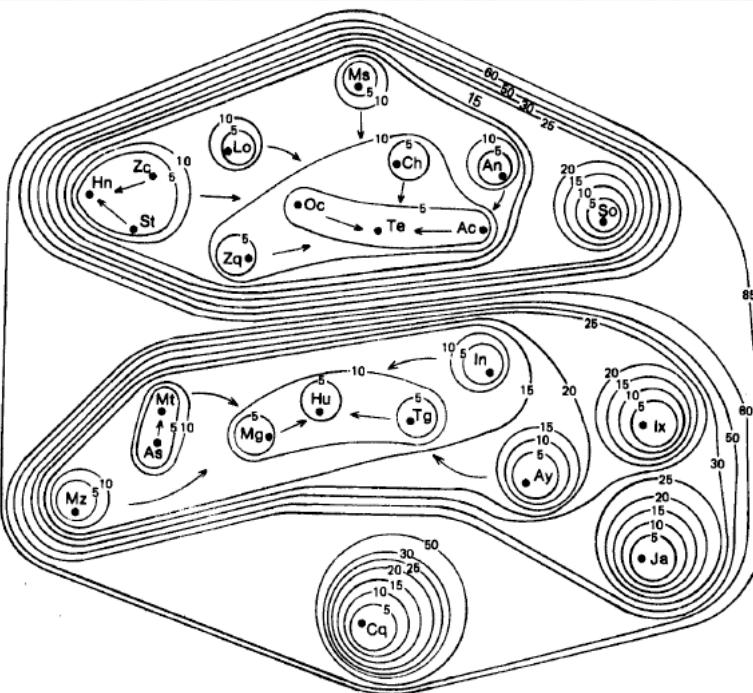


*Modular algorithmic complexity applied to
the Mazatec diasystem*

SCNL Workshop, 30-31 May 2016

**Jean Léo Léonard, Marco Patriarca, Antonella Gaillard-Corvaglia,
Pierre Darlu, Anirban Chakraborti & Kiran Sharma**

To Sarah Gudschinsky & Paul Livingston Kirk, for their contribution to Otomanguean (and general) linguistics and, indirectly... to Complexity Theory.



Source of the graph: Casad, 1974 - mutual intelligibility in Mazatec, according to Kirk's tests (1970)

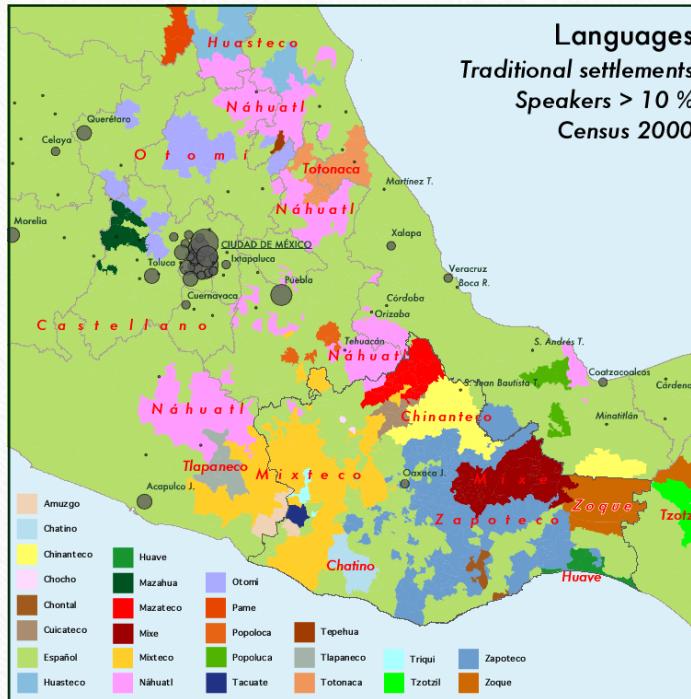
Map of the talk

- **1. Why Mazatec as a case study?**
- **1.1. Context**
- **1.2. Sarah Gudschinsky's models of dialect differentiation**
- **2. The Mazatec diasystem**
- **2.1. Cladistic sampling of three structural components**
- **2.2. Levenshtein distance (LD)**
- **2.2.1. LD sampling (111 cognates, all lexical categories)**
- **2.2.2. LD sampling (300+ cognates, only nouns)**
- **3. Conclusion & prospects**

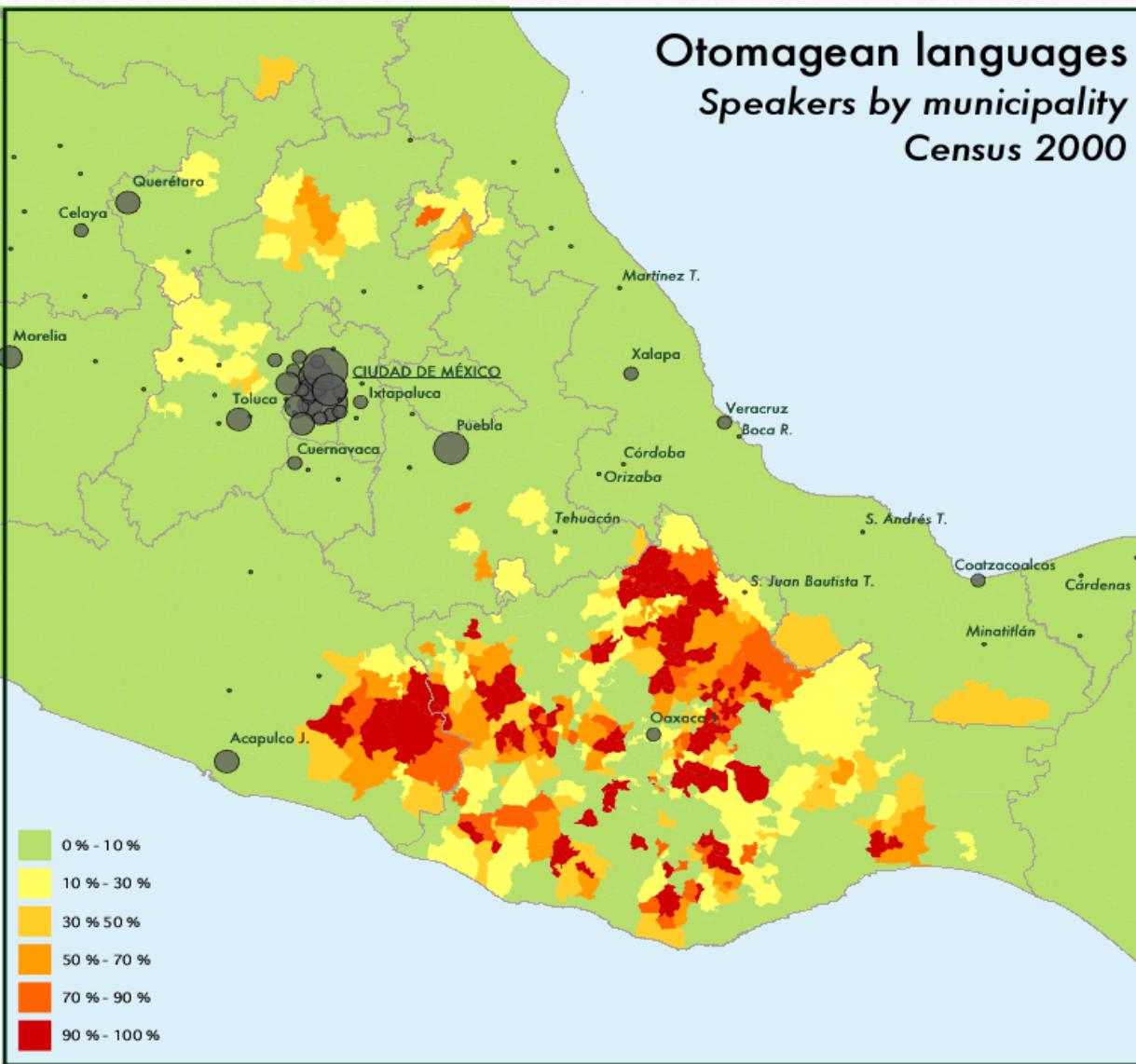
1. Why Mazatec as a case study?



