissertation, New Orleans, Tulane University)

A lire une liste de correspondances comme celle du tableau 5, qui reprend quelques séries de cognats identifiés par Calvin Rensch, qui compare le protomazatec à l'ixcatec et au njigua, les langues popolocanes semblent proches entre elles. La réalité est cependant bien plus complexe, et en dehors de listes soigneusement étudiées pour emboîter les pièces du puzzle, comme celles que Rensch a savamment conçues pour sa grammaire historique, le lexique présente des différences constantes²⁶. Le tableau 5 permettait de poser les jalons de la méthode comparatiste : la deuxième colonne mentionne le numéro de la série de correspondances dans la nomenclature de Rensch, et le mazatec est représenté par la protolangue reconstruite, afin d'éviter de papillonner d'une variété dialectale à l'autre. La justification de ces étymons était aisée à faire, en comparant avec les formes modernes de diverses variétés connues par les participants, originaires de Huautla mais aussi, pour quelques uns, de Mazatlán de Flores ou de San Miguel Huautepéc.

	Num. cognada Rensch 1976	Proto- mazateco	Ixcateco	Chocho (njigua)	Popoloca
dix	9	*te	'ute	te	te
vingt	57	*kan	xkan	kjan	ka
trois	238	*xjan	nijen	nie	ni

²⁶ J'ai eu l'occasion lors de mon séjour dans la Alta Mazateca de comparer avec la variété de Huautla, à l'aide d'une locutrice de mazatec, une liste comptant une centaine de mots et de syntagmes simples élaborée par l'équipe de UCLA pour l'étude phonologique du popoloca de San Felipe Otlaltepec. La dispersion structurale des deux langues en termes de lexique était constante pour plus de 80% des items du corpus, bien que l'affiliation au sein d'une même sous-famille ne fasse pourtant aucun doute. La méthode comparatiste ne saurait se contenter de listes triviales ou routinières, comme celles de Swadesh. La dialectique de cohérence génétique et typologique d'une part et de diversification forte du lexique et des structures grammaticales d'autre part entre langues otomangues est, de ce point de vue, un cas d'école. Que Morris Swadesh ait pourtant bâti sa théorie de la continuité du lexique fondamental en domaine mésoaméricain en dépit de la masse de données otomangues – qu'il connaissait bien par ailleurs – observable au seuil de sa porte qui contredisait cette vision des choses est un cas de cécité empirique intéressant, du point de vue de la (psycho)sociologie des sciences.

doux	260	*'inta	'inda	-	nta
salive	34	*ntatya	ndyatya	ndatya	nta'tan
marmite	39	*tiji, *ntiji	chiji	chigo (taza)	chi'
blanc	312	*tyawa	ruwa	rua	tjuwa
fil	136	*nachjun	chjun (algodón)	chjun	chjun'

Tableau 5. Liste de cognats popolocan, otomangue oriental
(d'après Rensch [1966]-1976 : 89-116).

Une fois ces rudiments de comparatisme otomangue présentés au public, nous avons récapitulé tous les éléments présentés dans la section 1.2.2. du présent article. Les figures 1 et 2, respectivement le faisceau d'isoglosses du consonantisme et l'organigramme de classification des aires dialectales mazatec par Sarah Gudschinsky, selon la série (a) de variables, inclus dans l'exemplier, furent analysés en détail. Nous sommes ensuite passés à la comparaison interne au réseau dialectal mazatec à partir des données de Kirk (1966). Le tableau 6 est un échantillon d'un extrait d'une base de données en cours de constitution, qui recense toutes les attestations figurant dans les séries de cognats de Kirk (*op. cit.*).

3.2.2. Éléments de méthodologie de la reconstruction du protomazatec

Dans le tableau 6, qui recense six variétés de mazatec, les innovations sont signalées en caractères gras. La numérotation en première colonne permet d'identifier les séries que nous allons rapidement analyser, en mentionnant entre parenthèses l'index de la série de cognats de la source, suivi de la traduction. Les formes ont été réécrites selon les conventions graphiques modernes du mazatec – celles utilisées dans le système éducatif. Ce microcorpus, qui a servi d'illustration de la méthode durant l'atelier de protomazatec, fait apparaître aussi bien des faits déjà observés plus haut que des phénomènes dont le survol de la classification de Gudschinsky dans la section 1.2.2. ne tenait pas compte, par souci de concision. On ne pourra guère commenter les tons des réflexes, car ils sont notés de manière sporadique dans la thèse de Kirk (1966), en dehors des étymons, pour lesquels l'auteur a reconstitué des séquences prosodiques. La notation retenue ici est celle inspirée du SIL/ILV, adoptée par les locuteurs : un accent aigu pour le ton haut (*á*), un accent grave pour un ton haut moyen ou M+ (*à*), aucun diacritique pour un ton moyen bas ou M- (*a*) et la voyelle soulignée pour le ton bas (*a*).

N°		N° Kirk 1966	Étimol.	Maza- tlán	Jalapa	Huautla de J.	Jiotes (Santa Maria)	Soyal- tepec	San Loren- zo
1	quatre	426	*ñunjun	ñunju	ñjun	ñ <u>un</u>	ñinjun	ñunjú <u>n</u>	ñinjin
2	gousse de maïs	329	*ninjin	ninjin	njin	nínj <u>ín</u>	ninjin	nínj <u>ín</u>	nenjen
3	papier	560	*xujun	xujun	xujun	x <u>un</u>	xun	xujun	xöjön
4	deux	128	*jau	jo	jo	jau	jo	jó	jwö
5	oisif-	690	*'a	-	'a	'a	-	'a	o
6	dix	570	*te	te	te	te	te	te	ta
7	proche	621	*tyisñaán	tiñan	tiñan	chian	chiñan	tiñ <u>a</u> n	chiñon
8	huipil	45	*tsú'ú	ts'u	ts'ú	ts'ú	ts'u	ts'u	tsö'ö

Tableau 6. Micro-corpus : du protomazatec à six variétés modernes, d'après les données de Kirk 1966.

