



**HAL**  
open science

## Le mazatec : un terrain-monde

Jean Léo Léonard

► **To cite this version:**

Jean Léo Léonard. Le mazatec : un terrain-monde. Leblic, Isabelle; Souag, Lameen. Du terrain à la théorie : les 40 ans du Lacito, Lacito, 2020, 978-2-490768-00-4. hal-04045960

**HAL Id: hal-04045960**

**<https://univ-montpellier3-paul-valery.hal.science/hal-04045960>**

Submitted on 25 Mar 2023

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

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

Jean Léo Léonard (Université Paris-Sorbonne, EA 4509, STIH)

## *Le mazatec : un terrain-monde*

### 1. Introduction

L'une des plus brillantes contributions de William Labov sur les mécanismes de diffusion du changement linguistique dans l'espace social commence comme dans la citation ci-dessous. Nous retrouvons là des observatoires qui nous sont familiers (des réseaux de centres urbains de différentes tailles, en Europe occidentale, avec des langues indo-européennes, germaniques, mais aussi des « objets » simples et faciles à manier, comme des variables telles que des diphtongues ou des phénomènes de palatalisation, avec des incidences systémiques faciles à observer, comme les chaînes de propulsion de de traction vocalique<sup>1</sup>, ou les vagues de palatalisations dans le lexique), et une méthode classique, qui oppose un centre innovant à une périphérie réceptrice, avec un rythme plus ou moins régulier de diffusion, conditionnée par le rayonnement des traits urbains prestigieux, qui déclassent progressivement les traits plus « rustiques » ou « provinciaux ». La métaphore de la cascade, avec sa dynamique scalaire régulière, l'eau s'écoulant des zones les plus « hautes », en amont, (ici, sur le plan symbolique, en termes de prestige), vers les régions situées plus en bas, en aval, est aussi crédible que rafraîchissante. L'alliance du modèle centrifuge (du centre vers la périphérie) et de la cascade ou de l'imbrication dynamique en poupée-gigogne constitue indéniablement une base solide pour l'observation d'un phénomène aussi complexe et délicat à analyser que le changement linguistique. En somme, même si l'on se situe en plaine, comme dans la région de Londres et ses villes alentour, le réseau de centres urbains plus ou moins « élevés » en prestige et plus ou moins discontinu ou diffus constitue un archipel vertical de sommets d'un réseau, dont le changement linguistique, pour sa diffusion, va emprunter des vecteurs (les arêtes d'un graphe, si on formalise sous forme de ponts et de traits un réseau de villes ou de dialectes/parlers urbains).

*“Granted that language is a social fact, and not the property of any individual, it follows that a linguistic change is equivalent to the diffusion of that change. An understanding of language change therefore demands an understanding of the mechanism of diffusion. It has long been observed that linguistic features spread outward from an originating center, but in a progressively weaker form as distance increases.*

*A major step towards the understanding of diffusion was made by Trudgill, in his studies of language change in the Brunlanes peninsula of Norway (1974). In Trudgill's gravity model of diffusion, change spreads from the largest to the next largest city, in a predictable order, the influence of one city on another being proportional to the relative sizes of the city and inversely proportional to the distance between them. The model showed a good fit to the variable (ae) involving the progressive lowering and backing of /ae/, and to the spread of London influence into Norfolk. Chambers and Trudgill 1980 develop the parallel palatalization of (sj) in Brunlanes with similar results” (Labov 2002)<sup>2</sup>.*

---

<sup>1</sup> Voir Fulcrand (2017) pour une synthèse théorique récente et abondamment documentée sur l'incidence systémique de cette catégorie de changement phonologique, et Labov (1994), qui consacre un chapitre entier aux « vowel shifts » dans la continuité des changements en chaîne, hérités de l'économie des changements phonétiques d'André Martinet, mais aussi de l'esprit des néogrammatariens.

<sup>2</sup> « Etant donné que la langue est un fait social, et non pas un bien individuel, il s'ensuit que tout changement linguistique équivaut à la diffusion de ce même changement. Toute analyse du changement linguistique requiert par conséquent l'analyse du mécanisme de diffusion qui le porte. De longue date on sait que les traits linguistiques se diffusent vers l'extérieur, hors d'un foyer initial, avec une inflexion progressive de la force des traits à mesure

On remarquera également l'usage de la notion de *gravitation*, dans le modèle de Trudgill cité par Labov, qui suit la même logique de poids, d'influence. Loin de moi l'idée de contester le caractère légitime et heuristique de ces concepts (et métaphores) de *centres directeurs*, *périphéries rétentrices*, *diffusion scalaire*, *gravitation*, *rayonnement*. Je pense qu'elles sont même incontournables, en tant que toile de fond, pour toute analyse des processus de diffusion du changement linguistique. En revanche, j'apporterai un bémol, pour souligner simplement que si ces notions sont des moyens, elles ne sauraient être des fins en soi. Elles demandent à être « falsifiées » dans le sens popperien du terme (autrement dit, mises à l'épreuve de faits nouveaux, afin de valider leur caractère *heuristique*, non *circulaire*)<sup>3</sup>. Elles demandent aussi à être raffinées, voire contournées et même... Détournées. C'est ce que nous allons faire à la lumière de faits.

Dans ce qui va suivre, je vais explorer, en hommage au Lacito et à son empirisme radical, alimentant une approche critique des théories et des modèles consensuels, sur la base de faits et d'observatoires nouveaux, voir insolites (dans le meilleur et le plus étymologique sens du terme, à savoir, inhabituel), un domaine dialectal non indo-européen, d'une grande complexité structurale, et en déployant un éventail inattendu d'observables. Je vais en effet fonder mon approche moins sur des variables que sur des paramètres et des sous-systèmes, ainsi que sur des pans entiers de cycles de changement linguistique (pour ne pas dire, de cycles évolutifs). Mon terrain sera no pas une région d'Europe, germanique, romane ou slave, mais le réseau dialectal d'une langue méso-américaine, de la « famille » otomangue, sous-branche orientale, à savoir, le mazatec.

Le mazatec, langue otomangue de la sous-famille popolocane, est parlé dans une région au carrefour de nombreuses langues et cultures méso-américaines, dans le bassin du Papaloapam, aux confins de la Chinantla. La région est aussi appelée Cañada<sup>4</sup> de Cuicatlán, et compte parmi les zones de plus haute biodiversité de Méso-amérique, voire d'Amérique. De 2010 à 2015, une équipe de chercheurs franco-italiens et mexicains a réalisé une série d'enquêtes dialectologiques afin de sonder le diasystème<sup>5</sup> mazatec, et de vérifier les données et les modèles existants pour les trois principaux dialectes connus (Huautla, Jalapa et Chiquihuitlán). Le résultat des enquêtes ainsi que de multiples approches de traitement des données (*Paradigm Function Morphology* [Stump 2001] pour rendre compte de la variation de la flexion verbale à échelle diasystémique, algorithme de Levenshtein [1966] pour le traitement dialectométrique des données de Paul L. Kirk, etc.)<sup>6</sup> ont fait apparaître l'un des réseaux dialectaux<sup>7</sup> les plus complexes et les plus diversifiés observables dans une région du monde pourtant connue pour son intense diversité linguistique. Sur bien des points, sont apparus des paradoxes empiriques d'un grand intérêt pour la recherche fondamentale mais aussi d'un point de vue interdisciplinaire. Dans chacun des quatre points qui va suivre, nous verrons le modèle de

---

que la distance augmente. Une avancée notable dans ce domaine a été réalisée par Trudgill, dans son étude sur le changement linguistique dans la péninsule de Brunlanes en Norvège (1974). Selon le modèle gravitationnel du processus de diffusion proposé par Trudgill, le changement procède de manière scalaire, de centres urbains majeurs en centres urbains mineurs, selon un ordre prévisible, l'influence de chaque centre sur ses voisins dépendant de la taille relative de chacun, de manière inversement proportionnelle à la distance qui les sépare. Ce modèle s'est avéré opérationnel pour la variable (ae), qui implique l'abaissement progressif et la vélarisation de /ae/, et pour expliquer le rayonnement de la variété londonienne sur une ville comme Norfolk. Chambers & Trudgill 1980 ont appliqué à la palatalisation de (si) dans la péninsule de Brunlanes le même traitement, avec succès ».

<sup>3</sup> Cf. Popper (1934).

<sup>4</sup> Ou « canyon ».

<sup>5</sup> Cf. Weinreich 1954 : le métasystème qui subsume l'ensemble des paramètres locaux de la variation dialectale d'une langue ou d'un domaine linguistique.

<sup>6</sup> Voir Heeringa 2004 pour le détail de la méthode de classification automatique en dialectologie à l'aide de cet algorithme.

<sup>7</sup> Ici, entendu comme terme quasiment équivalent de « continuum dialectal ».

Labow et Trudgill cité en amorce de cette contribution, sinon mis à mal, du moins défié sur l'une ou l'autre de ses prémisses

- (1) Alors que la phonologie des variétés de mazatec diffère relativement peu, les différences morphologiques (flexion verbale) sont déroutantes, et dépassent de loin tous les pronostics qu'on pouvait en faire sur la base des trois principaux dialectes décrits – *paradoxe de l'hyperdiversification flexionnelle* contre la diversification phonologique faible. Autrement dit, deux villages proches peuvent présenter des parlars quasiment équivalents du point de vue de leur système phonologique, mais différer radicalement à divers endroits de leur système de conjugaisons ou de *classes flexionnelles*.
- (2) Les aires dialectales mazatèques se caractérisent par leur caractère à la fois de continuité structurale à grande distance, entre basses terres et périphérie des hautes terres, et de diversité forte en relation de proximité – *paradoxe de la continuité morcelée ou scalaire (non linéaire)*. Ceci revient à dire que le *continuum dialectal* est à divers endroits aberrant : par exemple, les deux villages de San Lorenzo et de Huehuetlán sont très proches de San Jerónimo Tecoatl, au nord-ouest de l'aire mazatèque, et participent indéniablement de la même dynamique évolutive « néogrammairienne » (évolutions du vocalisme et du consonantisme) que ce petit centre urbain, mais les différences structurales sont pourtant très fortes – mon impression est même que Huehuetlán a neutralisé son système de tons, ni plus ni moins, ou l'a converti en accent intonné.
- (3) Les résultats des observations diasystémiques infirment ou controuvent souvent des faits que l'on croyait acquis du point de vue de l'ethnohistoire – en somme, les données parlent contre les « grandes narrations » ethnohistoriques reçues pour cette région, comme par exemple l'incidence du « superstrat mixtec », par intrusion d'élites dominantes exogènes. Cependant, il y a moyen de réconcilier les deux approches, en reconfigurant les scénarios de peuplement et de diversification des communautés. On peut parler du *paradoxe de l'interprétation décalée des centres directeurs* : s'il y a effectivement des phénomènes de diffusion sociolectale ou géolectale, c'est probablement dans la rivalité entre « town dialects » de différentes époques et centres récents de colonisation de terres à défricher – ce qui revient à suggérer l'idée d'un système de diffusion en cascade. Par exemple San Lucas Zoquiapan a pu être un centre directeur au cours des derniers siècles, mais semble avoir perdu toute incidence sur son arrière-pays, d'après nos observations de terrain.
- (4) Le diasystème semble avoir connu à date récente de fortes turbulences, dues à des ruptures abruptes dans les systèmes agraires et les modèles d'*eusocialité* (conditions d'équilibre et de réciprocité intra- et intercommunautaires ; le concept est de Edward O. Wilson, voir *infra*, section 6) : irruption des systèmes agraires mercantiles (*milpa* méso-américaine *versus* café, élevage et canne à sucre), construction de deux barrages hydroélectriques avec déplacement de villages entiers, relocalisés de bric et de broc dans des terres à défricher, et enfin, vagues migratoires dues à la dépression économique, résultat de ces déséquilibres démographiques et agraires, surtout après la crise du café des dernières décennies du siècle passé. On peut voir là un *paradoxe de la résilience*<sup>8</sup> et de l'*intrication* ou de l'*imbrication des systèmes locaux*. Ce dernier point peut expliquer nombre des aberrations évoquées dans les trois points précédents, dans la mesure où les turbulences qu'ont connues les communautés urbaines et rurales de différentes tailles de l'espace mazatec, sous l'impact des trois systèmes mercantiles (élevage, canne à sucre et café) ont provoqué des déplacements de populations, tantôt par segments (quartiers, hameaux), tantôt par migrations étagées sur plusieurs générations. En outre, la construction du barrage Miguel Alemán, de 1944 à 1954, a déplacé un grand nombre

---

<sup>8</sup> Par résilience on entend généralement le « regain » ou résistance à l'épuisement ou à l'éradication aussi bien de systèmes de culture (dans le domaine agricole) que de traits socioculturels dans une communauté humaine.

de petits agriculteurs de microfundios des basses terres, et les a relocalisés tout autour de la lagune artificielle, mais aussi jusque dans des villages nouveaux situés dans l'Etat de Veracruz. La miscigénation (ou *l'intrication*) des variétés des basses terres qui en a résulté est encore aujourd'hui bien visible, et en cours de développement, donnant lieu à tout un éventail de variétés mixtes<sup>9</sup>.

Conformément à l'esprit empirique qui préside dans les recherches et l'épistémologie du Lacito, c'est à travers des séjours de terrain, mais aussi grâce à un travail co-participatif de « recherche-action » avec les populations locales (cf. <http://axe7.labex-efl.org/taxonomy/term/12>, pour des ressources coopératives en ligne), que ces observations ont pu prendre forme, et faire l'objet de diverses stratégies de modélisation (description des systèmes de classes flexionnelles, dialectométrie, systèmes complexes, etc.). Nous chercherons à montrer comment un terrain complexe peut se décrire comme un *terrain-monde*, dans le sens braudelien et wallersteinien (cf. Wallerstein 2006) des *systèmes-mondes*.

Mais auparavant, je vais mettre en perspective cette exploration du « diasystème-monde » mazatec avec des questions fondamentales de dialectologie générale et, partant, de linguistique générale.

Lorsqu'un romaniste se trouve face à une aire de variation dialectale dense, il lui est toujours possible d'expliquer cette diversité par un jeu de substrats et de superstrats, soit entre peuples de même origine par exemple celtiques), soit d'origines différentes (par exemple, germaniques) - les deux communautés linguistiques convergeant de toutes façons dans un état de langue romane. Pourquoi la dialectologie a-t-elle si souvent recours à ces notions, comme Pierre Gardette invoquant les mânes des Ségusiaves au centre, des Eduens au nord, des Arvernes à l'Ouest, des Vellaves au Sud-Ouest et des Allobroges au Sud-Est pour expliquer la diversification dialectale du Forez, entre Saône et Loire au Nord, Haute Loire et Ardèche au Sud, Puy de Dôme et Lyonnais d'Ouest en Est ? Au détour des pages, les dialectologues du domaine gallo-roman aiment à conter la geste des anciennes tribus gauloises, qu'on imagine aussi divisées qu'intépides, formant une confédération informelle et fragile de fiers guerriers moustachus, avant que ne vienne se superposer l'avancée civilisatrice et unificatrice des Romains, disciplinés et glabres, qui finit par avoir raison des dialectes celtiques continentaux. La dialectologie fennique, faute d'invoquer la dialectique civilisatrice gréco-latine face aux peuples rustiques, a recours à d'autres poncifs : les allées et venues d'agiles trappeurs ou de sourcilieux marchands de zibelines dans les contrées boréales<sup>10</sup>. Non seulement les

---

<sup>9</sup> Une mission a été consacrée à cette question en avril-mai 2018, dans le cadre de l'opération PPC11 du LabEx EFL, confirmant cette observation. La situation socio-économique et sécuritaire de cette zone étant devenue plus difficile que jamais, suite à une succession de gouvernements peu orientés vers le social, depuis 2006, les données devront être complétées par d'autres missions à venir dans cette même zone. Fabio Pettirino, notre enquêteur, a pu constater que le mazatec de San Pedro (IX) se maintenait encore solidement dans certaines communautés mazatèques déplacées dans l'Etat de Veracruz, comme la commune de Nuevo Ixcatlán, ce qui ouvre des perspectives pour cette piste empirique, sur cette forme de changement linguistique induit par une « catastrophe » externe (la construction d'un barrage). La thèse de Diana Schwartz (2016) montre, de manière très stimulante pour la recherche, que contrairement à la vision essentialiste qui projetait une implantation ancienne des populations déplacées, la zone aujourd'hui inondée avait récemment été ouverte à des fronts pionniers de colonisation, qui la rendaient d'ores et déjà dialectalement intriquée, avant le déplacement et la relocalisation forcée de ces mêmes populations.

<sup>10</sup> Ainsi, dans sa monographie sur la phonétique historique du vepse, le (post)néogrammaire Tunkelo explique ainsi un phénomène de primarisation  $y > i, \emptyset > e$  qu'il observait sporadiquement en vepse central, explicable par des mécanismes universels, quoique endémiques, relevant de formes de neutralisation des conditions de marquage segmental : « *Le seul moyen d'expliquer [ce phénomène] est de supposer qu'au moins en vepse du nord, les formes en i (et en e) sont des vestiges et des indices d'un mélange entre caractéristiques dialectales survenu quelques siècles auparavant (sic !), - à une époque rustique où la vie de trappeurs (chasse, pêche, et surtout, l'exercice du commerce des fourrures) a pu mettre en contact avec les régions lacustres de chasse et de pêche comme Syväri et*

interprétations de la diversité dialectale passée et présente souffrent trop souvent de la facilité qu'est le recours aux strates et aux stratifications successives (la trinité *substrat*, *superstrat*, *adstrat*) sans scénario articulé de la périodisation et des modalités sociologiques et (glotto)politiques du contact<sup>11</sup>, mais les termes mêmes de bilinguisme et de diglossie sont souvent considérés comme allant de soi, alors qu'ils se réfèrent à une gamme de situations qu'il conviendrait d'analyser plus en détail, et surtout, de manière plus articulée, et avec grande prudence.