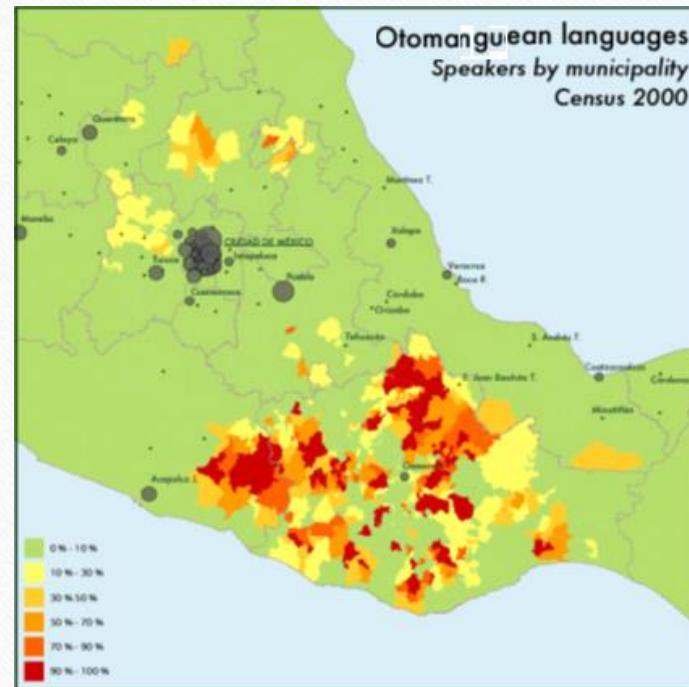
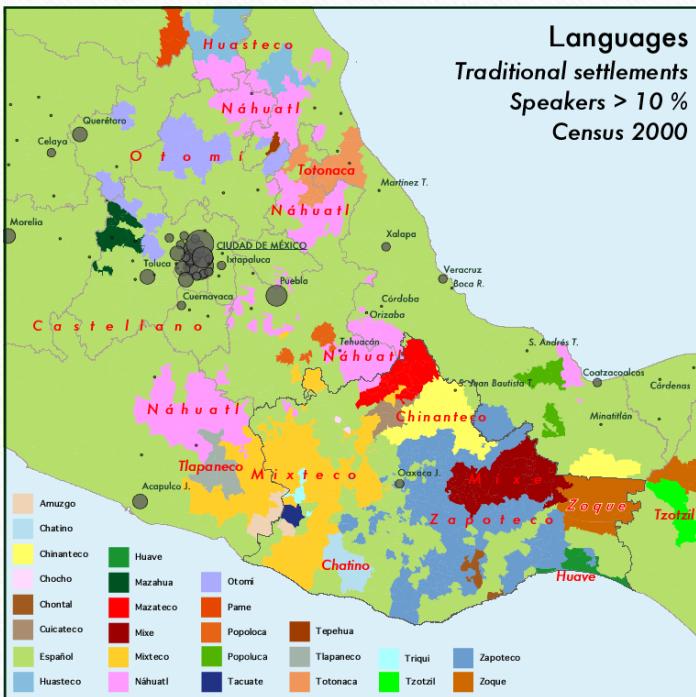
1.1. Context



Otomagean languages
Speakers by municipality
Census 2000



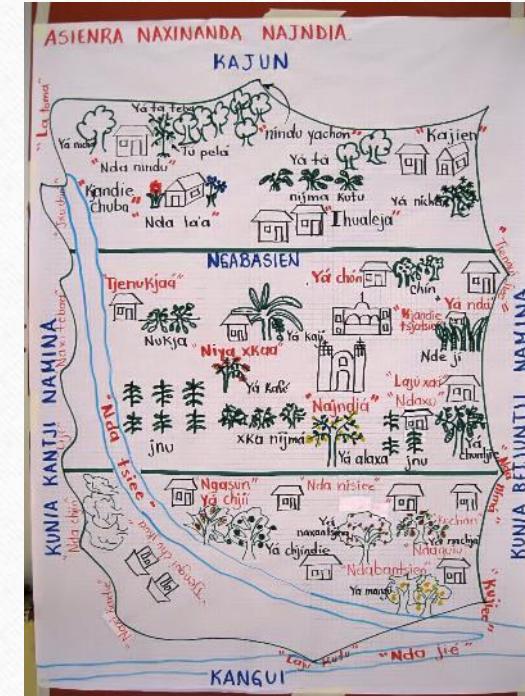
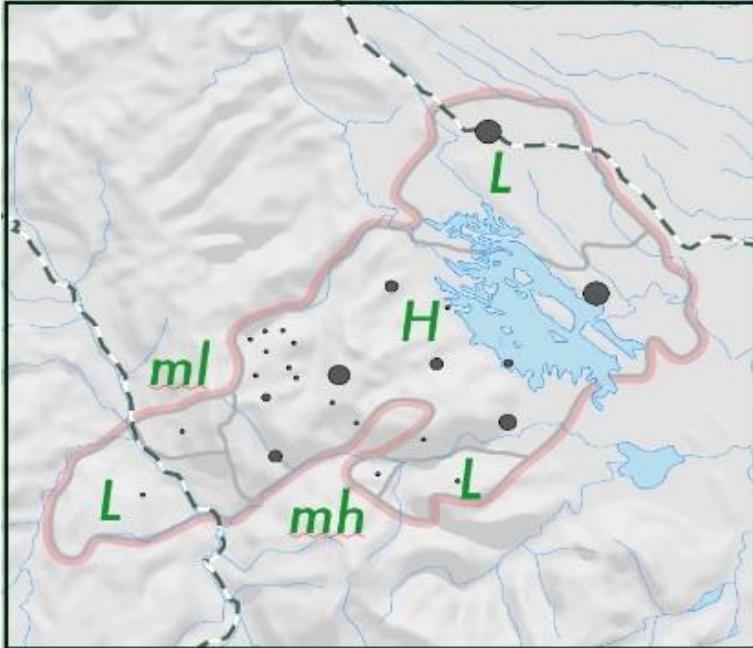
Otomanguean and other Mesoamerican languages (map: Vittorio dell'Aquila, CELE 2012)



Otomanguean Languages (Map: Josserand, 1983)



A Threefold Ecology:
Highlands vs. Midlands (or Piedmont) vs. Lowlands
Map: CELE;
Drawing: school masters, Mazatlán 2011



1.2. Sarah Gudschinsky's models of dialect differentiation



A milestone in the historiography of ecological linguistics or linguistic ecology / Ecolinguistics:

Gudschinsky Sarah, « Mazatec dialect history », *Language*, n° 34, 1958, p. 469-481.

MAZATEC DIALECT HISTORY

A STUDY IN MINIATURE

SARAH C. GUDSCHINSKY

Summer Institute of Linguistics

1. Introduction. It is well known that at least two conflicting forces contribute to the development of languages and dialects: regular sound change, which is studied by the techniques of comparative linguistics, and diffusion of linguistic innovations beyond their original source, which is studied by the techniques of linguistic geography.¹ The present paper offers a detailed study of these two forces and their interaction in the dialects of Mazatec.²

1.1. A procedure for the subgrouping of related languages or dialects on the basis of exclusively shared innovations was proposed by Brugmann and recently elaborated by Dyen (in a review of Otto Dahl's comparative study of Malagasy and Maanyan).³ In the present study of the Mazatec dialects, the procedure is applied to an Amerindian language.

The Mazatec dialects are so closely related that a single phonemic innovation may constitute the entire phonological evidence for the unique common history of a given subgroup. It is this very closeness of relationship and limitation of material that makes possible the main contribution of this paper—namely, a precise analysis, in miniature, of successive dialect splits and of the effects on these splits of the disturbing factors of shift in the boundaries of speech communities. This analysis is supported by a high correlation with the word geography of the area and with what is known of the political history of the Mazatecs.

The lack of a wealth of evidence is offset by a careful sifting of the material. Shared retentions of features of the parent language are excluded from this discussion. As Dyen pointed out, 'if we were to consider retentions, we should be using the criterion of proto-relationship also as evidence of subrelationship.'⁴ Trivialities are also excluded—phonetic changes which do not affect the phonemic system, and changes which are likely to have arisen by accidental parallel development, such as the reduction of phoneme clusters by loss.

1.2. Mazatec is spoken by nearly 90,000 people in a compact area in the north-eastern part of the state of Oaxaca, Mexico. This paper is based on data from ten dialects. Map 1 shows the approximate geographical location of the most prominent village in each dialect area. San Jerónimo Tecoati, San Antonio Eloxochitlán, San Mateo Husautla, Huautla de Jiménez, Santa María Jiotes,

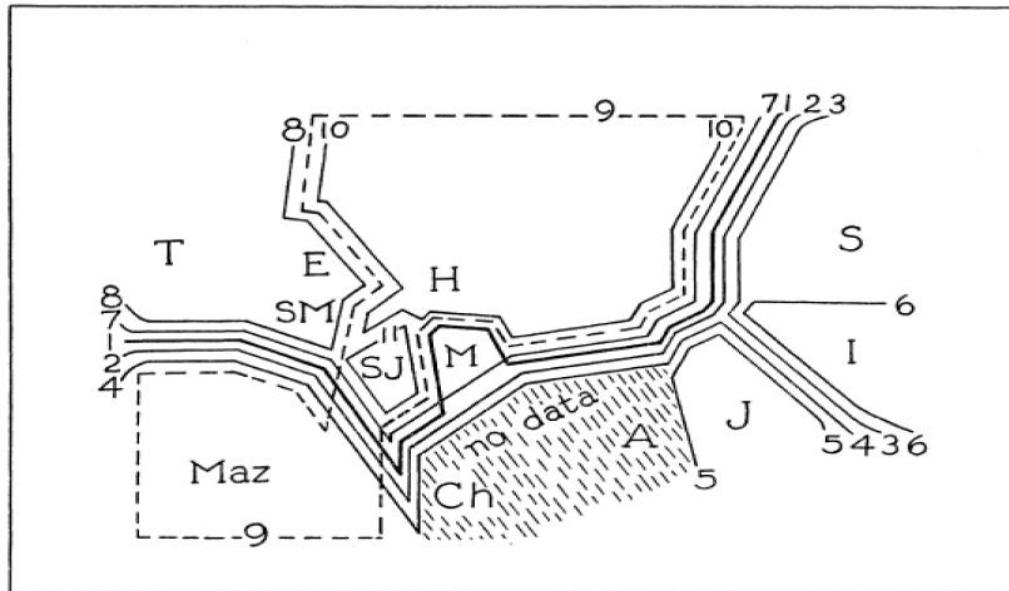
¹ See, for example, Leonard Bloomfield, *Language* 297-345 (New York, 1933).