Dans les séries numérotées de 1 à 3 (indexées 426, 329 et 560 dans le lexique étymologique que constitue la liste des séries de cognats in Kirk 1966), nous retrouvons la syncope en contexte homorganique à Jalapa (**ñunjun* > *ñjun*, **ninjin* > *njin* = « quatre ») mais bloquée par l'alternance oral/nasal dans **xujun* > *xujun* = « papier » ; une tendance à la compression syllabique conditionnée par le timbre à Huautla (**ñunjun* > *ñuun* et **xujun* > *xun* mais **ninjin* > *nínjin*) ; un cycle vocalique (ou *Vowel Shift*) à San Lorenzo (**ñunjun* > *ñinjin*, **ninjin* > *nenjen*, **xujun* > *xöjön*, où <ö> note /ɨ/) avec délabialisation de la séquence *u-u*, abaissement de la chaîne *i-i* et rétraction²⁷ dans la séquence *u-un*. Les séries numérotées de 4 à 6 (cognats de Kirk n° 128, 690, 570) sont représentatives de la variation du vocalisme mazatec dans des formes monosyllabiques : la diphtongue décroissante vélaire **au* peut se maintenir comme à Huautla, se monophthonguer comme à Mazatlán, à Jalapa et à Soyaltepec (**jau* > *jo*, *jó* = « deux » avec ton haut à Soyaltepec), ou connaître une inversion de constituance en devenant croissante, comme à San Lorenzo, où il faut aussi compter avec la rétraction vocalique (**jau* > *juö*). La voyelle basse *creaky* (car préglottalisée) se maintient généralement, mais se labialise à San Lorenzo (**'a* > *o* = « inactif », « paresseux »). La voyelle antérieure moyenne /e/ se maintient également, mais s'abaisse à San Lorenzo (**te* > *ta* = « dix »). Ces faits confirment la forte tendance de San Lorenzo à modifier de manière systémique son vocalisme aussi bien dissyllabique que monosyllabique, contribuant à signaler cette variété comme un centre d'innovation autonome. On voit également que les aires vocaliques fédèrent des zones des Hautes et des Basses terres que les aires consonantiques de Gudschinsky séparaient en ligne claire (neutralisation par le vocalisme des faisceaux d'isoglosses consonantiques). La série suivante n° 7 dans notre tableau (n° 621 de Kirk), relève des aires de Gudschinsky étudiées plus haut : on retrouve la dépalatalisation de l'occlusive coronale caractéristique des parlers des Basses Terres (**tyisñaán* > *tiñan*, *tiñaan* = « proche »), qui contraste avec l'affrication palatale des Hautes Terres (**tyisñaán* > *chian*, *chiñon*). Enfin, le n°8 (série de cognats n°45 chez Kirk) montre un fait de consonantisme stable (l'affriquée dentale **ts* dans **tsú'ú* = « huipil, blouse brodée »), avec un jeu de reconfiguration du gabarit CVCV avec hiatus [+GF], autrement dit une forme *CV'V, avec voyelle réarticulée (séquence **-ú'ú-* à ton haut) après apocope (**tsú'ú* > *ts'ú*, *ts'u*), pour laquelle on peut prédire des expressions de type *creaky* du noyau, sauf à San Lorenzo, où le hiatus est maintenu, et où la voyelle réarticulée subit la rétraction (**tsú'ú* > *tsö'ö*).

Une telle liste illustre de manière élémentaire la méthode de reconstruction. On part tout d'abord de grandes tendances ou de généralités, telles que : a) on ne peut rien attendre de San Lorenzo pour reconstruire le timbre des voyelles, puisque tout le vocalisme est engagé dans un *Vowel Shift* ou chaîne de traction systémique, b) Jalapa mais aussi huautla présentent des contractions de gabarits (syncope à Jalapa, monophthongaison à Huautla) qui les écartent comme variétés candidates pour la reconstruction des structures syllabiques, c) les variétés des Hautes terres sont palatalisantes, comme l'indiquent les affrications qui s'y produisent, si bien que l'on doit restituer des occlusives coronales davantage alvéolaires pour un état plus ancien. Cependant, la liste reproduite ici est incomplète, pour des raisons d'espace typographique – celle utilisée durant le stage comprenait d'autres variétés, comme Chiquihuitlán, qui permet de reconstruire la voyelle articulée dans la dernière série (*tsu'u* < **tsú'ú*). Plus en détail, dans le tableau 6 en (1) le vocalisme nasal de l'étymon **ñunjun* est identifiable par les réflexes de Jiotes, de San Lorenzo et de Soyaltepec, qui nous renseigne aussi sur le timbre. En (2), **ninjin* présente une séquence de voyelles nasales encore plus répandue dans le réseau dialectal, dont le timbre palatal haut est confirmé par la majorité des

²⁷ Dans les graphies modernes du mazatec, le graphème <ö> note une voyelle rétractée. La notation des linguistes est /ɨ/, aussi bien chez Kirk pour le mazatec que chez Jossierand pour le mixtec. Ne pas confondre avec la notation <ö> de l'amuzgo, qui fait pendant à <ë>, qui notent respectivement des voyelles moyennes mi-basses ou mi-ouvertes vélaire et palatale.

réflexes de la série. Il en va de même en (3) pour **xujun* (dont le palier de tons bas est inféré des formes attestées à Huautla et à Soyaltepec). En (4), la constituance syllabique de la diphtongue décroissante, se simplifiant en monophthongue (**jau* > *jo*) est davantage plausible qu'une monophthongaison, et cette hypothèse est confirmée par l'inversion de constituance à San Lorenzo, qui laisse penser qu'on part d'une diphtongue d'un certain type pour aboutir à une diphtongue d'un autre type (**jau* > *iwö*). En (5), la généralité de la voyelle basse étirée milite en sa faveur, contre l'innovation que constitue le rehaussement labialisé à San Lorenzo ('*a* > *o*). Il en va de même pour la voyelle moyenne palatale de **te* face à *ta* dans cette variété, en fonction du premier principe posé plus haut, qui écarte San Lorenzo des candidats fiables pour reconstruire les timbres vocaliques. Quant à la reconstruction des étymons 7 et 8 (**tyisñaán* et **tsú'ú*), elle suit également les principes évoqués précédemment. Cet exercice permet aussi bien de relativiser les aires de Gudschinsky sans pour autant les invalider, en montrant les phénomènes de tuilage et le caractère particulièrement innovant de certaines variétés sur des points précis du système phonologique, que d'appliquer les procédés élémentaires de la reconstruction : 1) identification des tendances générales, 2) plausibilité des changements phonétiques (plutôt **au* > *o*, *wö* que **o* > *au*, *wö*), 3) caractérisation marginale des variétés idiosyncrasiques comme Jalapa et San Lorenzo.

Il va de soi que le maniement de ces procédés ne s'apprend pas en deux jours, mais le fait de conscientiser le public à cette façon d'analyser des données dialectales en partant d'un terme commun eut pour effet salutaire de susciter une réflexion du groupe sur les fondements logiques de toute argumentation sur ce qu'on peut considérer comme une variété « ancestrale » ou « conservatrice » ou au contraire une variété « innovante », ainsi que sur le relativisme des frontières dialectales, dans un esprit de tolérance et de pluralisme face à la variation dialectale. Les participants firent observer que souvent les discussions entre instituteurs et promoteurs de l'éducation bilingue achoppaient sur cette pierre de discord qu'est le purisme local, ou le localisme campaniliste, chacun soutenant que sa variété était représentative du « vrai mazatec ». Le stage prit alors toute sa dimension d'*éducation populaire*, et aurait sans doute plu à un Antonio Gramsci, qui fut élève de Matteo Bartoli en son temps. Sans doute cette sincère distanciation du groupe a quelque chose à voir avec la dimension réflexive de la culture mazatèque, signalée par un anthropologue comme Boege Eckart (Boege 1988)²⁸. Lorsque le dialectologue réalise ce genre d'activité, où il consulte une assemblée de locuteurs praticiens de la langue sous des formes écrites, il rencontre souvent des résistances et des stéréotypes déclarés, qui semblent s'amonceler en une barrière insurmontable. Je ne nie pas cette réalité, et je me garderais bien d'idéaliser les Mazatèques ou le public de cet atelier – car des tendances au campanilisme furent observables à certains moments, sur d'autres questions –, mais le seul fait que cette réflexion distanciée ait pu émerger d'un travail purement technique à partir d'une liste réduite de données auprès d'un collectif d'instituteurs et d'animateurs culturels mazatèques n'en est pas moins encourageant.