Une tâche en revanche féconde consiste à modéliser les périodes de développement (ou de déploiement dans un espace géographique ou social) de tout diasystème, en partant tout d'abord de critères internes (par exemple, contraintes phonotactiques et morphotactiques, entre segments et morphèmes, du point de vue de la « double articulation » des structuralistes), voire en explorant des probabilités de remotivation et de réanalyse. C'est ce que nous allons faire ici, en prenant le mazatec comme observatoire neutre - heureux d'échapper à toute possibilité d'intrusion d'un substrat ou d'un superstrat, cette fois-ci non pas ségusiave ou arverne, mais chinantec ou mixtec.

En somme, cela revient à faire appel à deux types de modèles explicatifs, face aux données dialectales qui s'offrent à notre regard à travers des sources de seconde main (Kirk 1966) et de première main (données de l'ALMaz) : d'une part un *modèle de constituance morphotactique* (et, partant, morphologique), selon que les unités constitutives du lexique sont plus ou moins libres ou liées, d'autre part, un *modèle en cascade* (entendu cette fois, comme dynamique interne plutôt qu'externe), selon que les choix entre des formes détachées versus compactées ont une incidence sérielle les unes sur les autres, en généralisant les procédés d'un paradigme à l'autre.

La position du mazatec dans la constellation des langues autochtones (« idiomas originarios ») du Mexique est indiquée par un cercle dans la figure 1 ci-dessous.



Figure 1 : Le mazatec en Més-Amérique (bassin du Papaloapam).  
Cartographie : Vittorio dell'Aquila (CELE, 2010).

Les principales subdivisions dialectales sont indiquées dans la figure 2, qui est une carte schématique rendant compte de la classification des dialectes mazatecs à partir d'une synthèse des données de Kirk (1966) – cf. Léonard, dell'Aquila & Gaillard-Corvaglia (2012). La

---

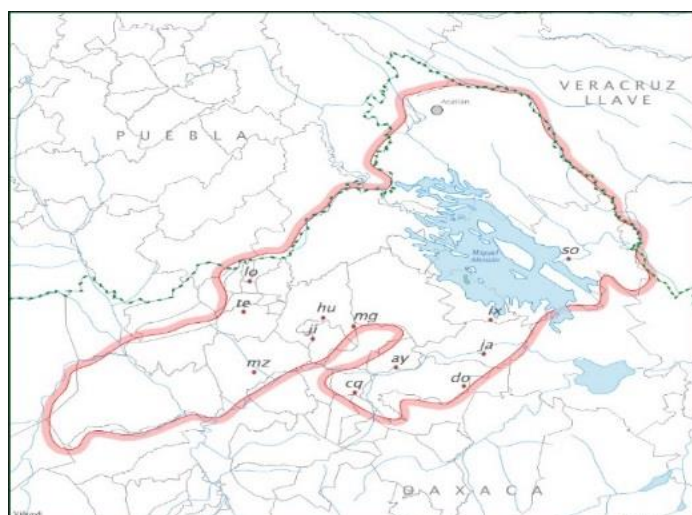
*le lac Onega des peuples éloignés, vivant sur les rives de la Baltique* » (Tunkelo, 1946 : 586). La traduction est de l'auteur de ces lignes.

<sup>11</sup> Il va de soi que réduire l'explication en dialectologie à la seule stratification des variétés et des langues en contact dans une aire serait caricatural, mais ces exemples gallo-romans et fenniques sont instructifs sur un mode de narration qui a eu longtemps cours. On consultera utilement Boberg & al. (2018) pour un panorama complet et très à jour des méthodes modernes en dialectologie, et Carrilho & al. (2013) pour un survol récent des approches en géolinguistique et sur la notion de frontières dialectales.

principale division est tripartite : elle oppose les parlers des hautes terres (par exemple, la partie « alta central » associé à l'aire « Noroeste A & B » dans le schéma de la figure 2) à des parlers des terres moyennes ou piémont (« Media »), eux-mêmes distincts, quoique désormais fortement convergents avec ceux des basses terres (« Baja »). Cependant, la gamme chromatique indique que hautes et basses terres ont formé un bassin davantage unitaire à date ancienne (couleur jaune), tandis que les parlers du piémont (couleur verte) peuvent être considérés comme une aire innovante plus récente, à partir d'une amphizone (autrement dit, une aire intermédiaire) entre les deux pôles (cf. figure 2 *infra*). Ceci dit, il importe de rappeler que cette classification, aussi justifiée soit-elle sur l'analyse de variables phonologiques et morphologiques, n'en reste pas moins une *narration* parmi d'autres. La classification dialectale est certes riche d'indices sociohistoriques et géohistoriques, mais il faut toujours se garder d'un seul point de vue. Les données, aussi complètes soient-elles, ne sont jamais exhaustives, et les résultats des regroupements peuvent varier d'une composante du lexique ou de la grammaire à l'autre<sup>12</sup>.



Figure 2 : Principales aires dialectales du mazatec : division quadripartite.



<sup>12</sup> Cf. Léonard & al. 2014, pour le tseltal (maya occidental), à l'aide d'une analyse dialectométrique par algorithme de Levenshtein, avec des données de première main, ciblées sur le noyau dur de variation diasystémique du réseau dialectal tseltal. Dans ce diasystème, les asymétries de configurations géolinguistiques entre la phonologie et la morphosyntaxe sont frappantes. Plusieurs strates historiques de diffusion des dialectes tseltal, et de contacts internes au sein du réseau dialectal, mais aussi plusieurs narrations possibles de *l'ethnohistoire* (pour autant que ce terme ait un sens), sont possibles.

Figure 3. Variétés documentées in Kirk (1966)<sup>13</sup>.  
Cartographie : Vittorio dell'Aquila (CELE, 2015).

Toujours selon une logique de « grandes lignes », il convient de distinguer une quatrième subdivision, dont la diversité interne est bien plus grande que celle suggérée ici par le vert mat de la carte schématique : de part et d'autre du Canyon de Cuicatlán (appelé Cañada dans la région) sont situées deux variétés difficiles à classer : d'une part, dans la partie occidentale du Canyon (Cañada occidental), le parler de Mazatlán Villa de Flores, qui présente des traits aussi bien des hautes terres que des basses terres ; d'autre part, la variété de Chiquihuitlán qui, elle, relève nettement du sous-diasystème oriental (piémont et basses terres : Media et Baja sur le schéma), et s'avère particulièrement distincte de toutes les autres variétés à bien des égards. Il est impossible d'entrer davantage dans les détails de ces divisions d'ordre général ; en revanche, dans le présent article d'hommages au Lacito, j'explorerai des détails de la trame diatopique<sup>14</sup> du mazatec, afin de suivre fidèlement la tradition de description à grain fin des langues, dans un souci de contextualisation anthropologique. Ce parti-pris privilégiant la recherche du général, voire de l'universel dans un *milieu humain contextualisé* – une écologie locale et régionale – me semble préférable à une logique de « grande fresque », et me semble caractériser l'esprit du Lacito. La figure 3 localise les douze variétés retenues par Paul Livingston Kirk dans sa monographie comparative des dialectes mazatecs (Kirk 1966) : toutes ces localités sauf Santo Domingo (Do, cf. note *infra* pour la liste des abréviations des localités) ont également fait l'objet d'enquêtes plus ou moins poussées par l'équipe ALMaz (*Atlas Lingüístico Mazateco*) du projet MAmP de l'Institut Universitaire de France, aujourd'hui continué par les opérations PPC11 et EM2 des axes 1 et 2 du LabEx EFL<sup>15</sup>.

Cependant, l'absence de Do dans nos données de première main est largement compensée par l'exploration de localités non documentées par Paul L. Kirk, notamment Santa Maria Chilchotla (2013)<sup>16</sup>, Santa Cruz Acatepec (2012), San Antonio Eloxochitlán (2011,

---

<sup>13</sup> Abréviations : Ay = Ayautla, Cq = Chiquihuitlán, Do = Santo Domingo, Ix = San Pedro Ixcatlán, Ji = Jiotes, Hu = Huautla, Ja = Jalapa, Lo = San Lorenzo, Mg = San Miguel Huautla, So = San Miguel Soyaltepec, Te = San Jerónimo Tecoaatl (conventions reprises de Kirk 1966).

<sup>14</sup> La *trame diatopique* peut se qualifier comme le *grain fin du continuum dialectal*, du point de vue de l'observation des données dialectales.

<sup>15</sup> La recherche présentée ici a bénéficié d'une aide de l'Etat gérée par l'Agence Nationale de la Recherche au titre du programme « Investissements d'Avenir » portant la référence ANR-10-LABX-0083 dans le cadre de deux opérations du LabEx EFL (PPC 11 : Complexité et diffusion des systèmes phonologiques et EM2, axe 7 : élicitations croisées). Elle est issue par ailleurs d'un projet quinquennal financé par l'IUF (Institut Universitaire de France), de 2009 à 2014. Pour plus de détails, v. <http://jll.smallcodes.com/home.page> et <http://axe7.labex-epl.org/taxonomy/term/12>.

<sup>16</sup> De manière insolite, une fois de plus, cette variété présente un intéressant paradoxe : alors que tous les locuteurs de Huautla que nous avons sondés au sujet de toute différence dialectale potentielle avec la variété de Santa Maria Chilchotla affirmaient mordicus qu'il s'agissait de la même variété, nous avons pourtant clairement rencontré des traits morphologiques, dans le système verbal (préverbes de classe 1A), qui relèvent typiquement de l'aire de San Antonio Eloxochitlán et de San Jerónimo Tecoaatl (TE), dont le parler de Santa Cruz Acatepec, évoqué dans la section 5 *infra*, est un bon exemple. Un tel trait ne tombe pas du ciel, car dans le diasystème mazatec, ce changement est conditionné non pas par un rehaussement  $e > i$  du vocalisme du préverbe  $b'é-$ , mais par un changement de classe flexionnelle. En outre, si ce rehaussement devait s'interpréter comme phonologique, il resterait à expliquer pourquoi il serait cantonné à ce préverbe, et dérogerait à l'inexorabilité attendue des changements phonétiques. Or, on sait que la zone de SM Chilchotla est une zone facétière de peuplement récent, par fronts pionniers agraires, qui s'est constitué en bassin de traits imbriqués, suite à l'afflux de locuteurs provenant aussi bien de Huautla que de Ayautla (dialectes à préverbes  $b'é-$  de classe 1A) que de l'aire dialectale du nord-ouest (San Lucas, San Antonio, San Pedro Ocopetatlillo, etc.) (variétés à préverbe 1A en  $bi-$ ). Là encore, le modèle centrifuge est infirmé par une dynamique de *miscigénéation*, ou imbrication dans un bassin de traits.



2013), Santa Maria Ateixtlahuaca (2011, 2016), San Pedro Ocopetatillo (2013, 2016), San Mateo Yoloxochitlán (2011), Soyaltitla (2013), La Toma (2015)<sup>17</sup>.

Les enquêtes ont été rendues possibles grâce au travail social bénévole réalisé par l'équipe des enquêteurs, qui ont tous réalisé des ateliers d'élaboration de matériaux pédagogiques en langues autochtones – principalement mazatec, dans une dizaine de variétés<sup>18</sup>. Des stages de plusieurs jours de tonologie, phonologie diachronique, morphologie flexionnelle, donnés à titre gracieux tout au long du projet (encore en cours, d'ailleurs, sous de multiples formes plus ou moins auto-financées), ont contribué également à rendre plus explicite et pertinent le travail de dialectologie et de linguistique que réalisait l'équipe au fil des ans (de 2010 jusqu'à ce jour et, espérons-le, au-delà)<sup>19</sup>.

## 2. Eléments de phonologie mazatèque

### 2.1. Généralités

Le protosystème du mazatec reconstruit par Gudschinsky (1958a) et repris par Kirk (1966) est le suivant pour les consonnes : six occlusives et mi-occlusives (ou affriquées) *\*t, \*tʃ, \*k, \*kʷ, \*ts, \*tʃ*, deux fricatives sibilantes *\*s, \*ʃ*, trois sonantes nasales *\*m, \*n, \*ɲ*, deux approximantes ou glides *\*j, \*w*, et deux glottales *\*h, \*ʔ*. Les conventions graphémiques actuelles (graphie ALFALEIM, ou *Alfabeto para la Lengua Indígena Mazateca*)<sup>20</sup> sont les suivantes : <*t, ti, k, ku, ts, ch ; s, x ; m, n, ñ ; y, b*> – les graphèmes sont indiqués entre crochets pointus. Quant aux voyelles, le système est pentavocalique (*\*i, \*e, \*a, \*u, \*o*), avec des jeux d'alternances distributionnelles pour les deux voyelles haute et moyenne postérieures /u/ et /o/ ; il faut compter également avec une classe de huit diphtongues (*\*ia, \*iu, \*ei, \*ai, \*ui, \*ue, \*ua, \*au*), qui partagent toutes une voyelle haute *i, u* comme pivot. Nous avons traité ces systèmes par ailleurs (v. Léonard *et al.* 2012, Léonard 2010b), rappelant qu'une liste de consonnes comme celle-ci est représentative des pivots contoïdes. Il faut compter également avec la qualité de voix (craquée, notée en API à l'aide d'un tilde souscrit :  $\underset{\sim}{C}$  et  $\underset{\sim}{V}$  ; soufflée, notée par deux points souscrits :  $\underset{\cdot\cdot}{C}$  et  $\underset{\cdot\cdot}{V}$ ) ainsi que de nasalité (occlusives prénasalisées, voyelles nasales). Le statut des voyelles dites « réarticulées » ( $\underset{\sim}{V}\underset{\sim}{V}$ ) ou de celles qu'on pourrait appeler « réaspirées » ( $\underset{\sim}{V}h\underset{\sim}{V}$ ) relève, selon l'auteur de ces lignes, d'une des propriétés prosodiques et syllabiques de la qualité de voix – ces noyaux complexes sont des voyelles respectivement craquées et soufflées en situation de plénoglossie<sup>21</sup>, autrement dit, en position forte<sup>22</sup>. Nous les noterons

---

<sup>17</sup> A ce titre, outre l'auteur de ces lignes, il faut remercier les enquêteurs de l'ALMaz : Antonia Colazo-Simon, Fabio Pettirino, Karla Avilés Janiré González, Jaime Ramón Calderón, et les informateurs, dont le nombre dépasse la centaine. Leur participation au projet a été d'autant plus spontanée et enthousiaste que beaucoup ont pu observer ou participer au travail social en linguistique appliquée mazatèque que nous avons développé bénévolement dans la communauté, aussi bien dans les centres urbains que dans les villages de l'aire mazatèque (et nahuatl : enclave de Santa Maria Teopoxco).

<sup>18</sup> Les résultats ainsi que la méthodologie de ces ateliers ou stages gratuits sont accessibles sur le lien de l'opération EM2 du LabEx EFL : <http://axe7.labex-efl.org/taxonomy/term/12>. Pour plus de détails quant aux méthodes et initiatives, v. <http://axe7.labex-efl.org/node/126>.

<sup>19</sup> Voir en particulier l'atelier de grammaire mazatèque de Huautla (<http://axe7.labex-efl.org/node/171>), et sa version nord-occidentale (<http://axe7.labex-efl.org/node/87>), ou sud-orientale (<http://axe7.labex-efl.org/node/123>).

<sup>20</sup> Fascicule didactique pour les écoles bilingues mazatèques disponible sur le lien du CEDELIO (*Centro de Estudios y Desarrollo de las Lenguas Indígenas de Oaxaca*) : <http://cedelio.edu.mx/BIBLIOTECA/index.php?libro=metodoalfaleim>.

<sup>21</sup> Ou « pleine voix ». Bien que quelque peu désuet, le terme de *plénoglossie* met bien en valeur la saillance de cette structure vocale déployée que sont les voyelles interrompues, ou réarticulées.

<sup>22</sup> Autrement dit, en position tonique, de mot plein sur le plan lexical, s'opposant à des allomorphes légers relevant de catégories fonctionnelles : par ex. la différence entre *ja'a*, thème aspectuel accompli de 'venir' en tant que verbe

donc VʔV et VhV plutôt que V̥V et V̥V. Nous souscrivons pleinement à la théorie de Golston & Kehrein (1998, 2004), selon laquelle la glottalisation se manifeste en mazatec par des traits autosegmentaux qui suivent le *principe de l'alignement senestre* (reporter la réalisation du trait sur la gauche du segment : C̥ → ʔC, C̥ → hC ; V̥ → ʔV, V̥ → hV. Dans la graphie, la glottalisation craquée pour les voyelles et implosive ou éjective pour les consonnes est notée par une apostrophe < ' >, tandis que l'aspiration ou la voix soufflée est notée par la « jota » < j >. L'inventaire des pivots contoïdes<sup>23</sup> et des autosegments du point de vue du diasystème figure dans le tableau 1 ci-dessous. Pour plus de détails sur les inventaires consonantiques de chaque variété de mazatec, cf. Kirk 1966 : 11-25, qui recense les spécificités de chacune des douze localités retenues dans son enquête.