² The Mazatec data were gathered in a series of field trips in 1949-1955 under the auspices of the Summer Institute of Linguistics. An earlier version of this paper was presented at the 1957 Summer meeting of the Linguistic Society of America.

³ Isidore Dyen, review of *Malgache et Maanjän: Une comparaison linguistique*, by Otto Chr. Dahl, in *Lg.* 29, 577-90 (1963). In his review Dyen quotes from Brugmann, *Zur Frage nach den Verwandtschaftsverhältnissen der indogermanischen Sprachen*, *Internationale*

Gudschinsky's phonological areas (1958)

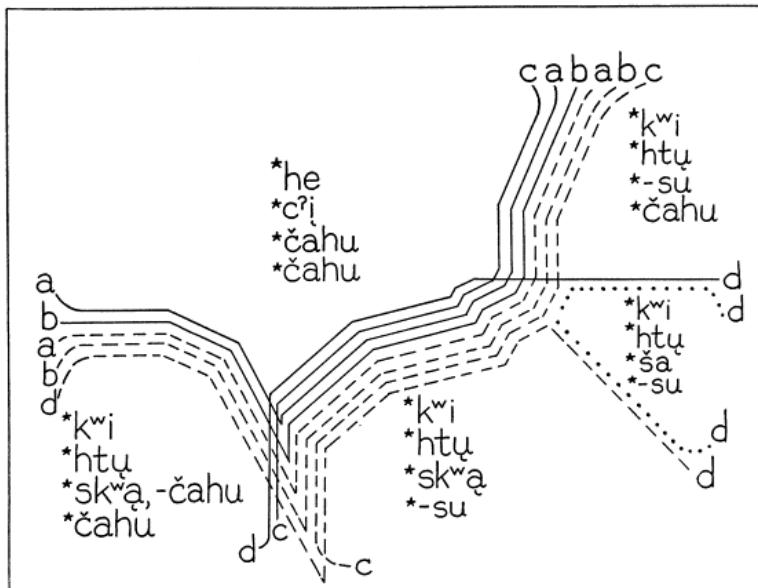
Gudschinsky Sarah, « Mazatec dialect history », *Language*, n° 34, 1958, p. 469-481.



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

Some Gudschinsky's isoglosses

Gudschinsky Sarah, « Mazatec dialect history », *Language*, n° 34, 1958, p. 469-481.



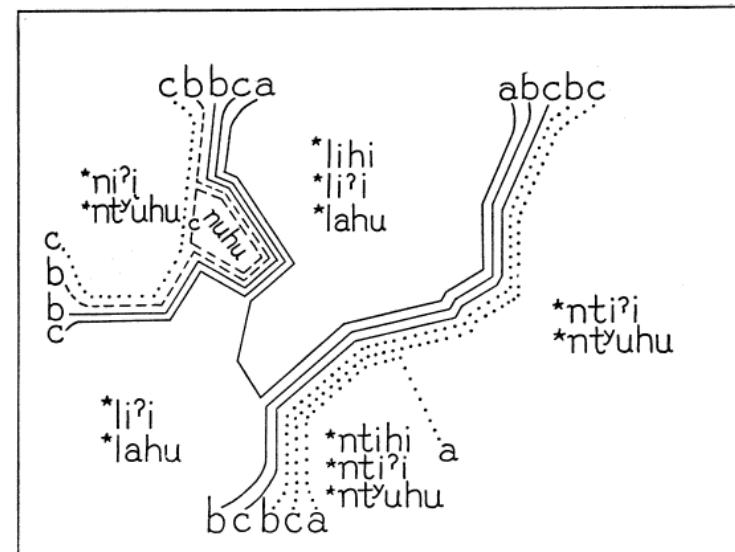
MAP 3. DISTRIBUTION OF LEXICAL PROTO-FORMS

FOR 'HE', 'BACK', 'ASHES', AND 'DUST'

- | | | | | |
|---------|-------------------------|---------------------------|------------|--------|
| line a: | solid *he | broken *k ^w i | 'he' | |
| line b: | solid *c [?] i | broken *h/t ^y | 'back' | |
| line c: | solid *čahu | broken *-su | 'ashes' | |
| line d: | solid *čahu | broken *sk ^w q | dotted *ša | 'dust' |

MAZATEC DIALECT HISTORY

477



MAP 4. DISTRIBUTION OF *l, *n, *nt

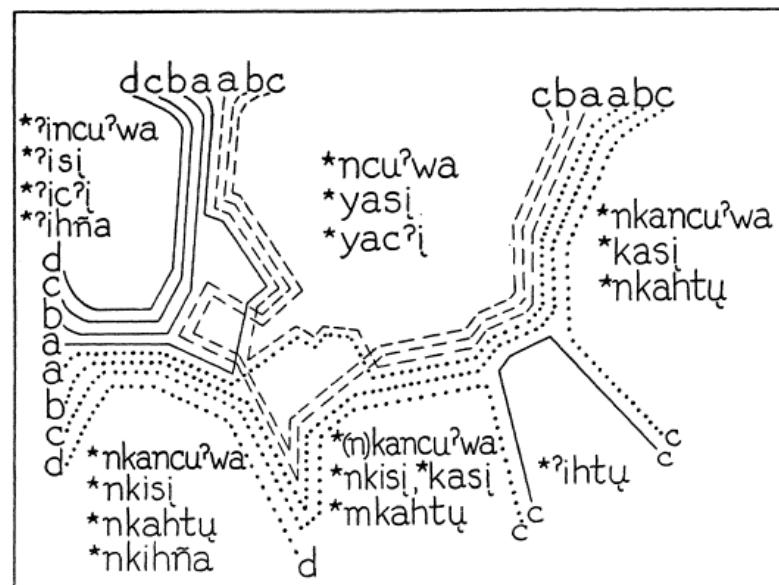
- | | | | | |
|---------|-------------|---------------|----------------|---------|
| line a: | solid *lihi | dotted *ntihi | 'grass' | |
| line b: | solid *li?i | broken *ni?i | dotted *nti?i | 'fire' |
| line c: | solid *nuhu | broken *nuhu | dotted *ntuuhu | 'stone' |

Gudschinsky's phonological areas (1958)

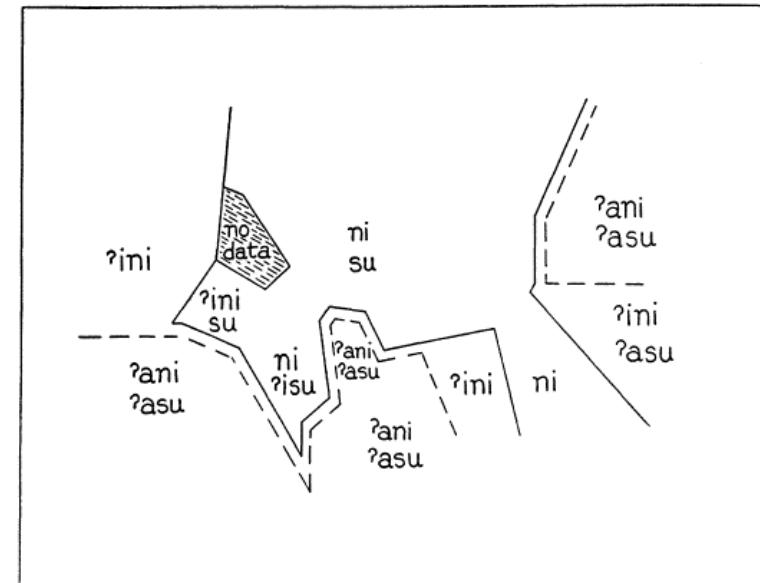
Gudschinsky Sarah, « Mazatec dialect history », *Language*, n° 34, 1958, p. 469-481.

478

LANGUAGE, VOLUME 34, NUMBER 4 (1958)