Une fois expliqué comment le linguiste reconstruit les étymons à partir de la comparaison des formes des dialectes modernes, et de là, projette les chaînes évolutives à partir de cet « état moyen » que représentent ces mêmes étymons, nous avons abordé le premier exercice d'application, qui consistait à compléter une liste de correspondances établie par Paul Kirk en notant les réflexes d'une ou plusieurs variétés dialectales des Hautes Terres en usage dans le public, et d'identifier les lois phonétiques à partir des étymons de Kirk – ou les processus phonologiques opérant en diachronie, comme nous l'avons fait précédemment dans l'analyse du tableau 4 supra.

²⁸ Cf. aussi Demanget 2006 : 26 et 279.

3.2.3. Résultats du premier exercice d'application

Cette activité se déroule en trois phases : 1) chaque groupe de deux ou trois participants complète avec les données de sa propre variété dialectale les colonnes vides de la liste prévues à cet effet sur une fiche préparée à l'avance par le formateur, et qualifie dans une autre colonne vide les changements phonétiques observables en se fondant sur l'étymon de Kirk ; 2) les participants de chaque groupe recopient ces données et ces diagnostics phonologiques sur des posters ou papiers grands formats (appelés *papelógrafos* ou *láminas*, accessibles dans toutes les papeteries²⁹) ; 3) L'un des participants de chaque groupe, ou chacun à tour de rôle, présente sous forme de microconférence le produit de son travail de compilation des données de sa variété dialectale et son analyse des changements phonétiques. Ces exposés oraux durent de 10 à 15 minutes par groupe, et furent entièrement enregistrés sur digitaliseur de son ZOOM H2, filmés à l'aide d'un caméscope numérique SAMSUNG MMC+ et photographiés avec un appareil numérique PRAKTICA Luxmedia 7403. Une copie de l'ensemble des fichiers son, vidéo et image est laissée ensuite à l'organisation commanditaire du stage (en l'occurrence, la Casa de Cultura Maria Sabina), qui reste également en possession de tous les posters produits par les participants. Il y a donc restitution immédiate aux intéressés de leur production, dans son intégralité. Les clichés 1a-c permettent de mieux comprendre comment se déroule cette activité, et de visualiser quelques posters : l'image 1a, à gauche de la série, montre un poster réalisé par un groupe de deux locuteurs – l'un de Huautla de Jiménez, l'autre de San Miguel Huautepéc, bourg proche de Huautla – ; le cliché 1b en haut à droite a été pris sur le vif durant la microconférence d'une participante de ce même groupe, qui compare les deux variétés ; le cliché 1c en bas à droite est un plan rapproché d'un autre poster, qui compare les données de Huautla de Jiménez et de Mazatlán de Flores.

Nous pouvons maintenant analyser les données contenues dans ces posters, reportées sur une microbase de données, qui alimente les deux tableaux ci-dessous³⁰ : 7.1. et 7.2, qui déclinent une trentaine d'items analysés à partir des étymons de Paul Kirk (référéncés colonne A, listés dans la colonne C), avec en regard les données d'au moins une variante des Basses Terres, celle de San Miguel Soyaltepec (colonne D, données recueillies par mes soins auprès d'Apolonio Bartolo Ronquillo en juillet 2010 à Nuevo Pescadito de Abajo Segundo, également par double élicitation, écrite et orale). Les colonnes suivantes contiennent les données des participants de l'atelier protomazatec de Huautla : locuteurs de San Miguel Huautepéc (colonne E), de Mazatlán de Flores (colonnes F et G), et de Huautla de Jiménez (colonnes H à K). Ces faits de langue sont tirées de la saisie des données présentées par les participants dans leurs posters servant de support à leurs microconférences, comme on peut en voir dans les séries de clichés 1a-c, 2a-c infra. Ce sont donc des données élicitées par écrit, bien que chaque forme ait été lue ou énoncée durant la microconférence par l'intervenant,

²⁹ Il est bien plus difficile de se fournir en papiers grand format de ce type en Europe sans engager des frais en papier dessin grand format. En Amérique centrale, le format poster papier Bond est un outil de travail quotidien pour les enseignants des écoles et les formateurs. Il y a une culture de la présentation de données sur posters faits à la main que nous ne possédons malheureusement pas en Europe. Cependant, la tendance au Mexique est désormais à utiliser les diaporamas Power Point, et plusieurs participants proposèrent de saisir directement les données sur tableur pour les projeter lors de la microconférence. A des fins pédagogiques et afin de laisser une trace papier de l'ensemble des travaux, j'ai cependant insisté pour que les présentations se fassent sur support « papélographique », c'est-à-dire sur posters grand format.

³⁰ Des bases de données bien plus fournies pourraient à l'avenir être constituées au cours d'ateliers de ce type dans d'autres localités. Cette possibilité a été envisagée avec des responsables de zones scolaire participant à ce stage. Bien entendu, ces données ne remplacent en aucune façon le travail de notation du linguiste, pour la publication d'un atlas, mais rien n'empêche aujourd'hui, grâce à l'informatique, d'enrichir la version en ligne de l'atlas linguistique mazatec en cours de conception d'une base de données d'attestations graphiques par des instituteurs et promoteurs culturels de la région. Il va de soi qu'un travail minutieux d'identification sociolinguistique des transcrits improvisés que seraient alors les participants de ces stages devrait être prévue en amont, ce que nous n'avons pas eu l'occasion de faire lors de cette étape expérimentale.

laissant ainsi une trace orale enregistrée sur dictaphone numérique ZOOM, nettement visible sur les clichés 2a-c. Par souci de concision, je n’analyserai dans ce qui suit que les formes écrites, telles qu’elles apparaissent dans les posters. Les formes orales restent utilisables pour des analyses ultérieures, mais elles ne diffèrent guère, dans l’ensemble, des formes que je vais commenter.

TE TAO	HUATEPEC (G.M.H.)	SIGNIFICADO
kui	kui	este, esto
chō	chō	animal
natsi	natsi	metate
niba	nitsja niu	comal
Xia	tjia	brazo
Kjin	kjua biya	lejos
hoo'biya	tsju	la muerte, el t'allá
tsjo	kichji	tostado
Kichji	nachjón	maduro, en su punto
nachjón	nandq	hilo
nandq	nandq	nopal
ndia	ndia	camino
ni'yo	ni'yo	hormiga
ndo	ndo	largo
lajoo	lojo	pedra
li	li	fuego
ti'ndis	ti'basenju ti-basenju	esta escuchando
chimo	tsindio	lindero - límites
chiaga	chiaga	Mañana, después, adelante
ngur'ba	ngi'ba	Mañana