En ce qui concerne l'inventaire tonal, l'état initial reconstitué par Gudschinsky et Kirk (*op. cités*) consiste en quatre tons de niveau (*level tones*) : Haut (H), mi-haut (h), moyen (M) et bas (B), et trois tons montants : Moyen-Haut (MH), Bas-Haut (BH) et Bas-Moyen (BM), soit les classes \*H, \*h, \*M, \*B et \*MH, \*BH et \*BM (Kirk 1966 : 8). Dans la graphie, le ton est noté sur les voyelles : par un accent aigu pour le ton haut, un accent grave pour le ton mi-haut, un soulignement pour le ton bas, et aucun diacritique pour le ton moyen (ex. < á, à, a, a >). Les contours tonaux sont reportés sur des séquences de deux voyelles selon les conventions de l'orthographe ALFALEIM, valant pour deux mores, même si le mazatec ne connaît pas l'opposition entre voyelles brèves et voyelles longues. Les quatre tons de niveau se répartissent sur deux paliers ou deux registres : élevé (H, h) et bas (M, B), si bien que les contours les plus marqués sont BH et MH par ordre décroissant et que le contour BM serait le contour *défaut* – très utilisé dans la flexion verbale, notamment pour le marquage subjectal 1SG et le marquage de l'inaccompli dans certaines classes.

Les racines sont dotées d'un ton lexical, qui implique les quatre tons ponctuels et deux contours à Huautla selon K. Pike (1948), qui définit six classes de ton de racine (CTR) : CTR/I = H ex. *b'éxá* 'envoyer', CTR/II = h *sítsjò* 'griller, torréfier', CTR/III = M ex. *b'éma* 'se cacher', CTR/IV = B ex. *síxkoá* 'casser', CTR/V = BM ex. *b'éñai* 'enterrer', CTR/VI = Mh ex. *bántià* 'ouvrir une brèche, guider'.

	LABIAL	CORONAL			DORSAL	LARYNGAL
		[apical]	[haut]	[rétrofl.]		
OCCLUSIVES	(p)	t			k	ʔ
AFFRIQUEES		ts	tʃ	t		
FRICATIVES	ϕ	s	f[ʃ]			h
SONANTES NASALES	m	n	ɲ			
LIQUIDES		l				
APPROX/GLIDES	w		j			

Tableau 1. Segments consonantiques en mazatec (*pivots* et autosegments<sup>24</sup>)

Le vocalisme se résume à un système pentavocalique (*i, u, e, o, a*) x 4 (voix modale, craquée, soufflée, et nasalité), avec distribution complémentaire de /o/ en relation à /u/ à

de mouvement, contrastant avec l'allomorphe *j'a=*, proclitique d'aspect ingressif accompli susceptible de s'adjoindre à n'importe quel verbe.

<sup>23</sup> Ce terme vaut pour « segment de type consonantique ». La valeur générique et sous-spécifiée de ce terme est ici particulièrement bienvenue, pour un inventaire consonantique aussi ambivalent que le mazatec, et la plupart des langues otomanges, avec des consonnes pivots s'enrichissant d'autosegments laryngés ou sonantiques.

<sup>24</sup> Nasalité, glotte fermée et glotte ouverte : ces traits autosegmentaux sont ici décrits comme des segments dans les trois cellules grisées du tableau.

Huautla et dans d'autres variétés, avec des idiosyncrasies distributionnelles (notamment à Chiquihuitlán). Une nette tendance à la diphtongaison ascendante de /e/ est observable à Mazatlán (Mz), mais aussi à San Pedro Ixcatlán (Ix). Jalapa ainsi que Chiquihuitlán (Ja & Cq) font alterner deux variantes complémentaires de /e/ : l'une mi-haute [e], l'autre très abaissée [æ]. Le parler de San Lorenzo (Lo) est le seul à attester un *vowel shift* (ou chaîne de propulsion et de traction) complet :  $i > e, e > a, a > o, u > u$ . On recense quatre tons de niveau à Huautla (H, h, M et B : le symbole |h| note le ton mi-haut), trois seulement à San Felipe Jalapa de Díaz (H, M, B), et entre quatre et six tons de contour à Huautla, selon différents descripteurs (**HB**, **BM**, **HM**, **Bh**, **BhM**, **BhB**) – en gras, ceux retenus ici.

Le tableau 2 *infra* donne quelques exemples de cognats, choisis en fonction de leur représentativité, selon une intéressante hypothèse de Sarah Gudschinsky, qui postule qu'en mazatec, la compression gabaritique<sup>25</sup> des lexèmes, pour la plupart construits sur le procédé de composition, rend compte de la plupart des isoglosses phonologiques (Gudschinsky 1956)<sup>26</sup>. Les tons ne sont notés que pour trois dialectes, dans les données de Kirk (1966), pour des raisons liées aux conditions matérielles de ses enquêtes de terrain : Hu, Ja et So, ce qui limite la comparabilité des données dans cette source – mais les structures tonales nous seront ici moins utiles que dans l'examen de la morphologie verbale, où nous les prendrons en compte, à partir de nos données de première main.

	Var.	A 'large'	B 'fourmi'	C 'ver'	D 'grotte'	E 'épis de maïs'	F 'jaune'	G 'aile'
1	MZ	<i>ntuhu</i>	<i>nĩʔju</i>	<i>ʃuntu</i>	-	<i>nʃutĩ</i>	<i>sinĩ</i>	<i>lahnka</i>
2	AY	<i>ntuhu</i>	<i>nĩʔiu</i>	<i>ʃuntu</i>	<i>nkiho</i>	<i>nʃutĩ</i>	<i>sinẽ</i>	<i>nka</i>
3	CQ	<i>ntuhu</i>	<i>ntĩʔiu</i>	<i>ʃuntu</i>	<i>nkeʔiho</i>	<i>nʃutĩ</i>	<i>sinẽĩ</i>	<i>thiunka</i>
4	JA	<i>ndhuʔ</i>	<i>nĩʔiuʔ</i>	<i>ʃhuʔntuʔ</i>	<i>ngiʔhoʔ</i>	<i>ʃuʔtĩʔ</i>	<i>siʔnẽʔ</i>	<i>ntshaʔnkaʔ</i>
5	DO	<i>nduhu</i>	<i>ndiʔiu</i>	<i>ʃhintu</i>	<i>ngiho</i>	<i>ʃutĩ</i>	<i>sinẽ</i>	<i>ntshanka</i>
6	HU	<i>ntuʔ</i>	<i>nĩʔjuʔ</i>	<i>ʃuʔntuʔ</i>	<i>nkaʔhauʔ</i>	<i>nʃuʔtĩʔ</i>	<i>siʔnẽʔ</i>	<i>hnkaʔ</i>
7	JI	<i>ntuhu, ntu</i>	<i>nĩʔju</i>	<i>ʃuntu</i>	<i>nkaho</i>	-	<i>sinẽ</i>	<i>hnka</i>
8	SO	<i>ntuʔhuʔ</i>	<i>nĩʔjuʔ</i>	<i>ʃuʔntuʔ</i>	<i>nkiʔhoʔ</i>	<i>nʃuʔtĩʔ</i>	<i>siʔnẽʔ</i>	<i>thjuʔnkaʔ</i>
9	IX	<i>ntuhu</i>	<i>ntĩʔiu</i>	<i>ʃuntu</i>	<i>nkihu</i>	<i>nʃitĩ</i>	<i>sinẽ</i>	<i>nka</i>
10	MG	<i>ntu</i>	<i>nĩʔju</i>	<i>ʃuntu</i>	-	-	<i>sinẽ</i>	<i>nka</i>
11	LO	<i>ntuhuw</i>	<i>nĩʔji</i>	<i>ʃintuw</i>	<i>nkohuw</i>	<i>nʃitẽ</i>	<i>sinã</i>	<i>nko</i>
12	TE	<i>ntuhu</i>	<i>nĩju</i>	<i>ʃintiu</i>	<i>nkiho</i>	<i>nʃitĩ</i>	<i>sinẽ</i>	<i>nka</i>

Tableau 2. Données relatives aux isoglosses, notation API (source : Kirk 1966).

Les cycles de Gudschinsky (ou CG) ont ceci d'intéressant qu'ils permettent de décrire mais aussi d'expliquer l'évolution phonologique du diasystème mazatec à travers un modèle d'ajustement morphonologique sobre et puissant, dont le détail est explicité dans la section 2, et les conséquences systémiques, de la section 3 *infra*. L'essentiel tient dans l'idée que les gabarits compositionnels de type Classificateur + Racine polyvalente par exemple pour les

<sup>25</sup> Un gabarit est un schème CVCV abstrait, ancré sur des positions squelettales, ou positions consonantiques (C) et vocaliques (V). Ce schème comprend des positions vides ou pleines, dont les interactions phonotactiques sont contraintes par des universaux. Cf. Scheer 2015.

<sup>26</sup> Je remercie Daniel Jonathan Vielma, de l'UAM, du *Séminaire de langues otomangues*, qui m'a fourni ce document, difficile d'accès, qui n'est autre que le mémoire de master (Master thesis, manuscrit inédit) de la jeune Sarah Gudschinsky – appelée à devenir ensuite la principale spécialiste du diasystème mazatec, avant Paul L. Kirk.

substantifs (mais c'est tout aussi vrai pour les verbes, selon un schéma Préverbe + Racine polyvalente) ont fusionné en lexèmes insécables : [Y[X]] > [YX]<sup>27</sup>.

Du point de vue de la morphologie lexicale dans sa dimension morphologique, on aboutit de fait, au terme des cycles de Gudschinsky, à des stems ou radicaux caractérisés par la qualité de voix (modale, craquée ou soufflée), comme dans la liste récapitulative ci-dessous :

*Caractérisation des radicaux (gradation et qualité de voix) dans les items du tableau 2 :*

Radical A : Radical **fortis**, à noyau réarticulé **breathy**

Radical B : Radical **composé**, de type phase 1, avec racine à voix craquée (*creaky*)

Radical C : Radical **composé**, de type phase 1, avec racine **prénasalisée**

Radical D : Radical **composé**, de type phase *n*, avec racine à voix soufflée (*breathy*)

Radical E : Radical **composé**, de type phase 1, avec racine **modale**

Radical F : Radical **composé**, de type phase *n*, avec racine **nasale**

Radical G : Radical **composé** ou **simple**, de type phase 1, avec racine **prénasalisée** et/ou **soufflée**

La synchronie du diasystème donne donc ces caractérisations, qui jouent sur des contrastes de qualité segmentale (*fortis* vs. *lenis*) et infrasegmentale ou autosegmentale (voix modale vs. craquée ou soufflée). Voyons quel en est le déterminisme diachronique, du point de vue des cycles de Gudschinsky.

## 2.2. Cycles de Gudschinsky (CG)

Les cycles de Gudschinsky relèvent, conformément aux souhaits exprimés dans l'introduction, à la fois d'un modèle de déterminismes internes, sur la base des phases de compression des séquences morphémiques au sein des lexèmes et, de manière plus générale, de la syntaxe, et d'un modèle en cascade, dans la mesure où chaque cycle de compression gabaritique et les mécanismes impliqués dans ces cycles, est susceptible d'avoir des conséquences en chaîne, selon la seconde logique de notre approche, qui est le modèle en cascade.

**CG-1** : Phase du protomazatec, avec clusters consonantiques suivis ou non d'une semi-voyelle, et de voyelles simples ou de diphtongues décroissantes (*au, ai*) | croissantes (*ia, ua*), selon le modèle de constituance de Pike & Pike (1947), qui stipule que les groupes consonantiques et vocaliques sont bel et bien des clusters inorganiques – pas de corrélation de voix craquée ou soufflée à ce stade ; les sonantes en position d'attaque initiale absolue par exemple sont précédées d'une occlusion ou d'une friction glottale de type *fortis*, purement distributionnelle, en vertu de leur position de début de groupe de souffle ou d'émission.

**CG-2** : une réforme métrique se met en place, à partir de morphèmes qui, dès le protomazatec, faisaient encore alterner des formes courtes et des formes longues, comme le doublet \**hʃanka / hʃa* 'vieux', ce qui alimente le lexique en doublons avec variation libre des gabarits CV | CVCV optionnels (le signe | ici valant pour « ou »).

**CG-3** : cette phase d'évolution gabaritique est identifiable au groupe consonantique interne, qui trahit une composition, comme dans *nahna, ʃahnu* en (1a-b) ci-dessous s'analysant comme suit, en termes de procédés de concaténation :

Proclise de classificateur nominal :

---

<sup>27</sup> Je remercie Shanti Ulfsbjorninn, dont l'exemplier de 2016 (v. références) a été une source d'inspiration, ainsi que plusieurs discussions à bâtons rompus préliminaires à cet article. Toutes les erreurs éventuelles sur l'application de son modèle compositionnel sont miennes.

(1a) *na=hna*  
CL<sub>fém</sub>=gousse  
'gousse de maïs'

(1b) *ʃa<sub>masc</sub>=hnu*  
CL=écureuil  
'écureuil',

Dans les exemples (1a-b), un spécifieur catégoriel (CL : CL<sub>fém</sub>, *ʃa<sub>masc</sub>*) est associé par proclise à une racine polyvalente ( $\sqrt{[i]}$ ), à laquelle il attribue la valeur substantivale.

**CG-4** : phase de compression prosodique des gabarits ou des pieds CVCV > CV<sup>T1T2</sup> par contour tonal (T = ton) équivalent à une résorption du matériel segmental :

Compression prosodique d'un noyau complexe ou *fortis* (voyelle réaspirée) :

(2) *\*t̥ihi > t̥i* dans les hautes terres centrales (Huatla).  
'pot, marmite'

**CG-5** : formes encore analysables comme des mots composés dans certaines variétés et comme lexèmes indécomposables dans d'autres :

*tofi > tui < \*tahwi, \*tuhwi*

(3) *\*tu=hwi*  
DEICT=aller  
'doucement, peu à peu'.

Dans le tableau 2 supra, les cognats de la colonne A *ntuhu* « large, grand » à gabarit CVhV (en réalité monosyllabique, avec noyau vocalique complexe, réarticulé, donc *fortis*) sont fidèles à la forme héritée *\*ntuhu*, sauf A6-7 qui attestent le CG4 (même si Kirk n'a pas noté le ton pour la variété Mg). On notera la forme ayant subi la traction vocalique *u > uu* pour Lo (*ntuhu > ntuhuu*), innovation post-CG, donc ultérieure aux cycles de compression, et autonome par rapport à ce paramètre évolutif.

En ce qui concerne *\*niʔjuʔ* « fourni », la phénoménologie observable sur la carte de la figure 4 oppose des formes fortes, ou *fortis*, avec occlusive prénasalisée (Do, Ix) – en réalité, il s'agit de sonantes nasales renforcées par postoralisation, mais il est impossible d'entrer ici dans les détails de ce processus (v. Marlett 1992 et Léonard 2010a, pour le Mixtec) –, qui relèvent de la logique de la phase CG-1, et des formes *lenis* de même phase, en termes de cycles, comme *nĩʔiuʔ*, contre des formes avec réduction vocalique comme *nĩʔji* à Lo.

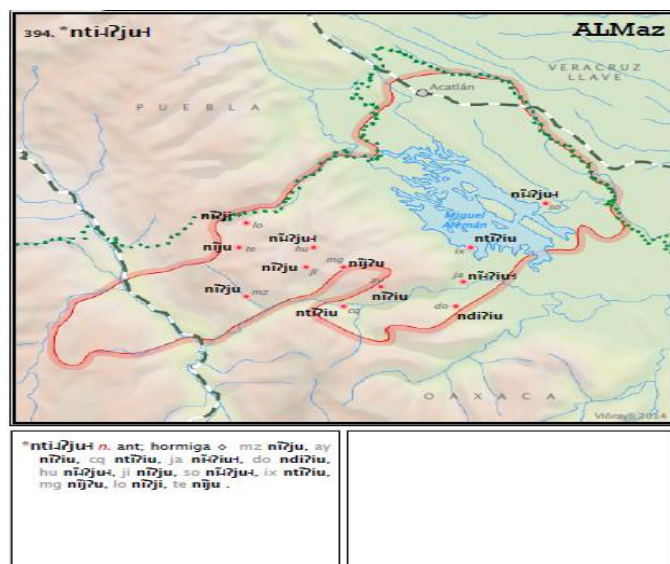


Figure 4. \*nti-ɬjuɬ « fourmi », ALMaz, données de Kirk (1966, cognat 394).  
Cartographie : Vittorio dell’Aquila (CELE).

Une forme comme *nɪʒu* à Mg suggère un réajustement, avec craquement du noyau tête (celui de la racine lexicale : *nɪʒu* = *nɪʒu*) selon une logique de phase CG-3. A Te, la forme *nɪʒu* est typiquement une forme modalisée, de type CG-5, voire ultérieure à cette étape. L’item \**nki-ɬhauɬ*, \**nka-ɬhauɬ* « grotte », formé d’un spécifieur positionnel/directionnel *nki* « (en/vers le) bas » (le vocalisme *nka* peut résulter d’une assimilation régressive ultérieure à CG-1 : \*\**nki-ɬhauɬ* > \**nka-ɬhauɬ*) associé par composition à une racine lexicale *hau*, oppose des formes héritées, pleines, comme *nka-ɬhau* (Hu), à des formes réduites comme *nkiho* (Ay, So, Te) > *nkihu* (Ix) ou *ngiho* (Ja, Do), en ce qui concerne le vocoïde de la racine dans ces formes, toutes de type CG-3, comparativement à une forme réinterprétée mais également réduite, comme *nkohwu*, à Lo – analysable comme *|nka+hw(a)u|*. Toutes ces formes se ventilent sur une gamme qui va de CG-1 à CG-3, qui estompe la dichotomie synchronie/diachronie par la mise à plat dans l’espace du diasystème, lorsqu’on examine les données de Kirk.

#### Conséquences des CG :

Les phénomènes de contraction signalés par Gudschinsky dans son mémoire de 1956 peuvent avoir pour conséquences

- a) l’effacement d’une syllabe toute entière,
- b) l’effacement ou la réduction d’une voyelle interne,
- c) des sonantes nasales intervocaliques et des effets de nasalisation des noyaux syllabiques environnants,
- d) des laryngales internes ou des clusters médians avec laryngales préconsonantiques, des semivoyelles médianes ou des clusters avec glides.