MAP 5. DISTRIBUTION OF *?i-, *ya-, *(n)ka/i-

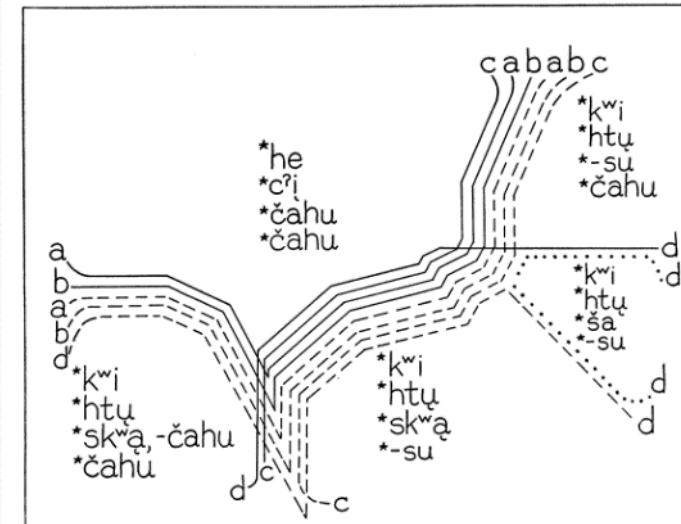
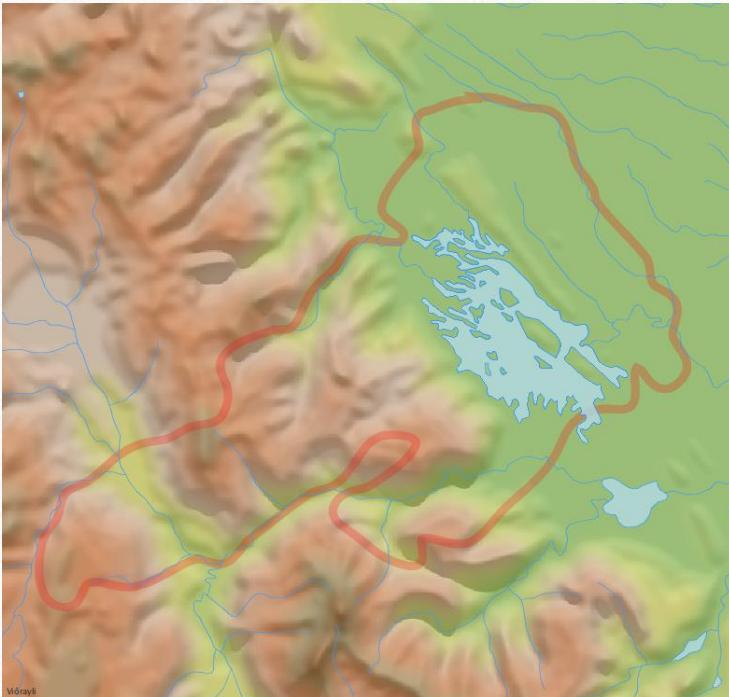


MAP 6. DISTRIBUTION OF *?i-, *?a-, *?u-.

solid line: limits of zero pre-position spreading from H
dashed line: limits of *?i- pre-position spreading from T

Gudschinsky's areas and ecological settings

Righthand map: Gudschinsky Sarah, « Mazatec dialect history », *Language*, n° 34, 1958, p. 469-481.



MAP 3. DISTRIBUTION OF LEXICAL PROTO-FORMS

FOR 'HE', 'BACK', 'ASHES', AND 'DUST'

line a:	solid *he	broken *k'wi	'he'
line b:	solid *c'i	broken *htu	'back'
line c:	solid *čahu	broken *-su	'ashes'
line d:	solid *čahu	broken *skwq	dotted *ša 'dust'

Gudschinsky's *Stammbaum* (1958), based on phonological variables

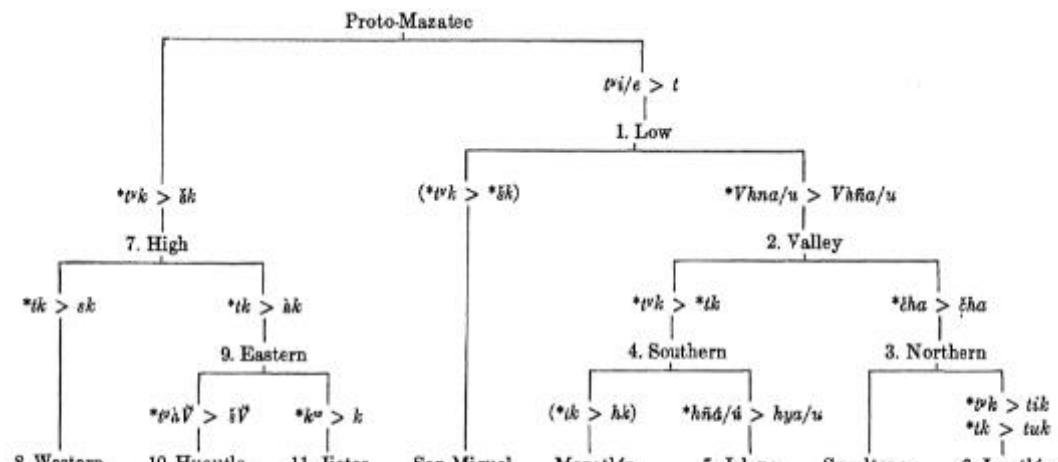


FIGURE 1. STEMMA OF THE MAZATEC DIALECTS

Numbers refer to isoglosses on Map 2

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

Some Gudschinsky's assumptions about the ethnohistorical background of phonological variation in Mazatec (1958: 480)

5. Linguistic history. On the basis of the foregoing comparative, historical, and geographical data, the history of the Mazatec dialects can be summarized as a succession of seven periods.

(A) A period of relative homogeneity, in which the free alternation of **a* and **u* in initial syllables occurred over the entire area.

(B) A period in which minor innovations, including phonological and lexical changes, spread from the central area, leaving relic islands on the periphery. It is probable that periods (A) and (B) preceded the period in which a lowland nation first flourished.

(C) A period in which Low Mazatec developed its characteristic phonological and lexical features. Apparently the speakers of the Mazatlán and San Miguel

Some Gudschinsky's assumptions about the ethnohistorical background of phonological variation in Mazatec (1958: 481)

dialects were part of the Lowland group at that time. This period probably corresponds to the era of the Lowland Mazatec Nation.

(D) A period in which the Valley dialect developed the characteristic features which distinguish it from the San Miguel dialect. It may have been later in the same period that the Southern Valley developed the features which differentiate it from the Northern Valley dialect. This period probably corresponds to the period of foreign domination.

(E) A period in which the High dialect developed its characteristic features, shared by the San Miguel and Mazatlán dialects. This period seems to have coincided with the historical period from 1300 to 1456, during which two kingdoms were set up and Mazatlán was joined to the highland. It was apparently also during this period that the Mixtec village which cuts off San Miguel from Low country was established.

(F) A period in which the Western High, San Miguel, and Northern Low dialects were differentiated. This period probably coincided with the period of Aztec rule.

(G) A period in which recent developments took place in individual dialects. This may overlap the end of the Aztec rule and the beginning of Spanish rule.

Three years before, Sarah Gudschinsky had published an important paper about (lexico)statistic skewing for isoglottic/geolinguistic analysis(1955)

LEXICO-STATISTICAL SKEWING FROM DIALECT BORROWING

SARAH C. GUDSCHINSKY

SUMMER INSTITUTE OF LINGUISTICS

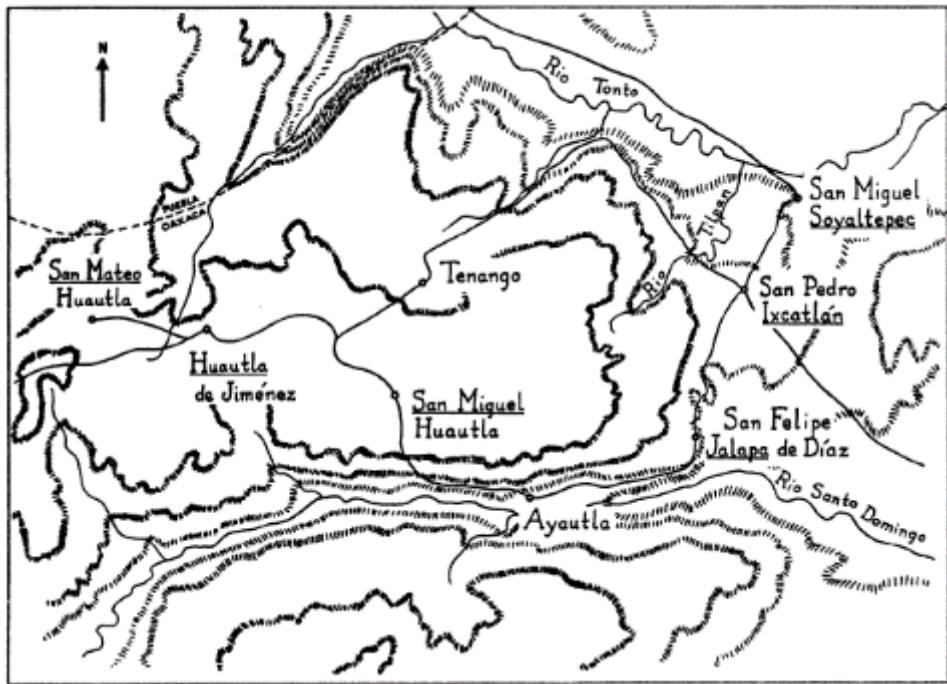
- 0. Introduction
- 1. The dip
- 2. Lexical dialects
- 3. Phonological dialects
- 4. Discrepancy between lexical and phonological dialects
- 5. Conclusion

0. Introduction. It is the purpose of this paper: (a) to propose a technique for discussing degrees of lexical relationship between dialects in terms of dips,¹ which, although based on lexico-statistical techniques, provide an objective measure of relationships free from inherent implications of absolute time (though not of relative time); (b) to show, for six closely related Mazatec² dialects, successive periods of lexical development as deduced from lexico-statistical data in terms of dips; (c) to show for the same dialects successive periods of phonological development as deduced from data obtained by traditional

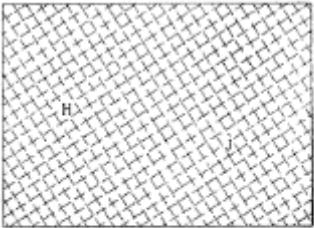
Huautla, San Miguel Soyaltepec, San Pedro Ixcatlán, and San Felipe Jalapa de Díaz.³ The approximate geographical location of these villages is shown on Map A. Huautla, San Mateo, and San Miguel are mountain villages located within two to four hours walking distance from each other. Soyaltepec, Ixcatlán, and Jalapa are lowland villages, located more than a day's walk from the mountain villages, and separated from each other by three to six hours walking distance.