TE TAO	V. HUATEPEC (G.M.H.)	SIGNIFICADO
Yaxo'kuale	ya-xu'kua'le, kuichon	Le van a pegar con un palo
ts'oba	nds'oba	boca
ti'ndso	ti'ndso	chivo
ncha	nchya	atole
nchotin	nchutin	elote
nchee	nyee	nitomol
tjao	njjo	viento
yo'ba	ya xj'je	gancho
tsokón	tsu'kun	Cañuela
ntsja	tsja	miedo
ts'a	ts'a	pelo
nds'a	tsja ts'a	tejón
tsja	tsja	amargo
hichja	ts'a	esta hablando
chya'hi	ts'ia	amargo
nga	ts'ia	amargo
xinde	ts'ia	amargo
nga	ts'ia	amargo
nchi	ts'ia	amargo
ts'a	ts'ia	amargo



ni'yo	ni'yo	ni'yo
ntoju	ndoo	ndoiú
ntygiú	lao	laju
nti'i	li'i	li'i
ntyéé	ti'rbé	ti'ndié
nkwajá	yaxókuale	yaxú'kualara
téntsú	ti'ndso	téntsú
nts'wa	ts'obaa	ts'uba
nchajá	nyia	nchajá
nt'iaé	nt'iaé	nt'iaé
ntyji	ya'ba	ya'ndebá
yatjai	yax'ia	yax'ie
xanti	chaki	nachajá
intya	ni'ga	ndia
jnka	xn'ga	langa
xinte	xn'ga	xindie
jaskan	skan	askan

Clichés 1a-c. Premier exercice de l’atelier de protomazatec, Huautla août 2010 : élicitation sur posters de données de la Alta et analyse des changements phonétiques.



Clichés 2a-c. Atelier de protomazatec, microconférence de Javier Garcia Martinez (SEP), Huautla août 2010

La présentation des tableaux 7.1 et 7.2 permet de référencer comme sur un jeu d'échec ou à la bataille navale chaque item élicité par écrit : 1-A du tableau 7.1 se réfère ainsi au numéro de série de correspondances 107 de la monographie de Kirk (1966), 1-C nous donne l'étymon *chu, traduit en 1-B comme « animal », 1-F une forme de Mazatlán *chu*, 1-H une forme *cho* de Huautla, concurrente de la forme 1-K *chó*, notée par un autre participant, dans la même localité. Autre exemple : 9-C traduit en 9-B comme « pierre », vaut pour *ntyaju, 9-E pour *lojo*, 9-K pour *lajao*, etc.

Les données produites avec ce genre d'exercice deviennent rapidement pléthoriques. Afin d'éviter l'atomisme dans l'analyse de ces résultats, je mentionnerai les grandes tendances, en indiquant les coordonnées tabulaires comme je viens de le suggérer (1-H contr 1-K, etc.). Je comparerai chaque fois qu'il sera nécessaire les données élicitées dans le cadre de l'atelier avec celles publiées par Kirk dans sa monographie sur la phonologie du protomazatec (*op. cit.*).

Suprasegmental : un ton bas en 1-C est confondu avec un ton haut en 1-K ou est noté par défaut comme ton moyen partout. J'appellerai ce phénomène, récurrent chez les praticiens de l'écrit de la langue, *l'inversion ou la neutralisation de palier tonal*. Là où Paul Kirk notait dans les années 1960 un ton bas à Huautla dans *chu > cho (*op. cit.* p. 57), le classificateur pour « animal », on voit des formes à ton moyen *chu* (en 1-D, G) et *cho* (1 H-J) ou, par erreur d'identification, notées avec ton haut comme en 1-K *chó*. On pourrait penser que cette tendance est plus forte pour les monosyllabes comme *chu « animal » en 1 et *ncha « cheveux » en 20 que pour les dissyllabes. Or, si l'identification du ton bas paraît en effet plus assurée en 5 avec *nanta, « nopal », en 9 avec *ntaju « pierre » et en 12 avec *chinka « cochon », on retrouve cependant des apories graphémiques analogues en 17 avec *nchutin « épis de maïs » : là où Kirk note deux tons bas dans son enquête à Huautla (*op. cit.*, p. 72), on trouve un ton haut graphié en 17-K dans notre corpus.

Une autre tendance des graphies, qui tend également à un réductionnisme de la notation, voire de ce qui est retenu de la perception pour la notation, est l'absence de notation des tons glissants ou tons à contours, ou bien une notation fluctuante par redoublement ou fausse diphtongaison de la voyelle. Le contour tonal BH en 3 dans *kjin « loin » n'est noté nulle part dans notre corpus, sauf en 3-D, où il est identifié comme ton haut. En 18 *nchě « nixtamal, pâte de maïs détrempé » est noté comme un ton bas à San Miguel Soyaltepec, comme double voyelle en 18-E, H, K et comme pseudodiphtongue en 18-G. Il est donc potentiellement distingué du ton haut, mais ne reçoit pas de caractérisation graphémique univoque. Enfin, le ton haut, indépendamment du nombre de syllabes, est soit correctement identifié, comme en 6 pour la série de *ntiyá « chemin », en 10 pour *nti'í « feu » et en 16 pour *nchajá « atole, gruau de maïs » – pour ce dernier item, Kirk (p. 71) notait à Huautla un ton glissant ascendant M+-H, non décelable ici. En conclusion, on voit que les tons cardinaux (haut et bas) sont davantage notés que les tons à contour, et que le plus robuste des deux est le ton haut, selon une hiérarchie de pertinence *H* >> *B* >> *Contour*. Cette généralité revêt une forte valeur prédictive pour diagnostiquer les fluctuations des compétences graphiques chez les locuteurs écrivant le mazatec.

	N° Kirk	Trad.	Etym.	SanM Soyal.	HTep	Maz.	Maz.	Huautla de Jiménez			
	A	B	C	D	E	F	G	H	I	J	K
1	107	animal	* <u>chu</u>	<i>chu</i>	<i>chu</i>	<i>chu</i>	<i>chu</i>	<i>cho</i>	<i>cho</i>	<i>cho</i>	<i>chó</i>
2	610	bras, épaule	* <u>tyja</u>	<i>tjiá</i>	<i>tjia</i>	<i>tjia</i>	<i>tjia</i>	<i>xja</i>	<i>xjaa</i>	<i>z'a</i>	<i>xja</i>
3	202	loin	* <u>kjin</u>	<i>kjin</i>	<i>kjin</i>		<i>kjin</i>	<i>kjin</i>		<i>kjin</i>	<i>kjin</i>