Ces séquences simples passent par différents seuils et/ou phases de compression de ce gabarit CVCV, pour aboutir par fusion morphologique à des allomorphes radicaux insécables. Les dialectes mazatecs divergent tous le long d’un continuum d’options structurales sur la manière d’associer ces composants simples CV<sub>SPEC</sub>+CV<sub>Racine</sub>. La variation dialectale va donc porter sur les processus suivants, dans le diasystème :

- phénomènes de joncture, de *sandhi*,
- de réduction de *timbre* vocalique ou de *manière* consonantique,
- simplification de clusters,
- réalignement des traits glottiques (constriction glottique vs. glotte relâchée vs. voix modale)

- un registre de quatre tons ponctuels ordonnés sur deux paliers prosodique, et quatre à six tons de contour.

### 3. Phases diachroniques et aires dialectales

Dans ce qui va suivre, je vais tenter de formaliser les CG en proposant des représentations gabaritiques. L'hypothèse de Gudschinsky déploie en effet un véritable éventail d'options diasystémiques, qui permettent de concilier diachronie (les phases) et synchronie (les aires). Nous verrons que la dichotomie saussurienne synchronie *versus* diachronie s'estompe, lorsqu'on envisage l'évolution à grain fin d'un diasystème avec cette méthode, en entrant dans le détail des faits de langues, à partir d'une base de données – celle que nous avons constituée à partir des séries de cognats (*cognate sets*) de Paul L. Kirk (*op. cit.*). Nous allons voir qu'une phase comme CG-2, pour laquelle l'auteure retient l'exemple de \***stʃa**l / cq **tʃanka**, avec coexistence ou choix successifs entre doublets lexicaux, est loin de se limiter à un simple polymorphisme, et implique des bifurcations en termes de morphologie lexicale – autant que de phonologie lexicale. Le modèle théorique que nous suivrons ici est adapté de la phonologie déclarative (Angoujard 1997, 2006) : il fonde l'analyse des positions squelettales sur des grilles rythmiques AN (Attaque, Noyau), indiquées par une étoile, auxquelles s'attachent des éléments prosodiques (tons) ou segmentaux (primitives phonologiques). Le modèle de constituance morphologique est inspiré d'Ulfsbjorninn (2016).

#### CG-2 : Eventail gabaritique

Les figures 5.1 et 5.2 rendent compte des phénomènes phonologiques mais aussi constructionnels (morphologie lexicale) qui interviennent pour un item que Gudschinsky cite comme exemplaire de la phase 2 – celle du polymorphisme CVCV / CV.

\***stʃa**l adj. 'old'; 'viejo' ◊ mz **htʃa**, ay **tʃa**, cq **tʃanka**, ja **tʃa**l 'old', 'viejo'; 'big', 'grande', do **tʃa**, hu **htʃa**l, ji **htʃa**, so **tʃa**, ix **stʃa**, mg **tʃa**, lo **kotʃonko**, te **stʃa**.

Protoforme, avec préfixation d'un marqueur adjectival

(4a) \***s-tʃa**l, \***f-tʃa**l  
ADJ-√

Réduction par débuccalisation du préfixe à Mazatlán : cet élément devient un indice relevant de la qualité de voix soufflée, fusionnant avec la racine (démotivation partielle) :

(4b) mz **h.tʃa**,  
ADJ.√  
'vieux, ancien'

Réanalyse et relexification, par conversion de la racine héritée (de statut X) en un préfixe mi-adjectival mi-classificateur nominal (<sup>h</sup>tʃa), de statut Y, associé à une racine polyvalente (nka) :

(4c) cq **tʃa+nka**  
CL/ADJ+√

Un exemple de conversion (N > A > V<sub>statif</sub>) à partir du type (4c) :

(4d) lo **kotʃonko** ← |**ka**= **tʃa+nka** | lit. 'vieilli'  
ASP<sub>ACMPL</sub>= CL/ADJ+√

(4e) <Y> = spécifieur *fortis*, <y> spécifieur *lenis*, <X> = tête catégorielle (partie du discours).

Les figures arborescentes qui suivent peuvent être considérés comme des arbres décisionnels, ou des espaces de construction morphologique, selon une logique implicationnelle, du haut vers le bas de l'arbre. En 5.1, les concaténations morphémiques sont explicitées. En 5.2, les structures gabaritiques sont symbolisées par des séquences de positions segmentales indiquées par des astérisques, valant pour des séquences CVCV. Ces représentations enrichissent la description des structures de constituants au sein des lexèmes.

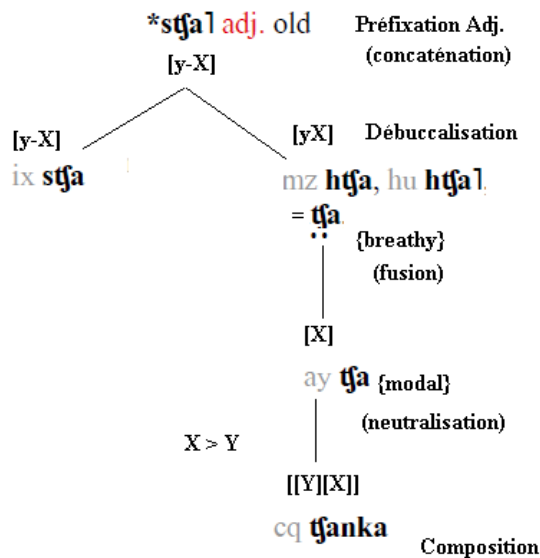


Figure 5.1. \*stʃaʎ adj. ‘vieux’ (‘ancien’) : cycles de concaténation et de composition.

Morphologie lexicale.

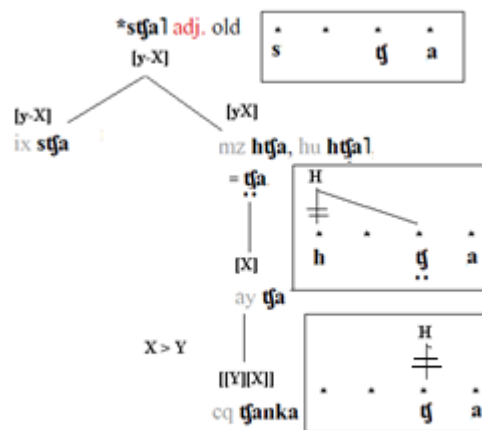


Figure 5.2. \*stʃaʎ adj. ‘vieux’ (‘ancien’) : cycles de concaténation et de composition. Grilles rythmiques et interactions entre constituants syllabiques.

Cette évolution se laisse décrire à l'aide des fonctions suivantes, en (5) *infra* : préfixation adjectivale (PFX\_ADJ) et borne senestre à voix soufflée (BORNE SENSTR & VX SOUFFL) dans un premier temps, suivi optionnellement d'un nouveau cycle de préfixation compositionnelle, qui décline l'ancienne racine lexicale déadjectivée en un spécifieur associé



à une nouvelle racine *nka*<sup>28</sup> (PFX COMPOS X > Y) – une autre analyse est possible, selon laquelle *nka* serait complémentateur et donc de statut Y, contre *ʃa* préservant son statut de racine lexicale, de rang X, mais cette hypothèse est moins cohérente avec la typologie de la langue, qui privilégie les formations de type SPEC+Racine polyvalente.

(5) Paramètres de CG-2 :

PFX\_ADJ & BORNE SENSTR & VX SOUFFL

PFX COMPOS X > Y

Le tableau 3 montre l'éventail de données accessibles dans la base de données ALMaz/Kirk66 (issue du traitement cartographique des listes de cognats de Kirk 1966) pour trois items, représentatifs des phases 3 et 4 (CG-3 & 4) du modèle évolutif de Sarah Gudschinsky.

Ex.	CG-3	CG-3	CG-4
	'écureuil'	'épis de maïs'	'marmite, pot'
MZ	<i>ʃahnũ</i>	<i>nāhnā</i>	<i>tihi</i>
AY	<i>ʃaʃnũ</i>	<i>nāhnā</i>	<i>tihi</i>
CQ	<i>ʃnũ</i>	<i>nāhnā</i>	<i>tihi</i>
JA	<i>ʃatuʔia</i>	<i>nāhnā</i>	<i>ndhi</i>
DO	<i>ʃatuʔia</i>	<i>nāhnā</i>	<i>tihi</i>
HU	<i>ʃahnũ</i>	<i>nāhnā</i>	<i>ti</i>
JI	<i>ʃahnũ</i>	<i>nāhnā</i>	<i>ti</i>
IX	<i>ʃihnũ</i>	<i>nāhnā</i>	<i>tihi</i>
MG	<i>ʃihnũ</i>	<i>nāhnā</i>	<i>tihi</i>
LO	<i>ʃijohnĩ</i>	<i>nōhnō</i>	<i>tehe</i>
TE	<i>ʃihnũ</i>	<i>nāhnā</i>	<i>tihi</i>

Tableau 3. Extrait de la base de données Kirk 1966, pour trois items (ex. CG-3 & 4)

CG-3 : Eventail gabaritique

CG-3 est identifiable au groupe consonantique interne aux lexèmes, qui trahit une composition, comme dans *nahna*, *ʃahnu*. L'extrait de la base de données ALMaz/Kirk 1966 suivant en suggère la gamme observable à l'échelle du diasystème : “*ʃa*l/\**ʃa*” pref. pers. ‘human’ (person prefix.) ; ‘humano’ (person prefix.) ◊ mz *ʃahnũ* n. ‘squirrel’, ‘ardilla’, ay *ʃanāji* n. ‘devil’, ‘diablo’, cq *ʃa* (male or female speaker with male referent), ja *ʃa*l (female speaker with male referent), do *ʃa*, hu *ʃa* (followed by person's name), ji *ʃahnũ* n. ‘squirrel’, ‘ardilla’, ix *ʃa* (female speaker with male referent), mg *ʃihnũ* n. ‘squirrel’, ‘ardilla’, lo *ʃijohnĩ* n.”

Comme l'indiquent les figures 6.1 et 6.2 infra, à cette phase sont associés les paramètres de borne senestre de racine forte (voix soufflée : BORNE VOIX SOUFFLEE X), la réduction du noyau de la syllabe du spécifieur (RED NOYAU Y) et l'expansion soit du spécifieur, soit de la racine (EXPANS Y | X).

(6) Paramètres de CG-3 :

BORNE VOIX SOUFFLEE X

RED NOYAU Y

EXPANS Y | X

<sup>28</sup> Racine de rang **lenis**, qui forme une paire minimale avec la racine **fortis** \**hnka* ↓ (à ton bas) \**hnka* ↓ (à ton moyen) « aile ».

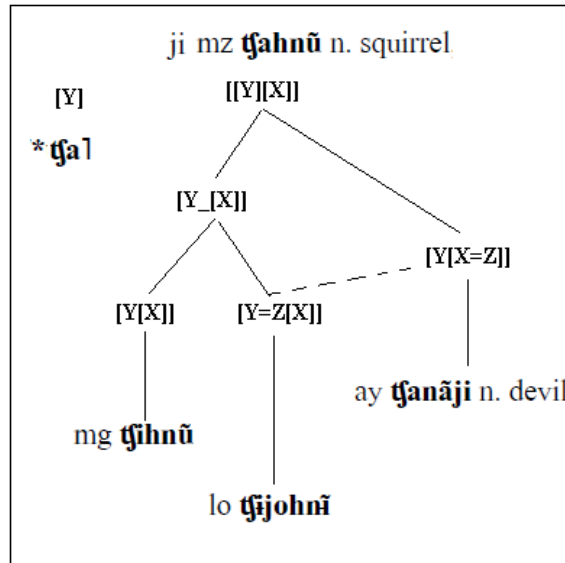


Figure 6.1. CG-3, composition et concaténation.

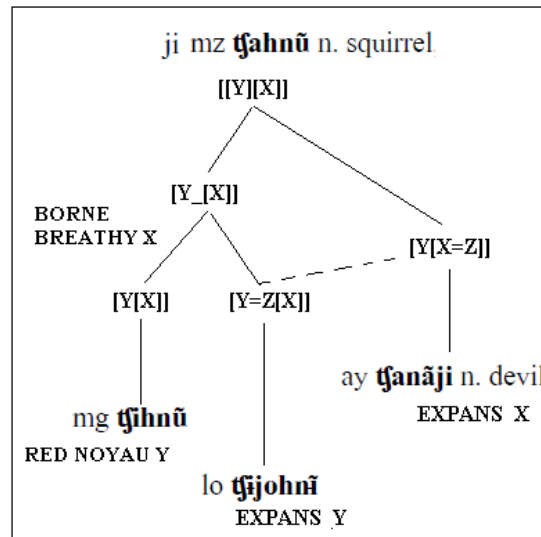


Figure 6.2. CG-3, paramètres.

*CG-4 : résorption  $CV^{T1}CV^{T2} > CV^{T1-2}$  par contour tonal :*

Le CG-4 rappelle l'allongement compensatoire, dans la mesure où des matériaux évacués du gabarit reportent une partie de leur substance sur un site préservé, de la droite vers la gauche du lexème – à la différence près que le mazatec ignore toute corrélation de durée, et n'a donc ni de consonne ni de voyelle longue. Les contours tonaux ancrent plusieurs morèmes (jusqu'à trois, selon les dialectes, mais le plus souvent seulement deux) sur une seule position squelettale. La résorption tonale par contour secondaire porte donc sur un seul segment au lieu de deux auparavant, et relève bel et bien de la compression gabaritique. Le cognat *\*ti|hi|/ \*nti|hi|* 'marmite, pot' offre un exemple particulièrement explicite de cette phase de réaménagement des gabarits dissyllabiques dans le diasystème.

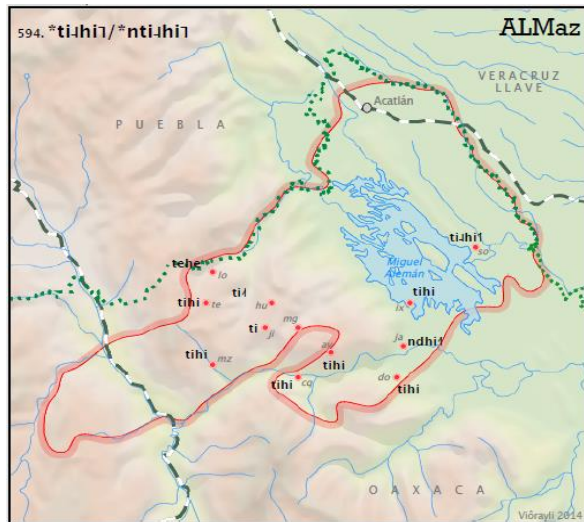


Figure 7. CG-4, cartographie de \**ntihí* ‘pot, marmite’

Ja : *ndiú* ↔ tons |MH| & Voix modale  
 Hu : *tîi* ↔ tons |Mh| & Voix modale  
 So : *tîhîi* ↔ tons |M-hH| & Voix soufflée

Les paramètres actifs sont le gouvernement tautosyllabique, au sein des noyaux syllabiques *fortis*, ou réarticulés (GVT VHV), autrement dit, par syncope du premier des deux noyaux vocoïdes, aboutissant à une qualité de voix soufflée :  $VhV > {}^hV \Leftrightarrow \check{V}$  (GVT VHV & VOIX SOUFFLEE), et l’émergence d’un contour secondaire au lieu d’un ancrage sur deux positions nucléaires au sein d’un noyau complexe (CONTOUR TON V | VHV), comme en (7). Les figures 7 et 8 présentent respectivement, d’une part la carte géolinguistique à partir de la base de données ALMaz/Kirk 1966, et d’autre part les options constructionnelles, sachant qu’ici la base *ntihi* ‘pot, marmite’ peut s’analyser comme *n=tihi*, où *n=* est un déterminant putatif (donc spécifieur Y) – démotivé –, tandis que *tihi* est la racine (X)<sup>29</sup>.

(7) Paramètres de CG-4:  
 CONTOUR TON V | VHV

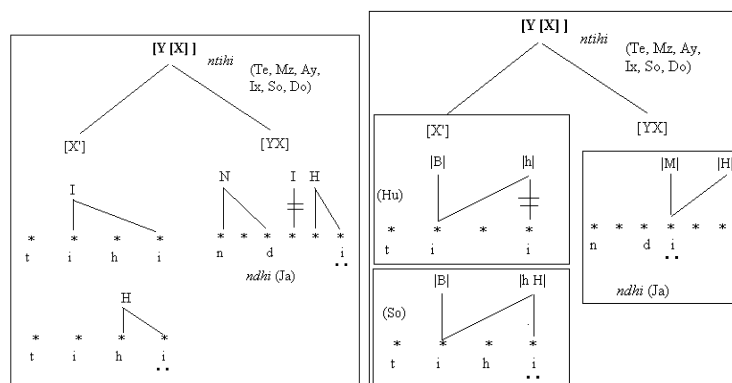


Figure 8. CG-4. Représentation des options structurales, morphologie et phonologie<sup>30</sup>.

<sup>29</sup> Pour des proclises devenues organiques du même genre dans une langue affiliée, voir Longacre (1957 : 54-55).

<sup>30</sup> Ici, une primitive phonologique comme {H} est notée H (pour la *breathiness*, ou voix soufflée), mais un ton est noté entre barres verticales : {H} = *breathiness*, mais |H| = ton haut.