An initial lexico-statistical study was made, comparing each of the six dialects with each of the others—a total of fifteen comparisons.⁴ The pairs of dialects are listed in Table 1 according to the percentages of apparent cognates they have in common, so that the higher on the scale a particular pair of dialects is listed, the fewer cognates the dialects have in common. Since, within the basic assumptions of lexico-statistics, the lower the percentage of

Gudschinsky's area under scrutiny (six varieties of Mazatec) – 1955.

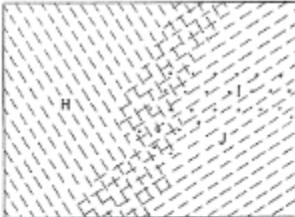


Gudschinsky's model (1955)



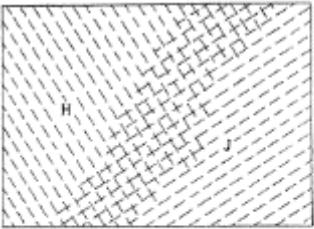
MAP B-1, of Period I

A postulated homogeneous speech community, with latent dialect differentiation of the poles indicated by the complete crosshatching of their respective symbols: \ Huasteca dialect, / Jalapa dialect.



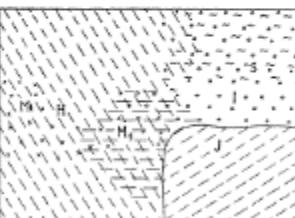
MAP B-3a, of Period IIIA

The emerging of a lexical pole within the Jalapa area is marked by the symbol O Ixcatlán, and identified by the lexical differentiation of the Ixcatlán dialect from the Huasteca dialect later than the differentiation of the Jalapa dialect from the Huasteca dialect. (See Map B-2 and Table 2).



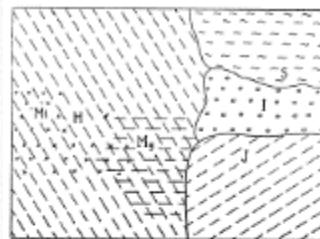
MAP B-2, of Period II

Here two distinct lexical poles have developed, symbolized by \ Huasteca, and / Jalapa, with undetermined borders symbolized by the crosshatching of the symbols in the center.



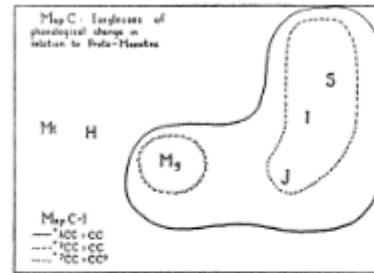
MAP B-3b, of Period IIIB

The development of lexical poles within the Huasteca and Ixcatlán dialect areas, and definition of the borders of the Jalapa dialect, O Ixcatlán, ~ Soyaltepec, X San Mateo, and — San Miguel.

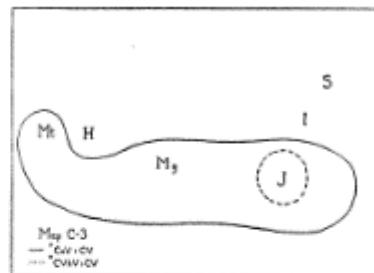


MAP B-4, of Period IV

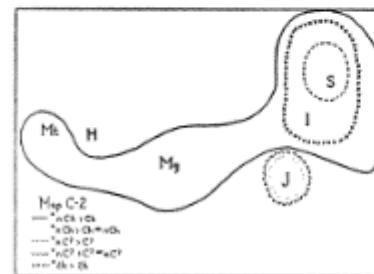
At this stage, the complete definition of the borders of the Soyaltepec and Ixcatlán dialects has occurred, but with continued transition area within the Huasteca dialect area.



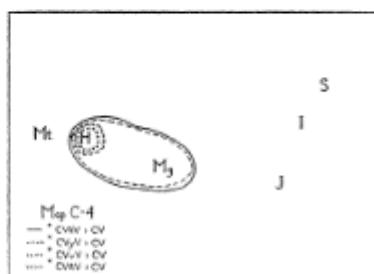
MAP C-1



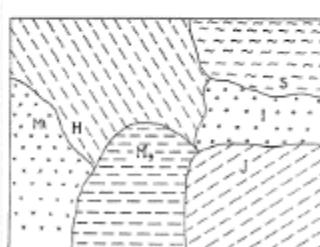
MAP C-3



MAP C-2



MAP C-4



MAP B-5, of Period V

The definition of all six dialects under consideration is complete by this period.

Gudschinsky's differentiation phases (1955)

I: Homogeneity, followed by the rise of Hu and Ja

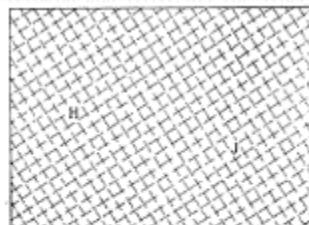
II: Emergence of a transitional buffer zone between Hu & Ja

IIIa: The lowland zone splits in two, with the emerging variety of Ix

IIIb: Both Hu and Ix areas diversify: Mt (San Mateo) emerges in the highlands, whereas So splits from Ix. In the buffer zone, Mg (San Miguel Huautepec) also emerges. Flows of lexicon and variables still pass from the Lowlands to the Highlands.

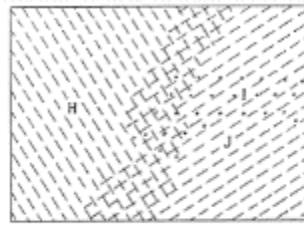
IV: Further and more clear-cut differentiation between Ix and So, in the Lowlands

V: Consolidation of the six dialects: sharper frontiers.



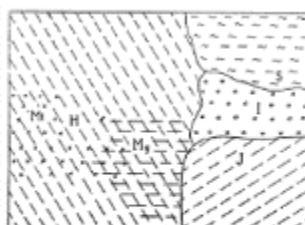
MAP B-1, of Period I

A postulated homogeneous speech community, with latent dialect differentiation of the poles indicated by the complete crosshatching of their respective symbols: \ Huastec dialect, / Jalapa dialect.



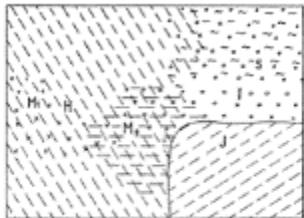
MAP B-2, of Period II

Here two distinct lexical poles have developed, symbolized by \ Huastec, and / Jalapa, with undetermined borders symbolized by the crosshatching of the symbols in the center.



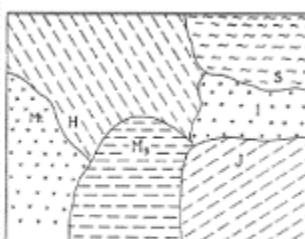
MAP B-3a, of Period IIIA

The emerging of a lexical pole within the Jalapa area is marked by the symbol ○ Ixcatlán, and identified by the lexical differentiation of the Ixcatlán dialect from the Huastec dialect later than the differentiation of the Jalapa dialect from the Huastec dialect. (See Map B-2 and Table 2).



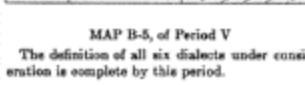
MAP B-3b, of Period IIIB

The development of lexical poles within the Huastec and Ixcatlán dialect areas, and definition of the borders of the Jalapa dialect, ○ Ixcatlán, ~ Soyaltepec, X San Mateo, and — San Miguel.



MAP B-4, of Period IV

At this stage, the complete definition of the borders of the Soyaltepec and Ixcatlán dialects has occurred, but with continued transition area within the Huastec dialect area.



MAP B-5, of Period V

The definition of all six dialects under consideration is complete by this period.



Sarah Gudschinsky

Relevance of Gudschinsky's model

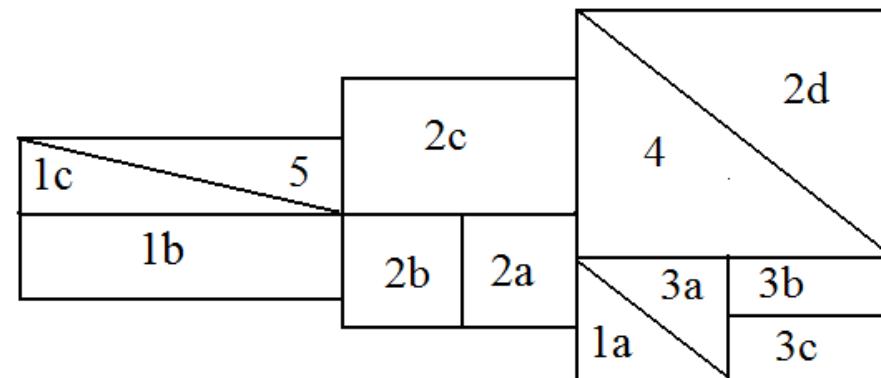
- Gudschinsky's argumentation has much in common with nowadays complex system theory on how automata get self-organised in space and time, through flow of interactions and/or introspection.
- She develops a gradual model, with buffer zones and flows of lexicon oriented from East to West (from the Lowlands to the Highlands).
- She advocates a model of punctuated equilibrium: phases of areal splits follow periods of heterogeneous equilibrium, with feature pools producing local poles, with different « whirling pace ».