4	23	grillé	* tsju	<i>tsjú</i>	<i>tsju</i>	<i>tsju</i>	<i>tsju</i>	<i>tsjo</i>	<i>tsjo</i>		<i>tsó</i>
5	271	nopal	nanta	<i>nanda</i>	<i>nanda</i>		<i>nanda</i>	<i>nanda</i>		<i>nanda</i>	<i>nanda</i>
6	392	chemin	* ntiyá	<i>ndiya</i>	<i>ndiá</i>	<i>ndiyá</i>	<i>ndiyá</i>	<i>ndiá</i>	<i>ndiaa</i>	<i>ndiaa</i>	<i>ndia</i>
7	394	fourmi	* nti'yu	<i>ni'yu</i>	<i>niy'u</i>	<i>ninyu</i>	<i>niyu</i>	<i>ni'yo</i>	<i>ni'yo</i>	<i>ni'yo</i>	<i>n'i yo</i>
8	397	large	* ntuju	<i>ndujú</i>	<i>ndu</i>	<i>ndujú</i>	<i>ndujú</i>	<i>ndo</i>	<i>ndoo</i>	<i>ndo</i>	<i>ndo</i>
9	406	Pierre	* ntyaju	<i>ndiajo</i>	<i>lojo</i>	<i>laju</i>	<i>laju</i>	<i>lajao</i>	<i>lao</i>	<i>lajao</i>	<i>lajao</i>
10	411	feu	* nti'í	<i>ndi'í</i>	<i>l'i</i>	<i>lí</i>	<i>li</i>	<i>li'í</i>	<i>l'i</i>	<i>lí</i>	<i>lí</i>
11	416	limite de terrain	* ntyú, * ntyú	<i>ndiya</i> <i>tsindiú</i>	<i>tsindiú</i>			<i>chinró</i>		<i>chi</i> <i>nró</i>	<i>chinyo</i> <i>o</i>
12	100	cochon	* chinka	<i>chinga</i>	<i>chinga</i>		<i>chinga</i>	<i>chinga</i>		<i>chinga</i>	<i>chinga</i>

Tableau 7.1. Résultats de l'élicitation par écrit sur posters à partir d'une liste d'étymons de Paul Kirk (1966)

Segmental :

Consonantisme : je m'intéresserai en premier lieu à l'affriquée rétroflexe, qui ressortit au modèle (a) des aires de Gudschinsky, avant d'aborder un autre ensemble de variables que nous n'avons fait jusqu'à maintenant qu'évoquer rapidement lors de la présentation des données du tableau 1 supra, mais qui joue pourtant un rôle notable dans la diversité phonologique du mazatec : la classe des prénasalisées. L'affriquée rétroflexe est notée en 2-J, mais avec une expression craquée (*creaky*), avec apostrophe ou « saltillo », au lieu de soufflée (*breathy*) : graphie <z'a> contre une notation équivalente à <zja> chez Kirk (op. cit., p. 58). Ailleurs à Huautla, les intervenants ont noté <xja(a)> en 2-H, I, K, conformément à Kirk, à un détail typographique près (<x> pour <z>, ce qui neutralise la notation de la rétroflexion en graphie). En 11, là où Kirk notait l'équivalent de <chinzu> à Huautla (p. 68), on a <chinró> en 11-H, J et <chinyoo> en 11-K. Or on attendrait en bonne logique chez cet intervenant la notation <chinzó> au lieu de <chinró> en 11-J, puisque le réflexe pour *tyja était noté <z'a> en 2-J, comme nous venons de le voir – quoiqu'une logique de distribution complémentaire entre attaque initiale et attaque médiane serait envisageable, mais le corpus n'est pas suffisant ici pour en juger. La notation <chinyoo> neutralise sinon l'affrication, du moins la rétroflexion dans la graphie, et peut soit s'interpréter comme une solution standardisante, soit comme un polymorphisme absent des données de Kirk.

	N° Kirk	Traduction	Étymologie	SanMSoyal.	HTepec	Mazatlán		Huautla de			
						F	G	H	I	J	K
	A	B	C	D	E	F	G	H	I	J	K
13	349	menton	* nki'wa	<i>ngui'ba</i>	<i>ngi'ba</i>		<i>ngui'ba</i>	<i>ngui'ba</i>		<i>ngui'ba</i>	<i>ngui'ba</i>
14	294	bouche	* ntsu'wa	<i>ts'ua</i>	<i>ndsub'a</i>	<i>tsu'ba</i>	<i>tsa</i>	<i>ts'obá</i>	<i>tso'ba</i>		<i>ndo'ba</i>
15	568	chèvre	* téntsú	<i>tendsu</i>	<i>tindsu</i>	<i>tendsú</i>	<i>tendsú</i>	<i>tindso</i>	<i>tindsó</i>		<i>tindso</i>
16	297	atole	* nchajá	<i>nchaja</i>	<i>nchyá</i>	<i>nchajá</i>	<i>nyaja</i>	<i>ncha</i>	<i>nyia</i>		<i>nya</i>
17	308	épis de maïs	* nchutín	<i>nchutín</i>	<i>nchutin</i>		<i>nyutín</i>	<i>nchotín</i>			<i>nyotín</i>
18	302	nixtamal	* nchě	<i>nche</i>	<i>nyee</i>		<i>nyea</i>	<i>nchee</i>			<i>nyee</i>
19	369	vent	ntjau	<i>tjao</i>	<i>ntjo</i>	<i>tjə</i>	<i>ntjo</i>	<i>tjao</i>	<i>ntjao</i>		<i>ntjao</i>
20	288	cheveux	* ncha	<i>tsja</i>	<i>tsja</i>		<i>ntsja</i>	<i>ntsja</i>		<i>nsja</i>	<i>ntsja</i>
21	303	Il/elle parle	* nchja	<i>tjichja</i>	<i>tíchja</i>		<i>tinchja</i>	<i>tjichja</i>		<i>jinchja</i>	<i>tinchja</i>

22	149	salé	* jntya	<i>ndia</i>	<i>ndia</i>	<i>ndiá</i>	<i>ndia</i>	<i>nra</i>	<i>nrja</i>	<i>nrjaa,</i> <i>njra</i>	<i>nra</i>
23	143	aile	* jnká	<i>tjiungá</i>	<i>ngà</i>	<i>langa</i>	<i>langa</i>	<i>nga</i>	<i>ng'a</i>	<i>ngaa,</i> <i>njga</i>	<i>nga</i>
24	523	adobe	* xi'nte (adobe)	<i>xinde</i>	<i>xind'e</i>	<i>xindie</i>	<i>xinde</i>	<i>xinde</i>	<i>xi'nde</i>	<i>rinde,</i> <i>xinde</i>	<i>xinde</i>
25	709	haut	* nka	<i>k'á</i>	<i>nk'a</i>		<i>nga</i>	<i>jnga</i>		<i>nga</i>	<i>n'ga</i>
26	705	mouillé	* nchi	<i>nchi</i>	<i>ny'i</i>		<i>nyi</i>	<i>nchji</i>		<i>nchyi,</i> <i>nyi</i>	<i>nyi</i>
27	485	poisson	* sti	<i>ti</i>	<i>tii</i>		<i>ti</i>	<i>jti</i>		<i>djti, jti</i>	<i>jti</i>

Tableau 7.2. Résultats de l'élicitation par écrit sur posters à partir d'une liste d'étymons de Paul Kirk (1966)