CG-5 : formes encore analysables comme des mots composés dans certaines variétés et comme formes indécomposables dans d'autres

Une fois que le système a produit de manière endémique (autrement dit, ventilé d'une variété à l'autre) son cycle de formes compressées, qui coexistent de manière tout aussi endémique avec des pans de lexique plus rétenteur des structures gabaritiques étendues ou complexes, le diasystème émergent s'avère composite. Il s'accommode de cette hétérogénéité : à son échelle propre, le diasystème est polymorphique. C'est cet état composite que représente la phase 5 des cycles de Gudschinsky. L'auteure prend pour exemple l'adverbe signifiant 'doucement' : \***hwi**l **adv.** 'slowly'; 'despacio'  $\diamond$  mz **tahwi**, ay **tahwita**hwi, cq **tahwita**hwi, ja **hwi**l**hwi**l, do **tuhwita**hwi, ji **tahwi**, so **tifi**, mg **tahwi**, lo **fifi**, te **tuhwi**, hu **tu**-**hwi**l, hu **tofi**, **tui**.

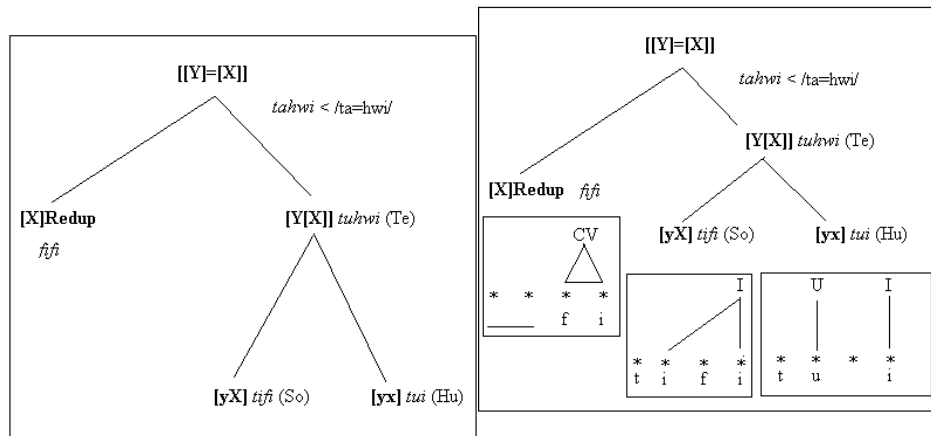


Figure 9. CG-5, morphologie et phonologie des aboutissements de \*to=hwi 'lentement'.

A ce stade, des processus additionnels de réduction segmentale interviennent, schématisés dans la figure 9 et déclarés en (8) : fusion mélodique (FUSION\_MELOD) ex. *tufi* > *tifi*, préservation de borne de racine ou de tête gabaritique (PRES\_BORNE) : *tu=hwi* > *tuh<sup>h</sup>phi*, reduplication senestre de tête gabaritique (REDUPL SENSTR) *fifi* > *fifi*.

(8) Paramètres de CG-5 :

FUSION\_MELOD  
PRES\_BORNE  
REDUPL SENSTR

En (9) figurent les principaux paramètres actifs dans le diasystème, selon la modélisation implicite de Sarah Guschinsky, réinterprétée à la lumière de la phonologie déclarative, à l'aide de paramètres *ad hoc*, qui trouvent leur cohérence dans des variables d'alignement (paramètres de bornes), de concaténation et/ou d'indicialité (affixes ou clitiques fusionnés dans le lexème), de réduction ou au contraire de renforcement ou expansion de structure interne ou de gabarit (variables RED vs. EXPANS) ou de copie (REDUPL).

(9) Récapitulation des principaux paramètres :

**Paramètres de CG-2 :**

PFX\_ADJ & BORNE SENSTR VX SOUFFL  
PFX COMPOS X > Y

**Paramètres de CG-3 :**

BORNE VX SOUFFLE X  
RED NOYAU Y

EXPANS Y | X

**Paramètres de CG-4:**

GVT VHV & VX SOUFFLEE

CONTOUR TON HV | VHV

**Paramètres de CG-5:**

FUSION\_MELOD

PRES\_BORNE

SENSTR

Ces paramètres ne sont pas nécessairement propres aux phases ou périodes postulées par l’auteure, qui a émis son hypothèse des cinq cycles de compression-unification lors de ses travaux pionniers sur la variation dialectale du mazatec, avec peu de données en main comparativement à la collecte de Paul Kirk (*op. cit.*), mais sa modélisation « en prose » – non formalisée – pose les jalons du modèle déclaratif que je viens d’exposer. Elle a le mérite de définir de manière analytique les options structurales qui se présentent dans le diasystème, et de les ordonner ou les hiérarchiser sur la ligne du temps (perspective chronologique).

On voit donc que ce qui opère ici, au sein du réseau dialectal, c’est davantage un jeu d’options et d’interactions entre solutions ou choix structuraux, entre différents parlars, qui sont autant de « décideurs » ou de « négociateurs », si l’on veut s’engager sur le sentier des métaphores socio-cognitives. Nous venons d’examiner en détail la texture structurale des réseaux d’options s’offrant pour la formation du lexique dans le diasystème mazatec. Ce faisant, nous avons davantage envisagé la variation du point de vue de systèmes de contraintes hiérarchisées (ou *paramètres*) et en relation d’implication que sous forme de lois phonétiques néogrammairiennes. Une telle démarche est résolument qualitative. Nous allons maintenant aborder la variation d’un point de vue strictement quantitatif, à l’aide d’une base de données morphologiquement unitaire, relevant du domaine structural que nous venons d’observer. Nous nous limiterons aux cognats substantivaux, afin de porter notre attention sur un sous-système homogène et prédictible, du moins davantage prédictible que le sous-système de la flexion verbale, que nous avons abondamment traité dans d’autres travaux, de nature structurale et qualitative (Léonard & Kihm 2010, 2012, Léonard & Fulcrand 2015, 2018, s.p.), et que nous aborderons ici-même dans la section 5, *infra*.

#### **4. Dialectométrie : cognats substantivaux de Kirk (1966)**

Un modèle à articulation logique, quasiment mécaniste, avait déjà été proposé par Gudschinsky au début de ses recherches sur le mazatec. Sarah Gudschinsky proposait déjà, peu avant son manuscrit de 1956, sur la base des données dont elle disposait alors – ses enquêtes ne portaient alors que sur six dialectes mazatecs, au lieu des douze sur lesquels a pu enquêter Paul Livingston Kirk, ou la trentaine de variétés que nous avons abordées entre 2010 et 2014 dans le cadre des enquêtes de l’ALMaz –, un scénario de fragmentation du réseau dialectal mazatec en cinq étapes (indiquées par  $T_n$  pour « Temps » *infra*) de diversification du mazatec (Gudschinsky 1955) :

( $T_1$ ) Homogénéité, suivie par l’émergence du dialecte de Huautla d’une part, dans les hautes terres, et de Jalapa d’autre part, dans les basses terres.

( $T_2$ ) Emergence d’une zone de transition entre les deux dialectes,

(T<sub>3a</sub>) Le dialecte des basses terres se divise en deux, faisant apparaître la variété de San Pedro Ixcatlán (Ix) ;

(T<sub>3b</sub>) Les aires de Huautla (Hu) et de Ix se diversifient : apparaît le sous-dialecte de San Mateo dans les hautes terres, tandis que la variété de San Miguel Soyaltepec (So) se sépare de Ix dans les basses terres. Dans la zone de transition, San Miguel Huautepéc (Mg) fait son apparition. Des flux de variables phonologiques et grammaticales et de lexique continuent de passer des basses terres vers les hautes terres.

(T<sub>4</sub>) Dans les basses terres, la distance entre Ix (piémont) et So (basses terres orientales) s'accroît.

(T<sub>5</sub>) La séparation entre les six dialectes en présence s'intensifie.

Dans un article ultérieur (Gudschinsky, 1958), l'auteure confirmera ce scénario, mais en y ajoutant des considérations ethnohistoriques liées aux phases d'hégémonie mixtèque et aztèque dans la région, que je ne retiendrai pas ici<sup>31</sup>. La carte schématique de la figure 2 *supra* fait clairement apparaître la tripartition entre hautes, moyennes et basses terres (ces dernières, en partie englouties à date récente par la retenue d'eau issue de la construction d'un barrage hydroélectrique, commencé en 1944 et achevé en 1954, qui a causé le déplacement de plus de 20 000 paysans). A ces trois plans ou segments du versant montagneux s'ajoute le Canyon de Cuicatlán et son réseau intriqué, nettement visible dans la partie sud de la zone mazatèque, qui est un facteur de diversification interne de plus – en outre, un coude est formé au sud de l'aire par l'intrusion d'une enclave mixtèque, celle de la variété de San Juan Coatzóspan<sup>32</sup>.

La matrice 1 ci-dessous présente les résultats du traitement de 311 cognats issus de Kirk (1966). L'équipe qui a réalisé ces mesures (Léonard *et al.*, 2016) dispose de résultats variés, y compris d'un calcul de distance de Levenshtein opérant sur un corpus plus diversifié en termes de parties du discours (noms, verbes, adverbes, etc.). Mais ces résultats sur une liste de cognats substantivaux présente l'avantage de correspondre aux phénomènes décrits précédemment ici-même à l'aide des cycles de Gudschinsky, et présentent un haut degré de cohérence structurale, pour ne provenir, précisément, que d'une composante du système linguistique et, partant, du diasystème – à partir de la même base de données.

	AY	CQ	DO	HU	IX	JA	AS/JI	LO	MG	MZ	SO	TE
AY	0,000	0,632	0,629	0,668	0,606	0,607	0,636	0,981	0,562	0,573	0,582	0,708
CQ	0,632	0,000	0,717	0,703	0,666	0,704	0,589	0,978	0,627	0,645	0,636	0,688
DO	0,629	0,717	0,000	0,689	0,585	0,334	0,643	1,000	0,608	0,639	0,620	0,703
HU	0,668	0,703	0,689	0,000	0,593	0,655	0,346	0,897	0,402	0,481	0,519	0,550
IX	0,606	0,666	0,585	0,593	0,000	0,599	0,616	0,937	0,574	0,639	0,519	0,586

<sup>31</sup> Ces détails sur les hégémonies successives dans cette région de Mésio-Amérique ne sont pas vraiment nécessaires ici, d'autant plus que les faits ne sont pas suffisamment documentés à ce jour sur le plan archéologique et ethnohistorique, alors que nous disposons d'excellentes données dialectologiques : celles de Paul Livingston Kirk (1966). Notre tentative d'analyse dialectométrique contribue à pallier à la carence de sources externes, concernant les configurations ethnolinguistiques dans le temps et dans l'espace, pour l'aire mazatèque.

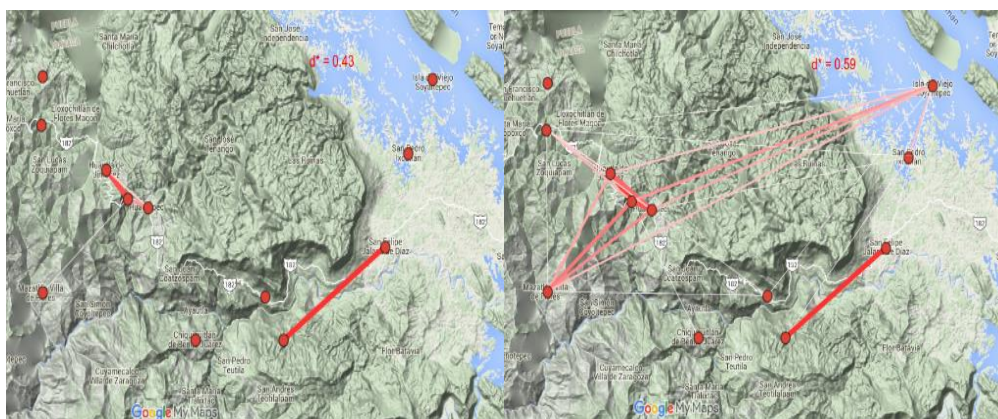
<sup>32</sup> Lors d'une enquête de terrain dans cette enclave, en août 2014, j'ai pu constater que deux voies reliaient cette micro-région des hautes terres à d'autres localités de langue mazatèque des basses terres : d'une part, à Chiquihuitlán, par un pont sous forme de filet suspendu, par un sentier de terre, d'autre part, à Jalapa de Díaz, par un chemin de muletier, qui passait autrefois derrière l'église de San Juan Coatzóspan. Ceci dit, étant donné les relations longtemps très conflictuelles entre Mazatecs de Huautla et Mixtecs à cette hauteur de la topographie de la région, pour des raisons de délimitations de terres exploitables, les liens intercommunautaires par cette enclave ont dû être restreints.

<b>JA</b>	0,607	0,704	0,334	0,655	0,599	0,000	0,617	0,945	0,594	0,604	0,585	0,675
<b>AS/JI</b>	0,636	0,589	0,643	0,346	0,616	0,617	0,000	0,841	0,377	0,426	0,462	0,502
<b>LO</b>	0,981	0,978	1,000	0,897	0,937	0,945	0,841	0,000	0,883	0,892	0,884	0,870
<b>MG</b>	0,562	0,627	0,608	0,402	0,574	0,594	0,377	0,883	0,000	0,446	0,490	0,539
<b>MZ</b>	0,573	0,645	0,639	0,481	0,639	0,604	0,426	0,892	0,446	0,000	0,511	0,567
<b>SO</b>	0,582	0,636	0,620	0,519	0,519	0,585	0,462	0,884	0,490	0,511	0,000	0,574
<b>TE</b>	0,708	0,688	0,703	0,550	0,586	0,675	0,502	0,870	0,539	0,567	0,574	0,000

Matrice 1. Matrice de distances de Levenshtein (LD), données Kirk (1966) : substantifs.  
Traitement algorithmique : v. Léonard *et al.*, 2016.

Nous allons maintenant retenir trois valeurs de différenciation : basse, moyenne et haute : nous retiendrons les agrégats diasystémiques étagés sur les valeurs 0,43 vs. 0,59 vs. 0,72. La première valeur (0,43) donne les dialectes les plus denses, en termes de similarité, puisque leur taux de différenciation est bas – il s’agit des variétés les plus résolument consensuelles, sur le plan structural. La deuxième valeur (0,59) laisse entrevoir le noyau dur des relations d’affinités au cœur du diasystème, faisant apparaître un treillis dense de traits partagés dans la longue durée, tandis que la troisième valeur (0,72) montre l’intrication de l’ensemble du diasystème.

Or, le résultat (moyennes normalisées de différenciation des douze localités) est à la fois congruent avec le scénario proposé par Sarah Gudschinsky il y a plus de 60 ans (Gudschinsky 1955, cf. *supra*), et bien plus précis (cf. les trois cartes de la figure 10) : l’indice bas (0,43) confirme la robustesse de deux dialectes : Huautla et les villages proches (Santa Maria Jiotes et San Miguel Huautepéc) dans les hautes terres, ainsi qu’un lien ténu mais positif avec Mazatlán Villa de Flores, dans la cordillère occidentale. A l’est, dans les basses terres, Jalapa de Diaz (JD) forme une paire avec Santo Domingo, le long de la rivière éponyme. Avec l’indice moyen (0,59), un treillis apparaît, qui relie les périphéries des basses terres orientales (SPI, SMS) avec celles des hautes terres occidentales (Mazatlán), en passant par le centre des hautes terres (Huautla), dont la trame se diversifie, en agrégeant le dialecte de San Jerónimo Tecoaatl (SJT), tandis que seules les variétés les plus périphériques, comme San Lorenzo et Chiquihuitlán, apparaissent comme des îlots non connectés (des singletons, comme San Lorenzo), ou de manière très sectorielle (Chiquihuitlán). Enfin, avec la valeur la plus élevée (0,72), l’ensemble du réseau se trouve connecté, hormis San Lorenzo, au nord-ouest, qui reste un indéfectible singleton, tandis que plusieurs treillis désormais connectent les périphéries du domaine – y compris Te, dont l’affinité lointaine, mais toutefois robuste à ce niveau de différenciation, avec les trois « town dialects » des basses terres que sont Ja, Ix et So, apparaît nettement.



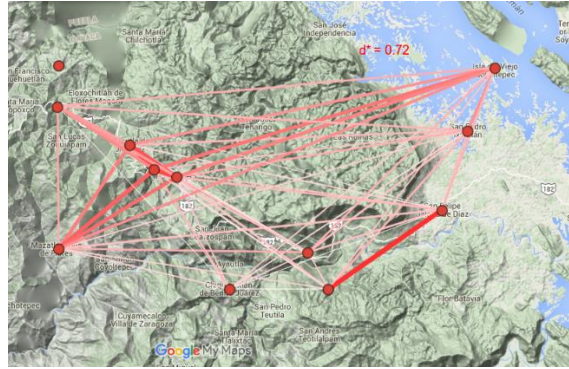


Figure 10. Trois saisies (ou *snapshots*) de distance de Levenshtein (0,43 vs. 0,59 vs. 0,72), substantifs, mazatec, données de Kirk (1966).

La coexistence de plusieurs centres directeurs (Huautla [Hu], Mazatlán Villa de Flores [Mz], San Pedro Ixcatlán [Ix], San Felipe Jalapa de Díaz [Ja] et San Miguel Soyaltepec [So]) crée un champ de force multipolaire (*polycentrisme*), et des innovations endocentriques (qui se diffusent en cascades, selon le modèle de Trudgill mentionné par Labov, au sein de ces microcosmes que constituent le parler urbain et le parler rural, voire de multiples sous-variétés rurales environnantes). Ces relais forts, de proche en proche, créent des conditions d'itérativité (ou de copie, avec d'éventuels remaniements) des innovations, et favorisent l'intrication (ou imbrication) des variables – dans le cas du mazatec et de sa flexion verbale, cette hiérarchisation à plusieurs niveaux de stratification –régiolectal et locolectal –, favorise la *multiréfraction*<sup>33</sup> (ou diversification par exaptation<sup>34</sup>) du système de classes flexionnelles verbales, par exemple : les systèmes locaux innovent, souvent en réinterprétant des modèles de paradigmes en vigueur dans les dialectes urbains voisins, comme nous allons le voir dans la prochaine section.