2. The Mazatec diasystem



Mazatec inner diversity: the diasystem (data from Kirk, 1966; cladistic analysis)

NB: Phases in the righthand chart below are not supposed to match Gudschinsky's phases (Gudschinsky 1955), but ours, on the basis of our cladistic sampling (phonology, lexical morphology and lexical phonology of a sample of verbs)



Some typological hints relevant here, 1

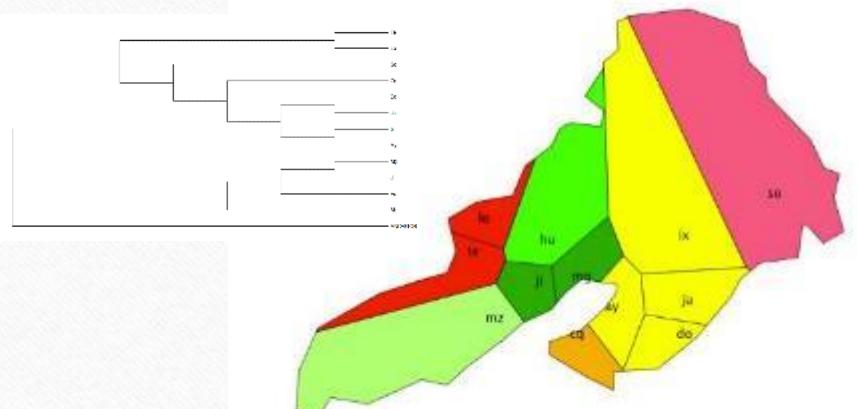
- **Tonal** language (4 level tones, 4-6 contour tones)
- **Voice Quality** language (breathy, creaky, modal)
- No dominant word order, but a trend toward **VSO, VOS**
- Strongly **Head-marking**
- Suffixal and enclitic **possessive inflection** – no other nominal nor adjectival inflection. Noun Inflection patterns morphologically similar to verb inflection.

Some typological hints relevant here, 2

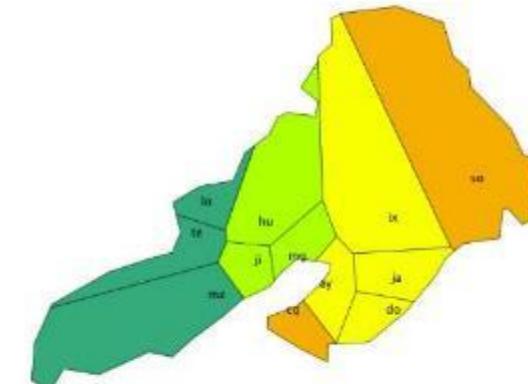
- **Aspect prevails on Tense**
- No clear-cut morphological division between transitive & intransitive
- **Light verbs** of motion / **trajectory** used in **verbal derivation** and TAMV marking and inflectional taxonomy (*taxogenesis*), i.e. making up the IC system
- **Conflative subsystems** of TAMV / **AGR.S.** marking (+/-1 SG / 3).

2.1. Cladistic sampling of three structural components

Diachronic Phonology, unpondered



Diachronic Phonology, pondered



All data from the sample, unpondered

Quand les caractères ne sont pas pondérés, on trouve,
par bandb 3 arbres de 710 pas

Tree length = 710

Consistency index (CI) = 0.6521

Homoplasy index (HI) = 0.3479

CI excluding uninformative characters = 0.4374

HI excluding uninformative characters = 0.5626

Retention index (RI) = 0.5304

Rescaled consistency index (RC) = 0.3459

Ces trois arbres diffèrent au niveau de la trifurcation,

par la position de MZ :

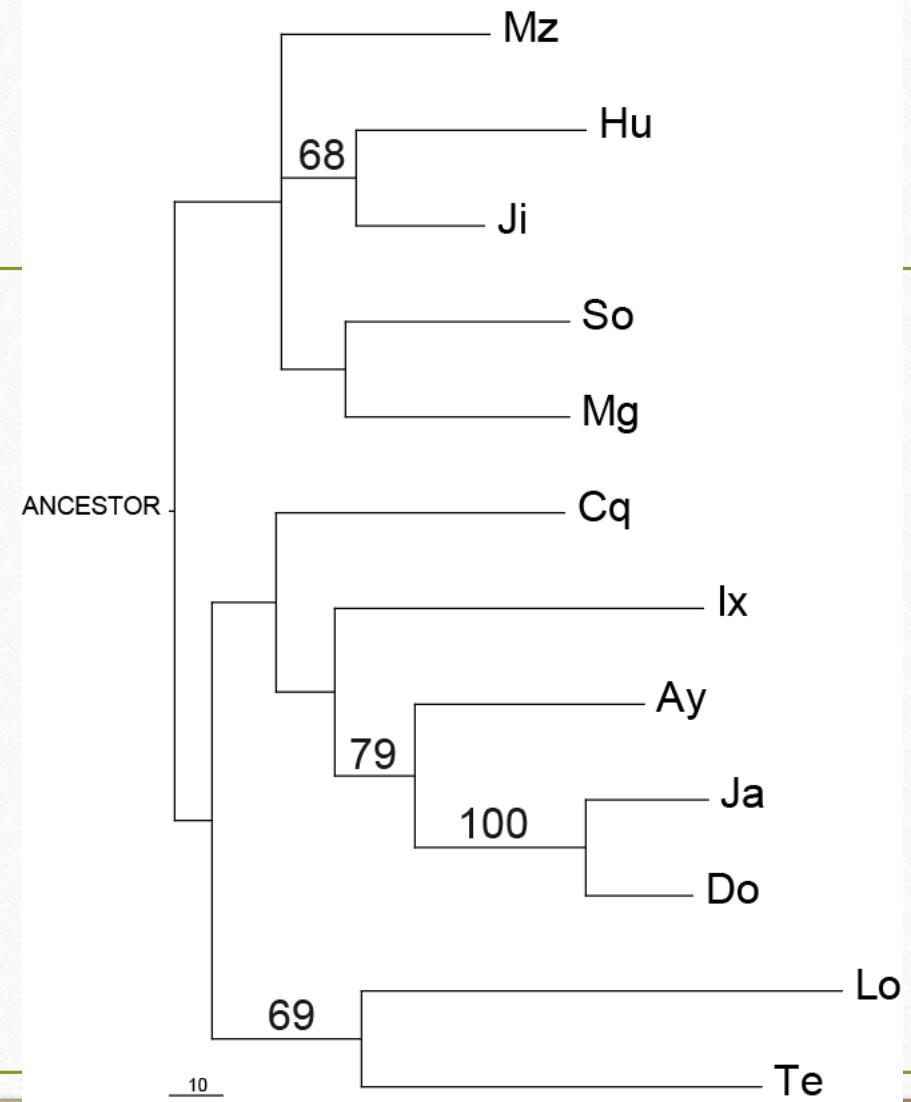
1) (MZ,((HU,JI),(SO,MG)))

2) ((MZ,(SO,MG)),(HU,JI))

3) ((MZ,(HU,JI)),(SO,MG))

L'option d'optimisation est le « branch and bound », c.a.d qu'elle examine TOUS les arbres possibles pour choisir le plus parcimonieux.

Les valeurs portées sont les Jackknife à 50% (c'est mieux que le bootstrap pour ce genre de données, même si les résultats sont quasi identiques).



Branch-and-bound search settings: Optimality criterion = parsimony

Character-status summary:

Of 467 total characters:

All characters are of type 'irrev.up'

133 characters have weight 1

334 characters have weights other than 1

4 characters are constant

271 variable characters are parsimony-uninformative

Number of parsimony-informative characters = 192

Branch-and-bound search completed:

Score of best tree found = 1609

Number of trees retained = 1

Character-state optimization: Accelerated transformation (DELTRAN)