Prénasalisées : en 7, 'nt dans **nti'yu* « fourmi » est partout simplifié en n ('nt > n / #_i), conformément aux données de Kirk (p. 63), où *'nt ne se conserve qu'à Chiquihuitlán et à Santo Domingo del Río, dans les Basses Terres. En revanche, en 8 **ntuju* « large » maintient l'attaque prénasalisée (*nt > nt ; nd / #_u), aussi bien dans les relevés de Kirk (p. 63) que dans notre corpus. Il n'en va pas de même en 9 et en 10 avec des séries phonolexicales (autrement dit, irrégulières, conditionnées lexicalement) pour lesquelles la prénasalisée coronale alterne avec une simple sonante nasale n et une latérale l (Kirk, op. cit., p. 67) : *nt > nt, nd, n, l. Le phénomène reste mal ou peu expliqué, surtout dans sa dimension phonolexical (une liste de mots restreinte, qui fait figure d'irrégularité et de niche de diffusion lexicale). Deux chaînes évolutives sont possibles³¹ : la première, **ntaju* « pierre », **nti'í* « feu » > **ndaju*, **ndi'í* > **nlaju*, **nli'í* > *laju*, *li'í* suppose une assimilation partielle de sonantinité entre la composante sonantique nasale et la composante occlusive coronale, mais reste théorique, la deuxième suppose une séquence évolutive **ndaju*, **ndi'í* > **daju*, **di'í* > *laju*, *li'í*, par dénasalisation, puis latéralisation de la coronale, par jeu de sous-spécification. Dans les deux cas, on retiendrait que la composante nasale peut s'amuir, suivant une logique de simplification de groupe consonantique compatible avec un modèle monosegmental. Cependant, la séquence **ndaju*, **ndi'í* > **nlaju*, **nli'í* > *laju*, *li'í* implique plus en amont dans l'analyse un modèle d'interaction de traits (cf. tableau 4 supra). Elle est également davantage convergente avec l'hypothèse de Marlett (1992) qui stipule que ce qu'on considère généralement dans les langues otomangues comme des occlusives prénasalisées sont en réalité des sonantes nasales postoralisées, puisque la sonantinité est le trait dominant dans un tel scénario évolutif : *nd > *nl > *(n)l > l. La notation *nt de Kirk, qui me semble en réalité très abstraite car attestée dans peu, voire aucune variété de mazatec moderne où c'est nd, n, d ou l qu'on entend, relève d'autant plus, a posteriori, du modèle monosegmental (n + t au lieu de ⁿd ou mieux, de n^d). Cependant, Kirk montre bien que *nt a pu donner n ou t et d dans le réseau dialectal mazatec (op. cit. p. 64-79). Les conséquences typologiques de ce changement ne sont pas du tout les mêmes selon qu'on fonde l'explication sur un modèle de traits complexes ou sur un modèle monosegmental. En 9 H-K on retrouve le même polymorphisme *lao* et *lajao* « pierre » que dans les relevés de Kirk, qui note l'équivalent de <*lau*, *lajau*>. C'est en 10 H-K que la variation entre trois formes contraste avec l'univoque notation de type < *li'í* > de Kirk (p. 67), qui converge avec 10-I dans notre corpus. Cette variation est sans doute autant due à la difficulté technique de graphier un monosyllabe doté d'une voyelle craquée qu'à une réelle allophonie de ce qu'on peut tenir pour une forme lexicale de type /li □/, d'où les notations *li'í*, *lií* « feu ». On voit là un exemple des apories ou des vacillations de

³¹ Je remercie Mamadou Keita, Paris VII, phonologue et spécialiste de linguistique niger-congo pour ses commentaires des faits otomangues, qui m'ont permis de mieux comprendre ces phénomènes, avec le recul et la distance que permet la comparaison typologique avec d'autres langues à prénasalisation.

l'écriture liées à la complexité du système de traits du mazatec, parfois irréductible au modèle de notation monosegmental.

Vocalisme : les graphies discrétisent bien davantage l'alternance diasystémique entre *u* et *o* pour la voyelle haute postérieure que ne le font les notations de Paul Kirk. Là où l'isoglosse est nettement visible dans la série 1 du tableau 7.1. (**chu* classificateur pour « animal »), la notation de Kirk (p. 57) est bien plus phonémique : il note /u/ partout, sauf à Lo (San Lorenzo Cuaunecuiltitla, proche de San Jeronimo Tecoaatl dans la Alta), où le changement *u* > *i* (**chu* > *chi*) par délabialisation apparaît d'autant plus catégoriel que Kirk ne retient pas d'autre forme de variation dans le réseau dialectal mazatec. La codification actuellement en cours de diffusion dans le système scolaire et l'édition creuse donc les écarts de ce qui peut être retenu comme un trait emblématique de la différence entre la Mazateca Haute et Basse, davantage phonétique que phonologique. De ce point de vue, les notations du participant de San Miguel Huauteppec, dans la colonne E, sont d'autant plus intéressantes que, selon cette logique, on s'attendrait à ce que ce locuteur de la Haute Mazateca note une voyelle abaissée *o* au lieu de *u*. Or, il note *u*, comme Paul Kirk. Dans le cadre très précis qui est ici le nôtre, d'un questionnement méthodologique sur comment la consultation et l'élicitation par écrit de données dialectales auprès de locuteurs praticiens de la langue et de sa codification peuvent guider le linguiste vers des pistes de recherche, ce détail est particulièrement intéressant. On peut s'attendre à une voyelle haute -ATR (*Advanced Tongue Root*) dans la Alta, de type [u] pour /u/, plutôt qu'à une voyelle mi-haute [o].

L'apport de l'atelier de dialectologie réside donc dans l'empirisme critique, au-delà de l'évident intérêt que représente l'exploration de la catégorisation phonémique des segments dans une langue aussi complexe en traits que le mazatec. L'empirisme critique consiste à ne rien prendre pour acquis, dans l'observation et l'analyse des données, à partir du principe que les données linguistiques sont des *construits perceptuels* (modèle de compétence : ce qu'on entend et interprète en notant), *catégoriels* (phonologie : ce qu'on analyse comme relevant de classes de sons) et *ergonomiques* (codification : ce qu'on note en fonction de conventions et de ressources graphémiques ou typographiques). En particulier, les résultats de cette activité avec des maîtres d'écoles et des promoteurs culturels mazatecs, est d'autant plus intéressant que le niveau de formation, de réflexion et de débat des participants est élevé. La problématisation de l'enquête dialectologique prend une dimension à la fois (ethno)méthodologique (construire les savoirs d'un point de vue endogène) et réflexive (distanciation face aux pratiques et aux savoirs reçus dans une culture ou une société donnée).

3.2.4. Résultats du second exercice d'application

Traiter les matériaux issus de cette dernière activité requiert autant de place que pour l'analyse des données et l'argumentation de la section 3.2.3. Je résumerai donc les faits saillants, dans la visée de cette contribution, qui est de montrer l'apport des ateliers de dialectologie dans un projet d'exploration de diasystèmes de langues non indo-européennes, typologiquement très marquées ou d'un haut degré d'idiosyncrasie³².