L'*intégration diasystémique* se fait aux niveaux *régiolectal* des hautes, moyennes et basses terres, mais aussi par contact de proche en proche tout le long du continuum dialectal. Le maillage est compact, car hormis les variétés périphériques du nord (Lo) et du sud (Cq), l'intrication de toutes les autres variétés, est forte – ce qui augmente la complexité du système, puisqu'il est fortement interactif, tous azimuts, au lieu d'être orienté par des canaux ou des branches capillaires, et unifié par des bassins en cascade. Là encore, nous sommes face à une narration alternative de l'histoire des aires dialectales du mazatec et, partant, des communautés humaines qui les portent (ou les diffusent ou les imbriquent, entre hasard et nécessité) : au-delà d'un « système-monde », c'est tout un monde de narrations possibles, quoique cohérentes ou mutuellement compatibles, qui émerge des différents points de vues retenus pour l'analyse. Derrière cette dynamique permanente, ou ce mouvement perpétuel, c'est la trame et la causalité des activités humaines à travers l'histoire qui se reflètent dans les données et dans les modèles qu'on peut tirer de la nature et de l'organisation de ces données, en fonction des principes universels et des paramètres locaux d'association de schèmes et de structures (gabarits et contraintes sur les gabarits, constituance morpholexicale, etc.).

## 5. Morphologie à grain fin

### 5.1. Éléments de morphologie verbale mazatèque

<sup>33</sup> J'entends par là une multitude de reflets d'un même sous-système ou d'une même structure, avec de légères variantes, selon une répartition endémique dans l'espace social ou géographique.

<sup>34</sup> L'*exaptation* diffère de l'*adaptation* en ce sens que ce n'est plus tant la fonction qui crée l'organe, mais l'organe qui crée la fonction, comme les plumes dans le règne animal, initialement déterminées par l'adaptation aux conditions thermiques chez certains reptiles, qui ont fini par servir au vol, chez ceux qui allaient devenir des oiseaux. La *métatypie*, en morphologie, rappelle le mécanisme d'exaptation.



Le tableau 4 ci-dessous présente les principaux paradigmes flexionnels (CF, ou Classes Flexionnelles) ainsi que les classes de ton de racine (CTR) dans la variété centrale des hautes terres (Huautla de Jiménez), en fonction des préverbes (PV) TAMV/Personne, pour le triplet opposant l'accord subjectal (ACC.S) aux trois personnes que sont la troisième singulier et pluriel (3), la première personne du singulier (1SG) et la première personne du pluriel inclusive (nous tous : 1 PL.INCL).

Le choix de ces trois paradigmes est motivé par l'efficacité de ses contrastes : la forme 3 donne le lemme, et ne subit aucune modification (ni affixation ni cliticisation) ; la forme 1 SG donne celle à contour tonal sur le préverbe, avec affixe *-a* toujours coalescent avec la voyelle de la racine lexicale ou voyelle finale et donnant lieu à un léger contour – noté par un doublement de la voyelle<sup>35</sup>) ; la forme 1 PL.INCL donne une forme également coalescente, mais avec suffixe *-à*, de ton mi-haut (h), ce qui permet d'observer des contrastes tonals dans la partie dextre des radicaux. En résumé, la cellule 3 donne le lemme, la cellule 1 SG deux contours tonaux, préverbal et suffixal, tandis que la cellule 1 PL. INCL permet d'observer à la fois les conditions de réduction du ton de préverbe sous contrainte OCP (*Obligatory Contour Principle*, cf. McCarthy 1986) et, surtout, un autre contour suffixal, dans le domaine dextre du radical. Enfin, les tiroirs TAMV retenus sont le neutre (équivalent du présent générique ou intemporel, ou de l'habituel), l'accompli (ACMPL) valant pour un prétérit, et l'inaccompli (INCMPL), valant principalement pour le futur (autrement dit, inaccompli irréel).

Le principal critère qui motive les subdivisions en sous-classes morphologiques sont les ajustements à la frontière entre proclitiques ou préfixes TAMV à l'accompli et à l'inaccompli.

Ces phénomènes de joncture sont de deux sortes, et concernent uniquement la classe des préverbes à attaque labiale (*b-*, issu de /w/, par consonantisation d'une approximante en position initiale, donc forte), comme suit :

- (a) la *sape* ou l'amuïssement de l'attaque ;
- (b) 3 NEUTRE *b'éxá* 'il/elle envoie' vs. ACMPL *tsak'éxá* 'il/elle envoya' ;
- (c) la *charge* de sonorité, ou vocalisation de l'attaque : 3 NEUTRE *b'éxá* vs. *k'òèxá* 'il/elle enverra'.

Ce phénomène semble panmazatec, bien que la glottalisation résultant de l'ajustement de joncture soit le plus souvent affaiblie dans les idiolectes modernes que nous avons pu observer sur le terrain.

PV	CF	CTR	ACC.S	NEUTRE	ACMPL	INCMPL	
<i>b'é-</i>	I.A	1/H/	3	<i>b'éxá</i>	<i>tsak'éxá</i>	<i>k'òèxá</i>	'envoyer'
			1SG	<i>b'eexáa</i>	<i>tsak'eexáa</i>	<i>k'òexáa</i>	
			1PL.INCL	<i>b'èxá</i>	<i>tsak'èxá</i>	<i>k'oèxá</i>	
<i>ba-</i>	I.B	3 /H/	3	<i>bané</i>	<i>tsakané</i>	<i>koanè</i>	'laver'
			1SG	<i>banée</i>	<i>tsakanee</i>	<i>koanee</i>	
			1PL.INCL	<i>banè</i>	<i>tsakanè</i>	<i>koanè</i>	
<i>bi-</i>	I.C 1 /LM/	3	3	<i>biyáa</i>	<i>tsakiyáa</i>	<i>koiyá</i>	'mourir'
			1SG	<i>biyáa</i>	<i>tsakiyáa</i>	<i>koiyáa</i>	
			1PL.INCL	<i>biyáà</i>	<i>tsakiyáà</i>	<i>koiyá</i>	
<i>ba-/bi-</i>	II. C2	1 /M/	3	<i>bate</i>	<i>tsakate</i>	<i>koate</i>	'casser'

<sup>35</sup> Comme déjà signalé *supra*, il n'y a pas de corrélation de longueur vocalique en mazatec, qui ne connaît que des voyelles brèves. La gémination vocalique à l'écrit est donc une pure convention graphémique. Nous la préférons de loin à la numération utilisée dans nombre de travaux de tonologie (tons énumérés de 1 à 4 pour de H à B, et contours 43 pour BM, 42 pour Bh, etc. comme dans Pike 1948 ; à mon humble avis, ces notations sont, à termes, illisibles, pour identifier des *patterns* discrets, objectif de tout phonologue).

			1SG	<i>bate</i>	<i>tsakate</i>	<i>koate</i>	
			1PL.INCL	<i>bichà</i>	<i>tsakichà</i>	<i>koichà</i>	
<i>bá-/fa-</i>	II.	3 /M/	3	<i>bájtsa</i>	<i>tsakájtsa</i>	<i>kòàjtsa</i>	‘mettre à cuire’
	C3		1SG	<i>b<sub>a</sub>ajtsà</i>	<i>tsakaajtsà</i>	<i>kòajtsàa</i>	
				<i>a</i>	<i>a</i>		
			1PL.INCL	<i>fajtsaà</i>	<i>tsakjajtsaà</i>	<i>kjoajtsaà</i>	
<i>b'a-/ch'a</i>	III	2 /H/	3	<i>b'akjá</i>	<i>tsak'akjá</i>	<i>k'oakjá</i>	‘porter (un vêtement)’
-			1SG	<i>b'akjáa</i>	<i>tsak'akjáa</i>	<i>k'oakjáa</i>	
			1PL.INCL	<i>ch'akjá</i>	<i>kich'akjá</i>	<i>ch'akjá</i>	
<i>sí-/nì-</i>	IV	5 /h/	3	<i>sítsjò</i>	<i>kisítsjò</i>	<i>sjítsjò</i>	‘griller’
			1SG	<i>sjítsjòa</i>	<i>kisjítsjòa</i>	<i>sjítsjòa</i>	
			1PL.INCL	<i>nítsjòà</i>	<i>kinítsjòà</i>	<i>sjítsjòà</i>	
<i>tsò- / mi-</i>	V	3 /M/	3	<i>tsòya</i>	<i>kitsòya</i>	<i>k'uitsòya</i>	‘enseigner’
			1SG	<i>tsòyaa</i>	<i>kixinyaa</i>	<i>xjinyaa</i>	
			1PL.INCL	<i>'miyà</i>	<i>tsak'inyà</i>	<i>k'uinyà</i>	
<i>kjo- / chjo-</i>	VI	3 /M/	3	<i>kjoya</i>	<i>kiskoya</i>	<i>skòya</i>	‘cuisiner’
			1SG	<i>kjoyaa</i>	<i>kiskoyaa</i>	<i>skoyaa</i>	
			1PL.INCL	<i>chjoyà</i>	<i>kichjoyà</i>	<i>chjoyà</i>	
<i>see- / jnta-</i>	VII	5 /BM/	3	<i>see</i>	<i>kisee</i>	<i>sé</i>	‘chanter’
			1SG	<i>see</i>	<i>kisee</i>	<i>kisee</i>	
			1PL.INCL	<i>jntaà</i>	<i>kijntaà</i>	<i>koijntá</i>	

Tableau 4. Huautla de Jiménez, CTR & CF. Données de Pike (1948) & enquêtes JLL 2011.

Le tableau 5 reprend un fragment du précédent, afin de mettre en évidence par la segmentation les procédés de formation (notés par +, qui associe les préverbes aux racines polyvalentes) et de flexion, notés par un tiret dans les cas d’affixation, et par un tiret double pour les clitiques : CFC I.a 3 NEUTRE *b'è+xá* ‘il/elle envoie’ vs. ACMPL *tsa=k'-é+xá* ‘il/elle envoya’ (avec effacement de l’attaque initiale de préverbe), et INCMPL *k'-oè+xá* ‘il/elle enverra’, (avec vocalisation de l’attaque labiale de préverbe). Comparer à 1PL.INCL *b'è+xá* ← *|b'è+xá-à|* ‘nous envoyons’, ACMPL *tsa=k'-è+xá* ‘nous envoyâmes’ [*tsaka=b'è+xá-à* |], INCMPL *k'-oé+xá* ← *|ku=b'è+xá-à|* ‘nous enverrons’.

PV	CF	CTR	ACC.S	NEUTRE	ACMPL	INCMPL
<i>b'è+</i>	I.a	1/H/	3	<i>b'è+xá</i>	<i>tsa=k'-é+xá</i>	<i>k'-oè+xá</i>
			1SG	<i>b'èe+xáa</i>	<i>tsa=k'-ee+xáa</i>	<i>k'-oèe+xáa</i>
			1PL.INCL	<i>b'è+xá</i>	<i>tsa=k'-è+xá</i>	<i>k'-oé+xá</i>
<i>ba-</i>	I.b	3 /H/	3	<i>ba+né</i>	<i>tsa=k-a+né</i>	<i>k-oa+né</i>
			1SG	<i>ba+née</i>	<i>tsa=k-a+née</i>	<i>k-oa+née</i>
			1PL.INCL	<i>ba+nè</i>	<i>tsa=k-a+nè</i>	<i>k-oa+nè</i>
<i>bi-</i>	I.c/1	3 /LM/	3	<i>bi+yáa</i>	<i>tsa=k-i+yáa</i>	<i>k-oi+yá</i>
			1SG	<i>bi+yáa</i>	<i>tsa=k-i+yáa</i>	<i>k-oi+yáa</i>
			1PL.INCL	<i>bi+yáà</i>	<i>tsa=k-i+yáà</i>	<i>k-oi+yá</i>
<i>sí-/nì-</i>	IV	5 /h/	3	<i>sí+tsjò</i>	<i>ki=sí+tsjò</i>	<i>sjí+tsjò</i>
			1SG	<i>sjí+tsjò-a</i>	<i>ki=sjí+tsjò-a</i>	<i>sjí+tsjò-a</i>
			1PL.INCL	<i>ní+tsjo-à</i>	<i>ki=ní+tsjo-à</i>	<i>sjí+tsjo-à</i>

Tableau 5. Segmentation des principaux paradigmes du tableau 5.1.1 ‘envoyer’ ‘laver’ ‘mourir’ ‘griller’ (Huautla de Jiménez).

Dans la classe flexionnelle IV, celle du causatif, pour le verbe *sítsjò* ‘il/elle grille/torréfie’<sup>36</sup> les gloses seraient :

<sup>36</sup> En situation d’élicitation, le meilleur stimulus en espagnol était *él o ella tuesta café* « il ou elle torréfie du café ».

(10) Paradigme +3/1SG : Préverbation *sí*+/*sí*+

- (10a) *sí*+*tsjò*  
FAIRE+gril  
P<sub>VCF</sub>IV.3+√.3  
'il/elle grille/torréfie'
- (10b) *ki=sí*+*tsjò*  
ACMPL=IV.nonNTR.3+√.3  
'il/elle grilla/torréfia'
- (10c) *sí*+*tsjò*  
IV.nonNTR/INCMPL.3+√.3  
'il/elle grillera/torréfiera'

(11) Paradigme -3/1SG : Préverbation *nì*+/*sí*+

- (11a) *nì*+*tsjo-à*  
IV.non3/1SG+√-1PL.INCL  
'Nous grillons/torréfions'
- (11b) *ki=nì*+*tsjo-à*  
ACMPL=IV.non3/1SG+√-1PL.INCL  
'Nous grillâmes/torréfiâmes'
- (11c) *sí*+*tsjo-à*  
INCMPL.IV.non3/1SG+√-1PL.INCL  
'Nous grillerons/torréfierons'

Nous avons noté P<sub>VCF</sub>IV pour la première forme (soit préverbe de classes flexionnelles IV), puis nous avons allégé la notation à l'aide de IV tout simplement dans les formes successives.

Notation simplifiée :

- (11d) *ki=nìtsjo-à*  
ACMPL=griller-1PL.INCL  
'Nous grillâmes/torréfiâmes'

Le système de classes flexionnelles (CF) est modélisé dans le tableau 6 infra, qui subdivise les critères pour la définition des « conjugaisons mazatèques » en fonction de (i) dans la première colonne à gauche, les critères morphologiques (préverbes à attaques labiales, sensibles aux phénomènes de sape et de vocalisation, sous l'effet de la proclise de marqueurs TAMV, versus autres types d'attaques préverbales, coronales ou dorsales, non alternantes) : classes I à III ; (ii) critères d'allomorphie et de supplétion : à partir de la classe IIC2, des alternances ont lieu en fonction de l'accord sujet, selon la dichotomie qui oppose 1SG et 3 SG & Pl à toutes les autres personnes (voir aussi dans la dernière colonne à droite, le paramètre de

« scission »<sup>37</sup>); (iii) de manière secondaire ou incidente, on peut tenir compte des sèmes attachés aux préverbes ou verbes légers, qui sont recrutés principalement dans les paradigmes lexicaux des verbes de position ou de mouvement et trajectoire (CF I à II) ou de causation (CF IV en particulier, à partir de la forme *lenis* du verbe ‘faire’).

Attaque initiale	Classe		Préverbe	Sémantique	Scission { □ 1SG,3 }
Labiale	I	A	<i>b'é-</i>	LOCATION	Non scindé
		B	<i>ba-</i>	GENERIQUE	
		C 1	<i>bi-</i>		
	II	C 2	<i>ba- / bi-</i>	MOUVEMENT	
		C 3	<i>bá- / fa-</i>		
III		<i>b'a- / ch'a-</i>	SURFACE		
Coronale	IV		<i>sí- / ni-</i>	CAUSATIF	scindé
	V		<i>tsò- / mi-</i>	Variable (hors corrélation)	
Dorsale	VI		<i>kjo- / chjo- ko- / cho-</i>		
			Racine nue		
Ensemble ouvert	VII		<i>see- / jnta-</i>	PROCES SIMPLE	

Tableau 6. Taxinomie flexionnelle du verbe dans le segment diasystémique des hautes terres centrales (données enquêtes JLL 2010-14).

Ces CF sont l'équivalent des conjugaisons du latin ou des langues romanes (conjugaison en *-a*, en *-e* et en *-i* notamment, ex. italien *cantare*, *correre*, *partire*, espagnol *cantar*, *correr*, *partir*), à la différence que la plupart des dialectes mazatec ne comptent pas moins de sept classes de conjugaison (en chiffres romains dans les tableaux 5.2-3), et en moyenne une dizaine. Certaines variétés comptent jusqu'à près d'une vingtaine de conjugaisons de ce type (notamment So et Cq, cf. Jamieson 1982 pour ce dernier dialecte). Comment le diasystème déploie-t-il cette diversité de taxinomies flexionnelles, et comment rendre compte de cette complexité ? **5.2. San Felipe Jalapa de Díaz**

Le dialecte de San Felipe Jalapa de Díaz (Ja)<sup>38</sup>, dans le piémont au sud-ouest des hautes terres, est le principal et le plus proche “rival” de celui de Huautla de Jiménez (25,395 habitants à Ja en 2005, contre 31 550 à Hu). Son système de classes flexionnelles est résumé dans le tableau 7. Dans les grandes lignes, à se fier aux quatre premières colonnes, ce système paraît très semblable à celui de Huautla (Hu, cf. tableau 6 *supra*), avec juste quelques ajustements de préverbes (cellules Vet VI, VIIa b). Mais ce n'est là qu'une impression de surface. Le détail des paradigmes révèle des structures d'un grand intérêt du point de vue des dynamiques de continuité et de rupture avec d'autres secteurs du réseau dialectal mazatec. Les informations stratégiques, de ce point de vue, figurent dans la colonne des « procédés locaux » : la classe flexionnelle IA atteste un marquage de l'aspect accompli lointain (équivalent du prétérit) par une alternance de préverbe *y'é-* vs. *b'é-* pour la forme de l'aspect neutre (ou forme lemmatique),

<sup>37</sup> Cf. le célèbre article de C. Jamieson sur les séries sous-conflatives comme « fil rouge » de l'ensemble de la flexion verbale du mazatec, à partir du système de Cq (Jamieson 1982). Voir aussi la modélisation que nous en avons induit selon une perspective PFM dans Léonard & Kihm 2010.