Consistency index (CI) = 0.6787

Homoplasy index (HI) = 0.3213

CI excluding uninformative characters = 0.4417

HI excluding uninformative characters = 0.5583

Retention index (RI) = 0.5313

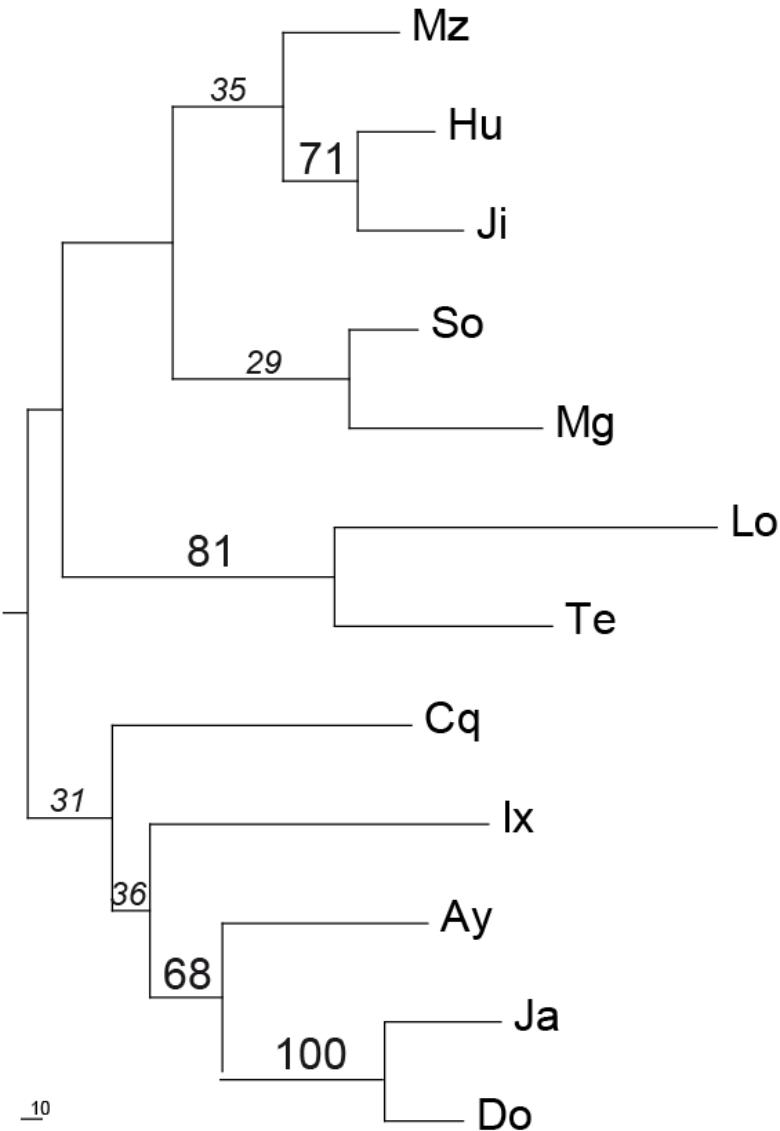
Rescaled consistency index (RC) = 0.3606

L'option d'optimisation est le « branch and bound », c.a.d qu'elle examine TOUS les arbres possibles pour choisir le plus parcimonieux.

Les valeurs portées sont les Jackknife à 50% (c'est mieux que le bootstrap pour ce genre de données, même si les résultats sont quasi identiques. En italiques, ceux qui sont inférieurs à 50%, donc très peu robustes mais encore représentables dans l'arbre).

Finalement, la pondération ne change pas radicalement la représentation. Sans doute faut-il aller dans le détail de la distribution des états de caractères pour voir s'il y a des modifications de leur position dans l'arbre. Cela dit, l'arbre est mieux résolu (en particulier pour la position de Mz). Si le couple Lo-Te change de groupe-frère, c'est à un niveau où, de toute façon, il n'y a aucune résolution.

All data from the cladistic sample, pondered



Phonological change (cladistic sample)

Branch-and-bound search completed:

Score of best tree found = 477

Tree description:

Optimality criterion = parsimony

Character-status summary:

329 characters are excluded

Of the remaining 138 included characters:

All characters are of type 'irrev.up'

29 characters have weight 1

109 characters have weights other than 1

2 characters are constant

70 variable characters are parsimony-uninformative

Number of (included) parsimony-informative characters = 66

Character-state optimization: Delayed transformation
(DELTRAN)

Tree length = 477

Consistency index (CI) = 0.6541

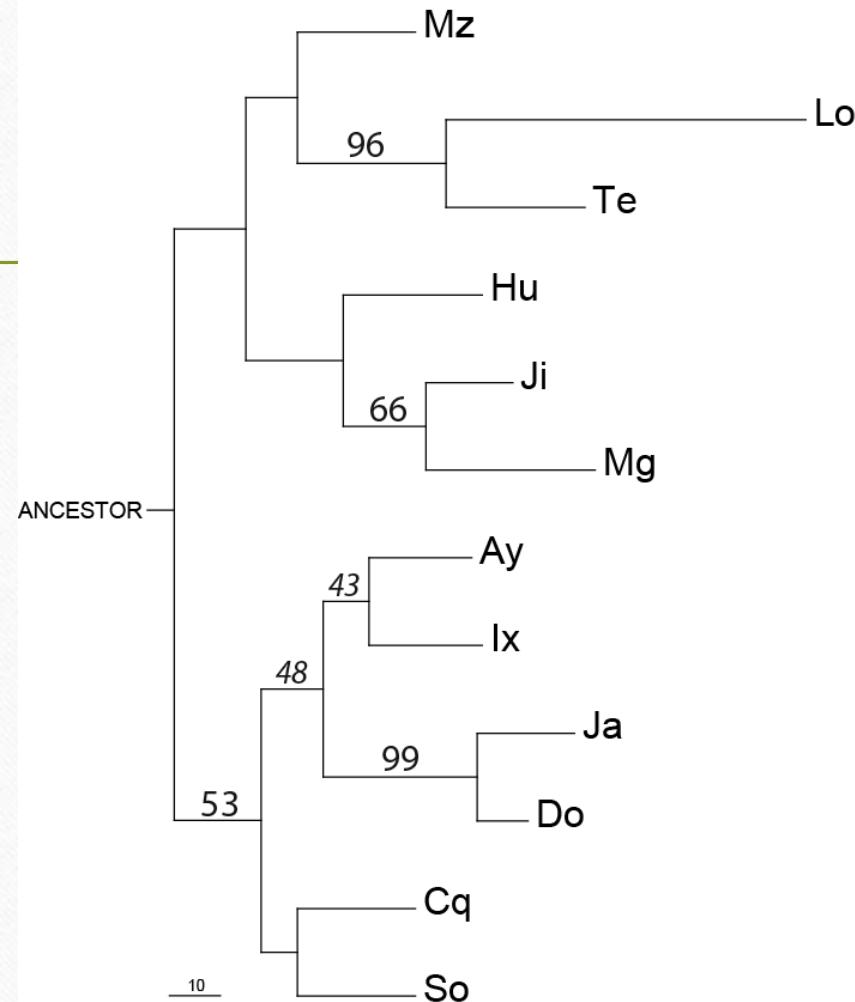
Homoplasy index (HI) = 0.3459

CI excluding uninformative characters = 0.4745

HI excluding uninformative characters = 0.5255

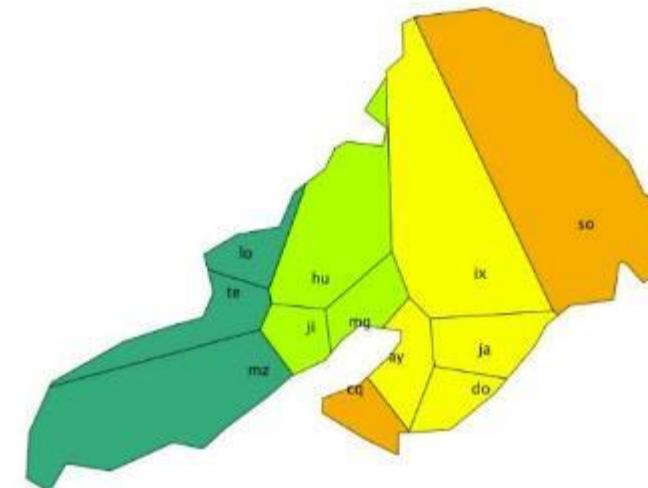
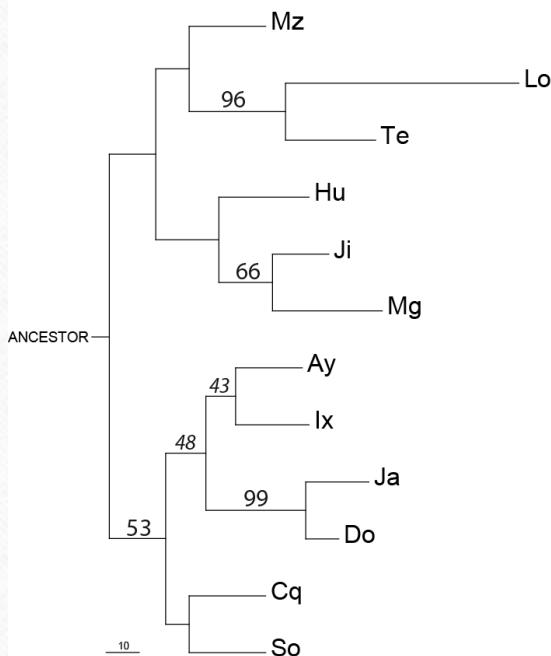
Retention index (RI) = 0.6504

Rescaled consistency index (RC) = 0.4254



Diachtonic phonology (pondered)

- see next slide for some examples of ponderation



Protoform (Kirk 1966)	Ex.	Glossing	Ex. Loc.	Processus et pondération pour traitement cladistique
*ntuhu >	ntu	« wide »	Hu	Apocope (réduction VhV) => 3
	ntuhuu		Lo	Rétraction de V haute vélaire => 2
	nduhu		Do	Voisement d'occlusive prénasalisée (<i>nC</i>) => 1
*nti?ju /ntij <u>a</u> / >	n̪i?ju /n̪i+j <u>a</u> /	« ant »	Mz	Simplification sonantique de <i>nC</i> => 3
	n̪ju		Te	Modalisation d'une attaque <i>creaky</i> => 1
*tsuntu >	tsintu	« worm »	Ay	Dissimilation séquence V hautes => 2
	tsintiu		Te	Propagation de V haute ou épenthèse => 2
*nt̪i?i >	nti?i	« fire »	Ay	Dépalatalisation => 2
	l̪i?i, l̪i /li <u>a</u> /		Hu	Latéralisation de <i>nC</i> palatal => 3
*nkahau >	nkaho	« cave »	Ji	Monophtongaison AU => 2
> *nkaihau (?) >	nkiho		Te	Dissimilation palatale intersyllabique
	nkeiho		Cq	Diphongaison décroissante palatale EI => 1