Le deuxième exercice consistait à remplir la colonne des protoformes (équivalent de la colonne C du tableau 7), à partir d'une microbase de données tirées d'une compilation des variantes dialectales de la Baja (nourrie par les données de quatre variétés de la Baja réunies dans Regino 1993), et de compléter en remplissant des colonnes comme en E, F ou H dans le

³² J'évite ici à dessein de parler de haut degré de *complexité*, même si j'ai auparavant mentionné la complexité du système de traits du mazatec. Mais comme le montrait la confrontation des tableaux 2 et 3 ou des modèles de Pike & Pike 1947 *versus* Golston & Kehrein 1998, la complexité est une notion toute relative, qui dépend essentiellement du point de vue ou de l'angle d'approche. Cf. Léonard & Kihm 2010 pour une application de cette distanciation vis-à-vis des notions de simplicité et de complexité par le choix des modèles d'analyse à la morphologie flexionnelle (verbale) du mazatec

tableau 7 (données de la Alta, à partir des idiolectes des participants). Outre l'analyse qu'on peut faire des données élicitées par écrit de cette manière, comme précédemment, l'intérêt était de voir comment les locuteurs allaient appliquer ce qu'ils avaient appris de la méthode, selon les principes énoncés en 3.2.2. (*Éléments de méthodologie de la reconstruction du protomazatec*). C'est là que nous eûmes des surprises. Les participants qui ont le mieux réussi cette activité n'étaient pas nécessairement des linguistes. Au contraire, le groupe le plus réaliste comprenait trois promoteurs culturels, qui disposaient de très peu d'information sur les principes de la linguistique et de l'analyse formelle de leur langue. Les ethnolinguistes ne furent pas toujours les plus exacts dans cette tâche, bien au contraire. En particulier, un groupe d'experts dont on pouvait attendre un maniement correct de la méthode ne fit qu'entériner le campanilisme, et se contenta pratiquement de recopier dans la colonne des étymons les variantes de sa propre variété locale, celle de Huautla. Le groupe des promoteurs culturels, en revanche, applica méticuleusement les trois principes que sont 1) le degré de généralité des formes, 2) la mise à l'écart des idiosyncrasies, 3) la simplicité des segments et la présupposition qu'un son complexe dérive d'un son simple ou de la fusion de deux sons simples (par ex. l'affrication d'une occlusive coronale suivie de glide palatal comme *ty). Le troisième principe avait été énoncé comme : « il est facile de transformer un bocal de poissons rouge en soupe de poisson, mais l'inverse est impossible ». Le travail minutieux a donc gagné sur la projection idéologique localiste, et la confrontation des deux points de vue, émanant de compétences très inégales au départ, s'est avérée riche en enseignements. Autrement dit, le point de vue ethnométhodologique a primé, en efficacité, sur le point de vue structuraliste, dans l'application de principes universels permettant de produire des étymons plausibles en fonction de ce que l'on connaît de la façon dont se produisent les changements phonétiques dans les langues du monde et dans les réseaux dialectaux.

Conclusion

C'est sur ces considérations paradoxales que je terminerai ce compte rendu méthodologique sur l'atelier de protomazatec à Huautla. Certaines langues ont eu, par la richesse et l'originalité de leurs structures et de leurs propriétés, une influence qui dépasse de loin les frontières de leur usage, de leur prestige local et de ce que peuvent imaginer leurs locuteurs, s'ils ne sont pas initiés à la linguistique. Une langue comme le mazatec a eu un impact considérable sur l'évolution de la linguistique de la seconde moitié du XXe siècle, suite à l'article de Pike & Pike 1947, qui fit l'effet d'une bombe dans le champ de la phonologie moderne, puisqu'il proposait, à partir de l'exercice de haute voltige que représentait l'analyse des clusters consonantiques à expression glottique et/ou prénasalisée du mazatec, un modèle de constituance des marges initiale et finale de syllabe que sont l'attaque et la coda, et du noyau ou centre de la syllabe qu'est la voyelle. Golston & Kehrein tentèrent de démontrer près d'un demi-siècle plus tard que, si l'intuition théorique des frères Pike (Eunice et Kenneth) était juste, c'est le détail de l'application du modèle aux objets d'analyse qui s'avérait sans doute erroné. Plutôt que des attaques branchantes cumulant de nombreux segments, ou des segments simples avec des combinatoires inattendues (*hnt, hnts, fh, sʔ, ?m, ?n*), ne pouvait-on analyser ces segments comme des catégories de consonnes et de voyelles dotées de traits secondaires mobiles ? Eunice et Kenneth avaient-ils bien distingué dans ces traits ce qui appartenait aux attaques et ce qui appartenait aux voyelles ? N'avaient-ils pas pris la richesse ou la complexité des formes de surface pour argent comptant ? Pourtant, en tant qu'évangélistes ayant pour objectif parallèle de traduire la Bible dans cette langue, ils avaient dû tenir compte des conséquences ergonomiques de ce modèle pour l'écriture. Il se trouve que la nature monosegmentale de cette approche permettait, au prix de légères incohérences en fonction de l'allophonie des formes de surface, de construire un système graphique. Ce modèle s'est même avéré suffisamment ergonomique pour que les codificateurs natifs à vocation

laïque qui ont repris le flambeau à la fin du XXe siècle, comme Juan Gregorio Rufino, puissent s'en accommoder, parvenant même, en radicalisant ce modèle, à codifier une langue riche en traits secondaires en une langue somme toute aussi facile à écrire que n'importe quelle autre. Mais il n'en reste pas moins que cette langue peut tout aussi bien s'analyser selon les critères de la typologie moderne en phonologie³³ comme une langue à segments dotés de traits secondaires *creaky* et *breathy*, de même que les occlusives et les affriquées prénasalisées ne sont probablement pas l'équivalent des occlusives sonores des langues indo-européennes, mais plutôt des sonantes nasales à expression post-oralisée (Marlett 1992 et Léonard 2010a). On peut très bien recueillir des données fiables et les analyser en fonction du premier modèle, comme le firent Guschinsky (1957, 1958) et Kirk (1966), et s'en contenter. Tout comme on peut tenter de revisiter ces données, cette modélisation, afin d'en recueillir de nouvelles, et de les éclairer sous un nouveau jour, ouvert sur les principes de la typologie phonologique moderne. On peut faire appel aux locuteurs de manière davantage participative que dans les enquêtes dialectologiques classiques. On se rendra alors compte combien le choix du double éclairage est alors avantageux, stimulant, riche en perspectives, tant pour la conception d'une forme d'atlas linguistique qui prend en compte des questions de typologie linguistique, que pour enrichir la pratique de la collecte et de l'analyse des données d'un empirisme critique actif, qui ne se contente pas de seulement référencer les sources parallèles en marge des cartes. C'est la voie que ce collectif de travail a tenté de tracer lors du déroulement de l'atelier de dialectologie mazatec, qui sera suivi de bien d'autres, dans le cadre de la réalisation de l'ALMaz (*Atlas Lingüístico Mazateco*). Non seulement les Mazatecs ont su tirer intelligemment parti des travaux de chercheurs étrangers sur leur langue, mais eux et leur langue n'ont pas fini de poser à la linguistique moderne des questions fondamentales. Preuve de plus, si besoin était, que la linguistique est autant au service des locuteurs que des langues elles-mêmes en tant qu'objets, autrement dit, au service de sujets pensants et agissants, en dépit du caractère arbitraire et aliénant de la diglossie.