<sup>38</sup> L'informateur ALMaz à San Felipe Jalapa de Díaz était Gilberto Martínez Álvarez, instituteur bilingue, âgé de 53 ans en 2013, formé à l'UPN de Tuxtepec (Universidad Pedagógica Nacional), auteur d'un livre d'école pour l'apprentissage de la lecture en mazatec de Jalapa. GAM n'a vécu que quelques années hors de Jalapa, notamment à Córdoba, dans l'Etat de Veracruz, et pour ses études à l'UPN. Les enregistrements ont eu lieu les 28 et 29-07-2013 au domicile de l'informateur (enquête par Jean Léo Léonard).

tandis que comme le montrait le tableau 5 *supra*, colonne 6, l'aspect accompli lointain se marque à Huautla (Hu) à l'aide d'une proclise tsak= enclenchant une sape de l'attaque labiale.

Attaque initiale	Classe		Préverbe	Procédés locaux	Items
Labiale	I	A	<i>b'é-</i>	CPL PV ( <i>y'é-</i> )	<i>b'éxá</i> 'envoyer'
		B	<i>ba-</i>	CPL ( <i>ts=</i> & b sape)	<i>bané</i> 'laver'
		C 1	<i>bi-</i>		<i>bixan</i> 'marier'
	II	C 2	<i>ba- / bi-</i>		<i>bak'antsank'a</i> 'courir'
	III		<i>b'a- / cha-</i>	CPL ( <i>kik=</i> & b sape) : Cmplx chaînes procl	<i>bajnu</i> 'rire'
Coronale	IV		<i>ts'í- / n'e-</i>	CPL ( <i>ki=</i> )	<i>ts'íntsju</i> 'griller'
	V		<i>tjú-</i>		<i>tjúntá</i> 'suer'
Dorsale	VI		<i>kja-</i>	SOUS-CONFLATION neutralisée	<i>kjaníya</i> 'jeter'
Ensemble ouvert	VIIa		<i>se-</i>		<i>se-</i> chanter
			Racine nue		
	VIIb		<i>fi- / mi- / ba-</i>	SOUS-CONFLATION brisée	<i>fi</i> 'aller'

Tableau 7. Taxinomie flexionnelle du verbe dans le segment diasystémique du piémont méridional (Ja) ; données enquêtes JLL 2010-13.

Ce trait dialectal, qu'on retrouve ailleurs dans les basses terres (So) mais aussi à la périphérie nord-occidentale des hautes terres (Lo), est noté CPL PV (*y'é-*) dans le tableau 7. Pour les sous-classes flexionnelles IB, IC1 et IIC2, on trouve le procédé consistant à marquer l'accompli lointain par une proclise *ts=* < *tsa=*, avec sape de l'attaque labiale du préverbe, phénomène recensé comme CPL (*ts=* & b sape) dans le tableau – là encore, traits des basses terres, partagé également avec Ix, en plus de So. Le marquage s'accompli continue de rompre avec les lemmes préverbaux pour la classe III, avec une proclise *kik=* associée également à la sape de l'attaque labiale préverbale – noté CPL (*kik=* & b sape) dans le tableau 7, ce qui revient à une complexification de la chaîne proclitique (noté Cmplx chaînes procl). On notera qu'il y a en outre croisement de *ki=* d'accompli lointain, solidement attesté pour la CF IV aussi bien à Hu qu'à Ja, avec un fragment de marquer aspectuel *k=*, qui n'est autre que le même déjà observé dans le proclitique *tsak=* à Hu, qui s'analyse simplement comme la combinaison de *tsa=* et de *ka=*, avec apocope. Or, *tsa=* est un morphème issu du complexe TAMV qu'on retrouve comme forme phonologiquement forte modale (conjonction 'si' hypothétique), ici affaiblie en *ts=* (à moins qu'il ne s'agisse plus simplement de *tsi=*, qui serait à proprement parler un marqueur d'accompli), et *ka=* est par ailleurs clairement attesté comme proclitique d'accompli proche. La CF VI diffère phonologiquement dans le détail de la CF correspondante à Hu, mais en outre, le mécanisme de sous-conflation qui opère à Hu n'apparaît pas à Ja, comme s'il y avait neutralisation de cette alternance (signalée par la mention « SOUS-CONFLATION neutralisée », dans le tableau). C'est aussi le cas pour une sous-série de la CF VII (numérotée VIIa), tandis que la sous-série VIIb, au contraire, ajoute une sous-série alternante de plus à la conflation – ce qui est appelé ici SOUS-CONFLATION brisée.

Il est d'une importance capitale de préciser que, si nous avons dû fonder la comparaison des deux dialectes sur la seule base des données recueillies par Paul L. Kirk dans sa liste de cognats – sa liste de plus de 850 items intègre pas moins de 140 entrées verbales –, nous n'y

aurions vu que du feu, et aurions jugé que les deux dialectes partageaient, à peu de choses près, les mêmes préverbes – ce qui n’est pas non plus faux en soi. Ce n’est qu’en enquêtant patiemment sur le système verbal, CF par CF et paradigme par paradigme (accord sujet et tiroirs TAMV) qu’apparaissent ces divergences, qui montrent deux choses d’un grand intérêt, du point de vue de la perspective qui est ici la nôtre : 1) la variété de Jalapa est résolument une variété des basses terres, qui converge de manière discontinue davantage avec So qu’avec sa proche voisine Ix (en cela, elle brise un « modèle de chaînes » qu’on pourrait préférer appliquer à l’observation de la variation géolinguistique, plutôt qu’un modèle arborescent) ; 2) le système de CF local de Ja est truffé de métatypies innovantes pour la flexion de l’accompli lointain : tout se passe comme si ce système « tricotait » sa différence avec Hu, en s’ingéniant à la fois à imiter (notamment par le recours à la sape de l’attaque labiale préverbale et les chaînes proclitiques, hors de la CF IA), et à inventer des solutions nouvelles, en croisant deux « compères » proclitiques : *tsa=* et *ki=*, le premier étant spécifique aux basses terres, le deuxième étant panmazatec. Mais en jouant de la sorte avec les procédés et les marqueurs disponibles dans l’inventaire diasystémique de ressources structurales, Ja ne fait pas que se démarquer de Hu, tout en convergeant fortement sur le plan architectonique du système de CF. Cette importante variété des terres moyennes, ou piémont, se démarque également, de manière plus conservatrice et plus convergente avec les hautes terres centrales (Hu, encore), avec son voisin innovant Ix – dont les mécanismes de métatypie ont été décrits dans Léonard & Fulcrand (s.p.).

### 5.3. Santa Cruz Acatepec

Santa Cruz Acatepec (désormais Ac)<sup>39</sup> relève du sous-dialecte Nord-occidental, variété périphérique du dialecte des hautes terres. Le bourg comptait pas plus de 1300 habitants en 2005. Autant Ja est un Goliath, avec ses 25,395 habitants en 2005, tout comme Hu (31 550 habitants, cf. *supra*), plus proche d’Ac, autant une bourgade comme Ac est un David, en comparaison. Nous avons entendu parler de ce lieu comme d’une zone difficile d’accès, marquée par des conflits de terres, et des problèmes de caciquisme. Nous avons eu la chance de pouvoir sonder le système des CF d’Ac lors d’un atelier de grammaire mazatèque qui s’est tenu dans un autre petit bourg des hautes terres, en 2012, en enregistrant un instituteur bilingue qui participait aux activités du groupe en formation.

Ce système flexionnel est très semblable à celui de Te (San Jerónimo Tecoaatl, principal bourg des hautes terres occidentales) ou de San Antonio Eloxochitlán ou encore, de San Lucas Zoquiapan. On en trouvera une description fine dans Bull (1984), avec une intéressante modilisation par cycles morphologiques du système flexionnel. Les différences avec le système de Hu sont plus importantes qu’il n’y paraît, et elles font converger cette variété par divers traits avec Ja, qui relève pourtant des moyennes et basses terres. D’abord, comme partout au nord-ouest des hautes terres, y compris dans le parler de Santa Maria Chilchotla, on observe une équivalence entre la CF IA de Hu et celle correspondant à IC1, à Ac : *bíxá* au lieu de *béxá* ‘envoyer’ (cette relation d’équivalence est indiquée par IA ⇔ IC1. Mieux encore, un verbe comme *bítsjò* ‘griller’ relève d’une catégorie hybride à A : sa flexion à l’aspect neutre présente un préverbe de CF IA, à attaque labiale (*bi-*), tandis que le tiroir de l’accompli lointain a un préverbe causatif qui le fait ressortir à la CF IV : *kì=sítsjò*. A l’inaccompli, sa flexion fait qu’il relève de nouveau de la CF IA : *kiitsjò* < *k=uitsjò*, avec assimilation par le noyau syllabique du

<sup>39</sup> L’informateur ALMaz à Santa Cruz Acatepec était Silvano Zaragoza Carrera, instituteur bilingue âgé de 31 ans, enregistré le 2 octobre 2012 au cours d’un atelier de grammaire mazatèque dans un collège de San Mateo Yolochochitlan (cf. <http://axe7.labex-efl.org/node/87> pour les données de cet atelier).

préverbe de l'attaque labiale ici vocalisée (cf. Bull, *op.cit.*). On comparera utilement avec l'entrée 'jouer', qui relève de la CF IV sans aucune hybridation, en préservant le préverbe causatif *si=* dans toutes les cellules de la conjugaison. Enfin, la forme lexicale de 'tomber' à Ac manifeste également un changement de CF, comparativement à Hu, où le verbe est *kaa* monosyllabique et donc relève de la CF VII (quoique sans scission) : ici, c'est une forme composée synthétique associant deux racines : *kàtjen* < *kà+tjen*, dont *tjen* vaut pour le sème de « du haut vers le bas » (comme le pléonasse 'tomber en bas'). En somme, ces quatre entrées attestent largement du jeu de différenciation que peut entretenir une variété du sous-dialecte nord-occidental des hautes terres aussi bien avec Ja, du piémont, qu'avec Hu, principal centre urbain des hautes terres centrales.

Il est intéressant de constater la relative robustesse des classes tonales de racines (ou CTR : troisième colonne du tableau 4 *supra*) entre les deux variétés des hautes terres (Ac et Hu), avec des aménagements qui restent fidèles aux paliers ou registres H et bas (par exemple maintien de H et h dans les racines de *bíxá* et de *bítsjò* à Ac), ou qui transposent les niveaux, comme le transfert B > h pour *kàtjen*, comparé à *kaa* à Hu, qui s'explique d'autant mieux que la première des deux racines lexicales est devenue spécifique de la deuxième – un phénomène analogue à des réagencements de constituance et de composition observés dans le domaine nominal dans la section 3.

		3	1 Sg.	2 Sg.	2 Pl	1 Pl.excl
<b>IA ⇔ IC1</b> <i>bíxá</i>	NTR	<i>bíxá</i>	<i>bixá</i>	<i>bixí</i>	<i>bixà</i>	<i>bixájìn</i>
'envoyer'	CPL	<i>tsikixà</i>	<i>tsikixà</i>	<i>tsikixí</i>	<i>tsikixà</i>	<i>tsikixájìn</i>
	INCPL	<i>kixà</i>	<i>kixà</i>	<i>kixí</i>	<i>kixà</i>	<i>kixájìn</i>
<b>IV ⇔ IA &amp; IV</b> <i>bítsjò</i>	NTR	<i>bítsjò</i>	<i>bítsja</i>	<i>bítsjí</i>	?	<i>bítsjòjìn</i>
'torréfier, griller'	CPL	<i>kìsítsjò</i>	<i>kisítsjá</i>	<i>tsikítsji</i>	<i>tsikítsjó</i>	<i>tsikítsòjìn</i>
	INCPL	<i>kiítsjò</i>	<i>kítsjaa</i>	<i>kítsjí</i>	?	<i>kítsjòjìn</i>
<b>IV. siská</b>	NTR	<i>siská</i>	<i>siská</i>	<i>siskí</i>	<i>siská</i>	<i>siskájìn</i>
'jouer'	CPL	<i>kìsiská</i>	<i>kisiská</i>	<i>kiniskí</i>	<i>kiniská</i>	<i>kiniskájìn</i>
	INCPL	<i>siská</i>	<i>siskáen</i>	<i>siskájì</i>	?	<i>siskájìn</i>
<b>VI. kàtjen</b>	NTR	<i>kàtjen</i>	<i>kàtjen</i>	<i>kàtjin</i>	<i>kátj'an</i>	<i>kátejìn</i>
'tomber, choir'	CPL	<i>kìskaàten</i>	<i>kiskàten</i>	<i>kiskàtj'in</i>	<i>kiskàt'an</i>	<i>kiskàtjejìn</i>
	INCPL	<i>skàtjen</i>	<i>skàtjen</i>	<i>skàtjin</i>	<i>skàtjan</i>	<i>skàtjejìn</i>

Tableau 8. Santa Cruz Acatepec: *bíxá* 'envoyer', *bi-/sítsjó* 'griller', *siská* 'jouer', *kàtjen* 'tomber'

Ce système est à la fois un modèle très évolué du modèle flexionnel des hautes terres (dont les tableaux 4 et 5 *supra* rendent compte dans les grandes lignes, à partir des données de Pike 1948) : il généralise la proclise de *ki=* à l'accompli lointain, ce qui lui permet d'éviter les fluctuations observées à Ja, et il tend à faire fusionner les désinences de personnes en position d'accord sujet suffixal (ex : Ac NTR 2SG *bixí* vs. Hu *b'èxái*, avec coalescence régulière de la désinence *-i* équivalent à la sape de la voyelle thématique *-a* et tout autre noyau thématique, dans toute la colonne NTR 2SG du tableau 7). Mais il innove et complexifie par l'hybridation des CF, et les jeux de chaise musicale d'une CF à l'autre. Là encore, on voit chaque localité « tricoter » sa différence de modèles flexionnels, non sans connaître le modèle voisin – d'autant plus qu'une variété comme Ac est en contact intense avec Hu, qui est le grand centre urbain

mazatec des hautes terres, avec son marché, ses commerces, ses administrations, ses nombreuses écoles et sa vaste gamme de services.

## 6. Conclusion

Sans fonder notre représentation graphique sur un seul résultat quantitatif – comme celui de la figure 10 supra, mais en tentant d’y condenser l’ensemble des observations de terrain de l’équipe ALMaz, je dirais que la trame de la diversité dialectale du mazatec s’organise selon un schéma comme dans la figure 11 ci-dessous<sup>40</sup>, à la manière d’une molécule, plutôt que d’une topologie en aires ou en ensembles emboîtés.

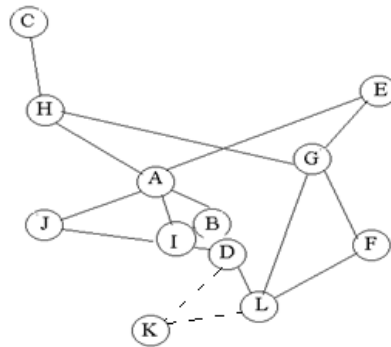


Figure 11. Réticularité des affinités dialectales au sein du diasystème mazatec.

Un tel schéma est certes très réducteur, mais il a le mérite de rendre compte de phénomènes visibles dans la plupart des espaces géolinguistiques, bien que problématiques pour le sens commun : (i) le fait que les dialectes semblent interagir à distance, et donc de manière discontinue dans l’espace, sous la double force, ou le double effet de marée des rétentions et des innovations partagées – tantôt par diffusion effective, tantôt par évolution parallèle ; (ii) la trame complexe des relations d’agrégation de proximité (comme les paquets (A, I, B, D dans le schéma) et de vecteurs à distance associant ces *grappes* ou chaînes de dialectes (comme la relation entre le paquet que je viens de mentionner, dans les hautes terres centrales, et la grappe ou chaîne (E-G-L-F), bien plus diffuse ou distendue ; (iii) pour aussi erratique ou isolé que puisse sembler une variété, elle entretiendra toujours des liens avec des comparses plus ou moins proches, comme c’est le cas pour K par rapport à D et L, ou C par rapport à H.

Enfin, et ce n’est pas là le moindre des paradoxes (en écho aux quatre dilemmes théoriques, du point de vue de la dialectologie générale, évoqués dans la section 1 du présent article), tout ce qui vient d’être décrit milite sur le plan méthodologique dans une direction qui va également contre le sens commun : c’est parce qu’on connaît de près, empiriquement et socialement une situation ou un « observatoire », que l’on peut d’autant mieux parvenir à en proposer une synthèse quasiment abstraite. Loin d’être irréductible à la théorie et aux modèles, la réalité des structures langagières et linguistiques, aussi complexe soit-elle, nécessite l’abstraction et la théorie, de manière aussi vitale que le poisson a besoin d’eau pour vivre – pas seulement pour « nager ».

Ainsi, le premier paradoxe – celui de la *phonologie continue et à gros grain vs. morphologie à grain fin* – semble se résorber en une tendance générale : la somme des interactions entre les dialectes à l’échelle de l’ensemble du réseau dialectal prévaut sur les

<sup>40</sup> C = Lo, H = Te, A = Hu, I = Ji, J = Mz, B = Mg, D = Ay, K = Cq, L = Ja, F = Do, G = Ix, E = So.



démarcations de proximité ou par agrégats vicinaux. Quand elles se produisent, ces démarcations opèrent par métatypie, en sélectionnant ou en croisant des paramètres simples, dans un éventail de critères taxinomiques et de procédés morphologiques limités, sur une gamme de classes flexionnelles clairement fondées (exemple de fondation robuste : la dichotomie des séries sous-conflatives en fonction de l'opposition entre Accord sujet 1SG & 3 versus autres personnes). Le deuxième paradoxe, celui des *périphéries congruentes vs. micro-chaînes à trame fine discontinue* a été évoqué au sujet de paramètres constructionnels, comme la préverbativité en *y'é-* à l'accompli lointain, partagée par des dialectes centraux comme Ja et So et des variétés périphériques comme Lo. Ces convergences s'expliquent par des rétentions de paramètres systémiques forts : la préverbativité en *y'é-* est là encore au cœur de ce qui fonde le système de CF mazatec : *y'é-* est un verbe léger de mouvement et trajectoire dont le signifié est 'porter'. Rien d'étonnant à ce que sa prégnance soit forte. Mais rien d'étonnant non plus à ce que chaque agrégat vicinal tente sa chance, à essayer d'autres solutions, comme le fait Ja avec des chaînes proclitiques TMAV qui convergent partiellement avec celles de Hu, opposant deux centres urbains majeurs – l'un du piémont (Ja), l'autre, des hautes terres centrales. On peut préserver ce qui marche bien dans une sous-classe, tout en expérimentant dans d'autres sous-classes – douce liberté des groupes vicinaux.

Le troisième paradoxe – celui de *l'équilibre ponctué des centres directeurs* – est consubstantiel à la formation des grappes, ou des groupes vicinaux : chacun se constitue comme un agrégat gravitant autour d'un centre densément peuplé et économiquement fort, comme Mz, Hu, Ja, Ix et So. Ce sont eux qui mènent la danse, ou donnent le *la*, au diapason. Mz et Hu comme centres *cafetaleros*, Ja comme centre d'une zone d'élevage, Ix et So, par le système canne à sucre et les activités développées autour du barrage Miguel Aleman, connecté à un centre agro-industriel comme Tuxtepec (deuxième ville de l'Etat de Oaxaca, après la ville éponyme). Même un petit bourg comme Ac ne fait que graviter autour de Te (San Jerónimo Tecuatl), ancien relais *cafetalero*, lui-même issu de l'ancienne hégémonie micro-régionale de San Antonio Eloxochitlan, aujourd'hui lieu de résilience d'une économie davantage vivrière et de subsistance autour de la milpa – caractéristique partagée par Mz également. Toute la grappe ou la chaîne de dialectes Ji, Mg, Ay et Te a gravité autour du système café dirigé depuis Hu, de l'indépendance aux années 1980. Ce centre s'est reconverti en ville de services et gros bourg commercial, et s'alimente également des économies de survie de hameaux (ou *dependencias*) des alentours, comme San Andrés Hidalgo, resté hors du système café et ayant plus longtemps conservé sa tradition *milpera*. Enfin, le paradoxe de la *robustesse des structures et de l'intrication diasystémique* peut se décrire comme le fait que le diasystème mazatec est bien plus compact que diffus du point de vue des facteurs externes de cohésion, aussi bien géographiquement que démographiquement, mais aussi du point de vue de ses mécanismes, dont une clé est le jeu cyclique de compression-unification des gabarits morpho(no)logiques, comme nous l'avons vu aussi bien pour la formation des noms que pour la formation des verbes : les cycles de Gudschinsky aussi bien que l'architecture des systèmes de classes flexionnelles fondée sur la préverbativité de verbes légers de mouvement ou de causation, sont autant de mécanismes puissants et simples qui, relayés par la compacité des réseaux intercommunautaires dans l'aire mazatèque garantissent la robustesse du système (des procédés simples et transparents) et l'intrication des solutions locales pour sérier et accommoder les lexèmes. Pour s'en rendre compte, il fallait allier la collecte des données sur le terrain, associée à une historiographie des modèles théoriques ayant permis à différents auteurs de décrire divers dialectes mazatec (Pike 1948, Kirk 1966, Gudschinsky 1956, 1958a-b, Jamieson 1982, Bull 1984), ainsi qu'un travail réflexif sur les structures grammaticales et phonologiques, en coopération avec les locuteurs<sup>41</sup>.

<sup>41</sup> Cf. très précisément, les ateliers suivants, qui se sont avérés indispensables pour modéliser le système de CF des dialectes du mazatec : à Huautla en 2012 <http://axe7.labex-efl.org/node/171>, à San Mateo Yoloxochitlan en 2012

Le graphe de la figure 11 supra et les explications qui rendent compte de cette constellation hybride, issue d'une *synthèse* entre de multiples points de vue (qualitatif et quantitatif) d'une part, et entre les faits de langues observables dans diverses composantes du lexique (substantifs et verbes) mais aussi des sous-systèmes morphologiques (systèmes de classificateurs nominaux et de classes flexionnelles verbales, ou « conjugaisons »), d'autre part, confirme l'avantage des modèles de chaînes, articulées en fonction de liens ethnolinguistiques, socioéconomiques, politiques et écologiques (cf. François 2011, 2017), sur les modèles centrifuges (Labov, Chamber & Trudgill). En revanche, le principe de la « cascade » et de la dynamique de ricochets par relais reste valide, mais c'est sa verticalité (psycho)sociale qui est discutable, non pas en soi, sur le plan purement idéologique (il serait naïf de projeter un égalitarisme de principe dans certaines sociétés plutôt que dans d'autres), mais plutôt parce que c'est plutôt, comme le fait remarquer le biologiste Edward O. Wilson « *un mélange complexe et très calibré d'altruisme, de coopération, de compétition, de domination, de réciprocité, de défection et de tromperie* » qui prévaut dans les sociétés humaines sur la simple et triviale « concurrence » ou « compétition » (Wilson, 2013 : 34). En outre, ce schéma vient apporter encore une *narration* sur la dynamique évolutive du réseau dialectal mazatec, du point de vue de la géohistoire, qui s'imbrique avec les précédentes sans les contredire, ni dans le détail, ni dans les grandes lignes, tout en reconfigurant les objets (les variétés dialectales) et les relations mutuelles que ceux-ci entretiennent entre eux.

Je crois que c'est là l'une des grandes leçons que nous donne le Lacito : pas de linguistique théorique ouvrant des horizons nouveaux sans une imprégnation dans l'épaisseur empirique du terrain et de la vie des langues dans les sociétés dont elles sont le creuset socio-cognitif. C'est ce que l'équipe de l'ALMaz a tenté de faire dans le projet résumé à grands traits dans le présent article, en associant à la fois des modèles théoriques en phonologie et en morphologie, des méthodes co-participatives de linguistique appliquée, des méthodes quantitatives de mesure de similarité géolectale, et une connaissance de l'historiographie de la description des différents dialectes de la langue aussi bien que du milieu humain de l'aire mazatèque, constitutive d'une « langue-monde », sur le plan de l'écologie humaine, environnementale et diasystémique.

**Remerciements** : Ce travail a bénéficié partiellement d'une aide de l'Etat gérée par l'Agence Nationale de la Recherche au titre du programme "Investissements d'Avenir" portant la référence ANR-10-LABX-0083, dans le cadre des opérations PPC11 (axe 1) et EM2 (axe 7), ainsi que d'un financement IUF, de 2010 à 2014, dans le cadre du projet MAmP (cf. <http://jll.smallcodes.com/home.page>). L'auteur remercie également les relecteurs du comité éditorial de cette édition en ligne des contributions au colloque de jubilé du Lacito, dont les remarques ont été stimulantes pour la conception finale de ce texte.

## Références

- ANGOUJARD J.-P. 1997, *Théorie de la syllabe*, Paris, Editions du CNRS.
- ANGOUJARD J.-P. 2006, *Phonologie déclarative*, Paris, Editions du CNRS.
- BOBERG Charles, NERBONNE John, & WATT Dominic (eds.) 2018. *The Handbook of Dialectology*, Oxford, Wiley Blackwell.
- BULL, Brian. 1984. Aspect Formation of San Jeronimo Mazatec Verb. *SIL-Mexico Workpapers* 6, 93-117.
- CARRILHO Ernestina, MAGRO Catarina & ALVAREZ Xosé (eds.) 2013. *Current Approaches to Limits and Areas in Dialectology*, Cambridge, Cambridge Scholars Publishing.
- CHAMBERS, J.K. & TRUDGILL Peter 1980. *Dialectology*. Cambridge, Cambridge University Press.

---

<http://axe7.labex-efl.org/node/87>, à Corral de Piedra, près de Temascal et donc de So, en 2013 <http://axe7.labex-efl.org/node/123>.

- FULCRAND, Julien 2017. *Les changements en chaîne historiques confrontés à la phonologie moderne. Propulsion et traction modélisées par deux approches de préservation*, thèse de doctorat, Lille III.
- FRANÇOIS, Alexandre 2011. "Social ecology and language history in the northern Vanuatu linkage: a tale of divergence and convergence", *Journal of Historical Linguistics* 1(2), 175–246.
- FRANÇOIS, Alexandre 2017. "Méthode comparative et chaînages linguistiques : Pour un modèle diffusionniste en généalogie des langues", Jean-Léo Léonard (éd.). *Diffusion : implantation, affinités, convergence*, XXIV, Peeters, *Mémoires de la Société de Linguistique de Paris*, p.43-82, 2017, <<http://www.peetersleuven.be/boekoverz.asp?nr=10559>>. <hal-01692136>.
- GARDETTE, Pierre, [1970]-1983. "Rencontre de synonymes et pénétration du français dans les aires marginales", in *Revue de Linguistique Romane*, 1970, 34, 280-305, repris in HORIOT, Brigitte, SIMONI, Marie-Rose & STRAKA, Georges (éds) : Pierre Gardette : *Etudes de géographie linguistique*, Strasbourg, Klincksieck, p. 521-546.
- GOLSTON, Chris & KEHREIN, Wolfgang, 1998. "Mazatec onsets and nuclei", *International Journal of American Linguistics*, 64.4, p. 311-337.
- GOLSTON, Chris & KEHREIN, Wolfgang, 2004. "A prosodic Theory of laryngeal contrasts", *Phonology* 21, 1-33. <<http://zimmer.csufresno.edu/~chrisg/>>.
- GUDSCHINSKY, Sarah, 1955. "Lexico-Statistical Skewing from Dialect Borrowing", *IJAL* 21(2), 138-149.
- GUDSCHINSKY, Sarah C., 1956. *Proto-Mazatec Structure*, Master of Art Thesis, manuscript, Université de Pennsylvania.
- GUDSCHINSKY, Sarah, 1958a. *Proto-Popotecan. A Comparative Study of Popolocan and Mixtecan*, *IJAL*, 25(2).
- GUDSCHINSKY, Sarah C., 1958b. "Mazatec dialect history", *Language* 34, p. 469-481.
- HEERINGA, W, 2004. *Measuring dialect pronunciation differences using Levenshtein distance*. Ph.D. Dissertation, University of Groningen.
- JAMIESON, Carole Ann, 1982. "Conflated subsystems marking person and aspect in Chiquihuitlán Mazatec verb", *IJAL* 48(2), p. 139-167.
- KIRK, Paul Livingston, 1966. *Proto-Mazatec phonology*. PhD dissertation, University of Washington. Accessible, entre autres liens sur <http://axe7.labex-efl.org/node/145>.
- LABOV, William 1994. *Principles of Linguistic Change. Volume 1: Internal Factors*. Oxford: Basil Blackwell,
- LABOV, William 2002. Pursuing the Cascade Model, n D. Britain and J. Cheshire (eds). *Social Dialectology: In Honor of Peter Trudgill*. Amsterdam, John Benjamins, p. 9-22.
- LEONARD, Jean Léo, 2010a « Diasystème et spécification de traits en tu'un savi (mixtecan, otomangue orientale) », *Bulletin de la Société Linguistique de Paris* 55 : 265-326.
- LEONARD, Jean Léo, 2010b. "Enquêtes exploratoires pour l'ALMaz (*Atlas Lingüístico Mazateco*). Elicitation croisée, entre typologie et codification d'une langue otomangue", Grenoble, *Géolinguistique* 11, pp. 59-109.
- LÉONARD, Jean Léo & Alain KIHM, 2010. "Verb inflection in Chiquihuitlán Mazatec: a fragment and a PFM approach", in MÜLLER S. (ed.), *Proceedings of the HPSG10 Conference*. CSLI Publications (available online on <http://csli-publications.stanford.edu/>).
- LÉONARD, Jean Léo & Alain KIHM, 2012. "Classes flexionnelles du mazatec et diasystème. Empirisme critique et formalisation", *Bulletin de la Société de Linguistique de Paris* 56, p. 379 - 446.

- LEONARD, Jean Léo, DELL'AQUILA, Vittorio & GAILLARD-CORVAGLIA, Antonella, 2012. "The ALMaz (*Atlas Lingüístico Mazateco*): from geolinguistic data processing to typological traits", *STUF*, Akademie Verlag, 65, p. 78-94.
- LEONARD, Jean Léo, POLIAN Gilles, HEINSALU, Els, PATRIARCA, Marco 2014. "Variación dialectal de la morfología tseltal (Maya occidental) en los ámbitos morfológico, fonológico y léxico: un enfoque holístico del diassistema", in Léonard, J.L. & Kihm, A. (eds), 2014. *Patterns in Mesoamerican morphology*, Paris, Michel Houdiard Editeur, p. 280-303.
- LEONARD, Jean Léo & FULCRAND Julien 2015. "Tonal Inflection and dialectal variation in Mazatec", in Palancar, E. & Léonard, J. L. (eds) *Tone & Inflection*, Trends in Linguistics. Studies and Monographs, 296, Mouton de Gruyter, p. 165-195.
- LEONARD, Jean Léo & FULCRAND Julien 2018. "*Inflectional Class Shifts in the Mazatec diasystem: innovation, contact and metatypy*", *STUF* 71-3, 2018, Mouton de Gruyter, pp. 429-473.
- LEONARD, Jean Léo, Els HEINSALU, Marco PATRIARCA, Anirban CHAKRABORTI & Kiran SHARMA, 2016, *Modular algorithmic complexity applied to the Mazatec diasystem*. In Léonard, Jean Léo, Didier Demolin & Karla Janire Avilés González (eds.), 2016, *Proceedings of the International Workshop on Structural Complexity in Natural Language(s) (SCNL)*. Paris, colloque, mai 2016, Université de Paris 3 - Labex EFL (PPC11), accessible sur <http://axe7.labex-efl.org/node/353>.
- LEONARD, Jean Léo & Julien Fulcrand (sous presse). « Complexité de la flexion verbale dans les diassystèmes basque et mazatec : notes diassystémiques », in *Lapurdum* 2015 19 : 389-403.
- LEVENSHTAIN, Vladimir 1966. 'Binary codes capable of correcting deletions, insertions, and reversals'. *Soviet Physics Doklady* 10(8). 707-710. Berlin, Springer.
- LONGACRE, Robert 1957. *Proto-Mixtecan*, Indiana University Research Center in Anthropology, Folklore and Linguistics 5; IJAL 23-4.
- MCCARTHY, John, 1986. "OCP Effects: Gemination and Antigemination", *Linguistic Inquiry* 17, p. 207-263.
- MARLETT, S. 1992. "Nazalization in Mixtec languages", *IJAL*, 58.4: 425-435.
- PIKE, Kenneth L. & PIKE, Eunice 1947. "Immediate constituents of Mazatec Syllables", *IJAL* 13:78-91.
- POPPER, Karl 1934. *Logik der Forschung. Zur Erkenntnistheorie der modernen Naturwissenschaft* : 1973. *La Logique de la découverte scientifique*, trad. fce par Nicole Thyssen-Rutten et Philippe Devaux, Paris, Payot.
- SCHEER, Tobias 2015. *Précis de structure syllabique, accompagné d'un appareil critique*, Lyon, ENS Editions.
- SCHWARTZ, Diana Lynn 2016. *Transforming the tropics: development, displacement, and anthropology in the Papaloapan, Mexico, 1940S-1970s*, PhD dissertation, Dpt of History, Chicago, University of Chicago.
- STUMP, Gregory T. 2001. *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge, Cambridge University Press.
- TRUDGILL, Peter 1974. "Linguistic change and diffusion: description and explanation in sociolinguistic dialect geography", *Language in Society* 3, p. 215-246.
- TUNKALO, E. A., 1946. *Vepsän kielen äännehistoria*, Helsinki, Suomalainen Kirjallisuuden Seura.
- ULFSBJORNINN Shanti 2016. "Compounding and Proclisis in Italian and English. Syntactic or phonological variation" Queen Mary Linglunch, 09/03/2016 (exemplier).
- WALLERSTEIN, Immanuel. [2004]-2006. *Comprendre le monde. Introduction à l'analyse des systèmes-monde*, trad. Camille Horsey & François Gèze, Paris : La Découverte.

WEINREICH, Uriel, 1954. "Is a structural dialectology possible ?", *Word*, 4, pp.388-400. WILSON, Edward O. 2012 *The Social Conquest of Earth*, trad. fce par Marie-France Desjeux *La conquête Sociale de la Terre*, Paris, Flammarion.