Références

- Benítez Fernando, *Los indios de México*, vol. 3, México, ERA, [1970]-1993.
- Benton Joseph, 1999. « How the Summer Institute of Linguistics has developed orthographies for indigenous languages of Mexico », accessible sur <<http://www.sil.org/americas/mexico/ilv/L001i-SILOrth.pdf>>.
- Bradley Henry & Josserand Kathryn, « El protomixteco y sus descendientes », México D.F., *Anales de Antropología*, n°19, 1982, p. 279-343.
- Boege Eckart, *Los mazatecos ante la nación. Contradicciones de la identidad étnica en el México actual*, México, Fondo de Cultura Económica, 1988.
- Bull Brian « Aspect Formation of San Jeronimo Mazatec Verb », SIL-Mexico Workpapers, n° 6, 1984, p. 93-117.
- Cazes, Daniel (s.d.) « Glotocronología hña-maclasinca-meco (Otopame) », manuscrit, CNRS.
- Clements, Nick G. & Hume, Elisabeth. « The internal organization of speech sounds », in Goldsmith John (ed.), *The Handbook of Phonological Theory*, Oxford/Cambridge, Blackwell, 1996, p. 245-306.
- Demagnet Magali, *La chair des Dieux est-elle à vendre ? Chamanisme, tourisme et ethnicité en terre mazatèque (Mexique)*, Thèse NR, Université Paris X, 2006.
- Fernández de Miranda Maria Teresa (ed. Piper M. & Bartholomew). *El protozapoteco*, México, El Colegio de México, [1966] 1995.

³³ V. Ladefoged & Maddieson 1996 pour une introduction à la typologie phonologique des langues du monde.

- Duke, Michael « Writing Mazateco: Linguistic Standardization and Social Power », accessible en ligne sur http://www.utexas.edu/courses/stross/ant392n_files/dukem.htm, (sans date).
- Golston Chris & Kehrein Wolfgang, « Mazatec onsets and nuclei », *International Journal of American Linguistics*, n° 64.4, 1998, p. 311-337.
- Golston Chris & Kehrein Wolfgang, « A prosodic Theory of laryngeal contrasts », *Phonology* n° 21, 2004, p. 1-33. < <http://zimmer.csufresno.edu/~chrisg/>>
- Gudschinsky Sarah, « Mazatec dialect history », *Language*, n° 34, 1958, p. 469-481.
- Gudschinsky Sarah, *Proto-Popotecan. A Comparative Study of Popolocan and Mixtecan*, *IJAL*, n° 25-2, 1959.
- Haspelmath Martin, Dryer Matthew, Gil David & Comrie Bernard (eds.), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, accessible en ligne sur <http://wals.info/feature/>, [2005]-2008.
- IJAL* = *International Journal of American Linguistics*, Chicago.
- Jamieson Carole, « Conflated subsystems marking person and aspect in Chiquihuitlán Mazatec verb », *IJAL*, n° 48, 1982, p. 139-167.
- Jamieson Carole, *Diccionario mazateco de Chiquihuitlán*, Tucson, SIL, 1996.
- Jamieson Carole, *Gramática mazateca. Mazateco de Chiquihuitlán de Juárez*, México D.F, SIL, 1988.
- Josserand Kathryn, *Mixtec Dialect History*, Ph. D. Dissertation, New Orleans, Tulane University, 1983.
- Kaufman Terrence, « Early Otomanguean Homelands and Cultures: Some premature hypotheses », manuscrit, University of Pittsburg, 2006.
- Kirk Paul Livingston, *Proto-Mazatec Phonology*, Ph. D. dissertation, University of Washington, UMI, 1966.
- Kirk Paul Livingston, « Intelligibility Testing: The Mazatec Study », *IJAL*, n°36-3, p. 205-211.
- Ladefoged Peter & Maddieson Ian, *The sounds of the world's languages*, Blackwell, Oxford, 1996.
- Léonard Jean Léo, « Classes naturelles, traits, complexité et simplicité segmentale en tu'un savi », *Bulletin de la Société de Linguistique de Paris*, 2010a (sous presse).
- Léonard Jean Léo, « Classes naturelles, traits, complexité et simplicité segmentale en tu'un savi et en mazatèque (otomangue orientale) », Colloque Réseau Français de Phonologie, juillet 2010, *in memoriam* Nick Clements, 2010b.
- Léonard Jean Léo & kihm, Alain, « Verb inflection in Chiquihuitlán Mazatec: a fragment and a PFM approach », handout, HPSG Conference, Paris, accessible sur <<http://csli-publications.stanford.edu/HPSG/2010/leonard-kih.pdf>>, 2010.
- Léonard Jean Léo & dell'Aquila Vittorio « Haudricourt & Juilland 1949 revisité. Perspectives géolinguistiques et post-structuralistes », *CILPhR València*, 2010 (à paraître).
- Longacre Robert, *Proto-Mixtecan*, Indiana University Research Center in Anthropology, Folklore and Linguistics, n°5, 1957.
- Longacre Robert, « Swadesh' Macro-Mixtecan Hypothesis », *IJAL*, n° 27.1, 1961, p. 9-21.
- Longacre Robert, « On Linguistic Affinities of Amuzgo », *IJAL*, n° 32-1, 1966, p. 46-49.
- Longacre Robert, « Systemic comparison and reconstruction », in McQuown, Norman (ed.) 1967. *Handbook of Middle American Indians*, vol. 5, Austin, University of Texas Press, 1967, p. 117-159.
- Mak Cornelia & Longacre Robert, « Proto-Mixtec Phonology », *IJAL*, n°26, 1960, p. 23-40.
- Marlett Stephen, « Nazalization in Mixtec languages », *IJAL*, n° 58.4, 1992, p. 425-435.

- McMahon David, *The Impact of a Federal Dam on a Mexican Indian Community in the Papaloapan Basin*, 1971, traduction de Carmen Viqueira, *Antropología de una presa: los mazatecos y el proyecto del Papaloapan*, México, INI, 1989.
- Pike Kenneth & Pike Eunice, "Immediate constituents of Mazatec Syllables", *IJAL*, n°13, 1947, p. 78-91.
- Polian Gilles & Léonard Jean-Léo, « La morphologie dans ALTO (Atlas Linguistique du Tseltal Occidental). Réseau dialectal et systèmes à décideurs multiples », Grenoble, *Géolinguistique*, n° 11, 2009, p. 149-201.
- Rendon Juan José, *Diversificación de las lenguas zapotecas*, Oaxaca, Ciesas, IOC, 1995.
- Regino Juan Gregorio, *Alfabeto mazateco*, Oaxaca, Ciesas, IOC, CDCNC, 1993.
- Rensch Calvin, *Comparative Otomanguean Phonology*, Ph.D. Dissertation, University of Pennsylvania, Indiana University Publications, Language Science Monographs, n° 14, Bloomington, Indiana University Press, [1966]-1976.
- Silverman Daniel, Blankenship Barbara, Kirk Paul & Ladefoged Peter, « Phonetic Structures in Jalapa Mazatec », *UCLA Working Papers*, 1987, p. 113-130, repris in *Anthropological Linguistics*, n° 37-1, p. 70-88.
- Silverman Daniel « Laryngeal complexity in Otomanguean vowels », *Phonology*, n° 14, 1997, p. 235-261.
- Soustelle Jacques, *La famille Otomí-Pame du Mexique central*, Paris, Institut d'Ethnologie, 1937.
- Weinreich Uriel, « Is a structural dialectology possible ? », *Word*, n°4, 1954, p. 388-400.