Tree description:

Optimality criterion = parsimony

Character-status summary:

294 characters are excluded

Of the remaining 173 included characters:

All characters are of type 'irrev.up'

55 characters have weight 1

118 characters have weights other than 1

108 characters are parsimony-uninformative

Number of (included) parsimony-informative characters = 65

Character-state optimization: Delayed transformation (DELTRAN)

AncStates = "standard"

Tree length = 581

Consistency index (CI) = 0.7745

Homoplasy index (HI) = 0.2255

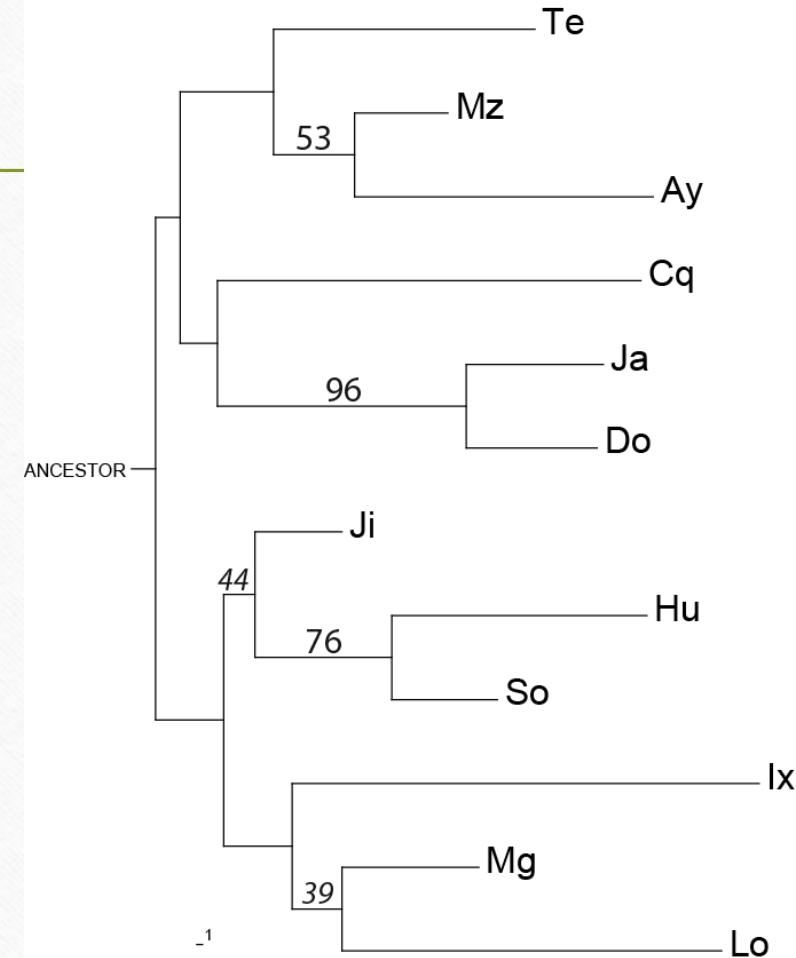
CI excluding uninformative characters = 0.5148

HI excluding uninformative characters = 0.4852

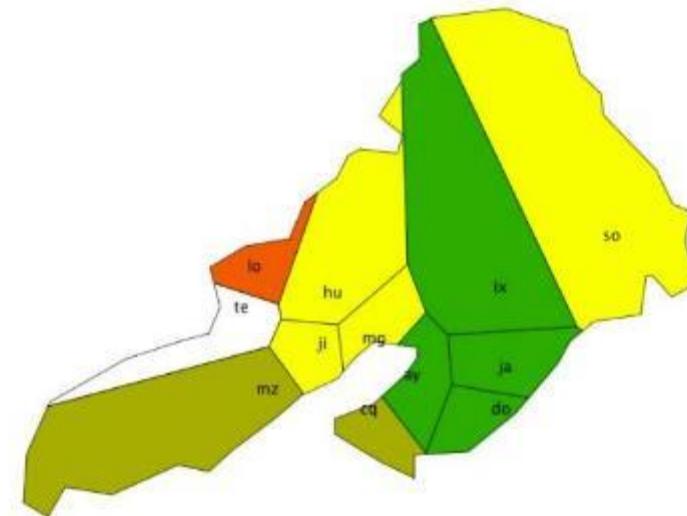
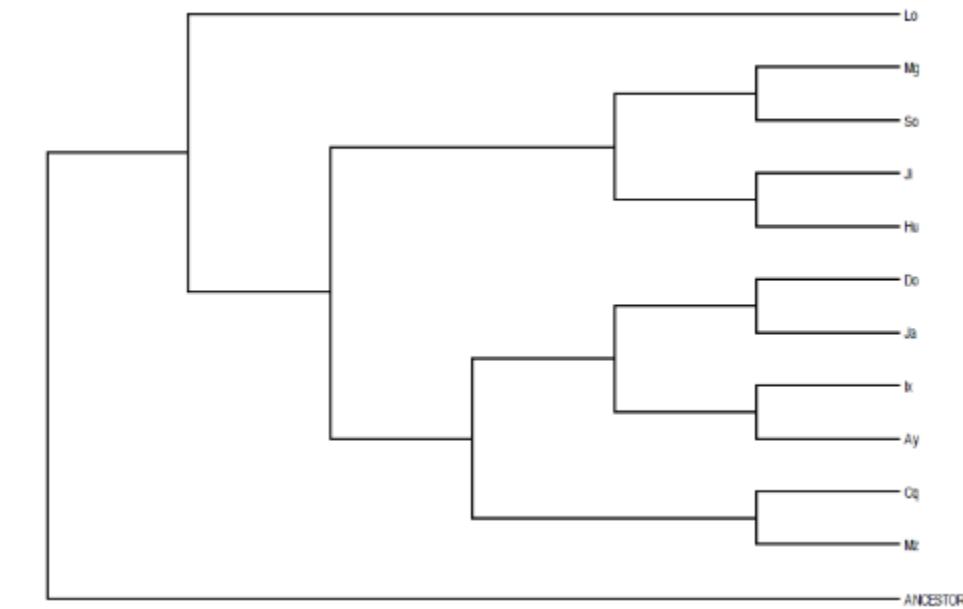
Retention index (RI) = 0.5747

Rescaled consistency index (RC) = 0.4451

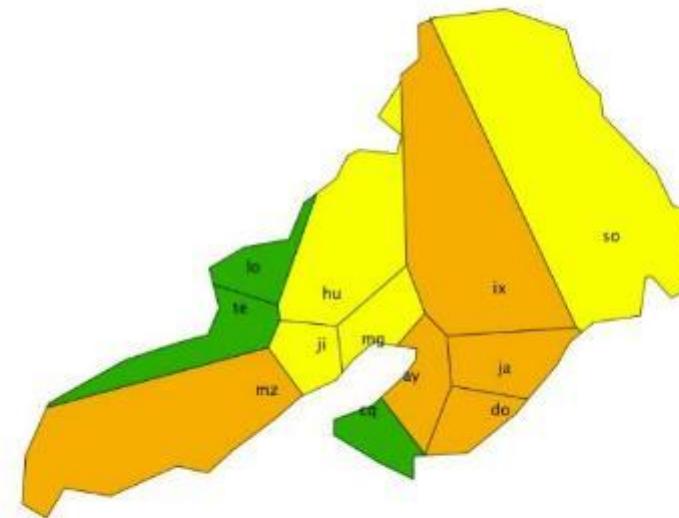
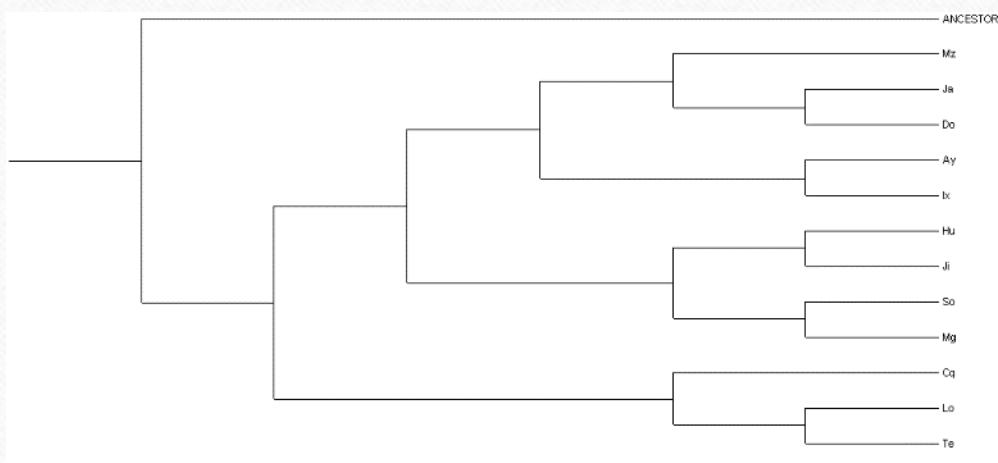
TAMV Preverbation, pondered



Preverb-Root morphophonological adjustments, unpondered

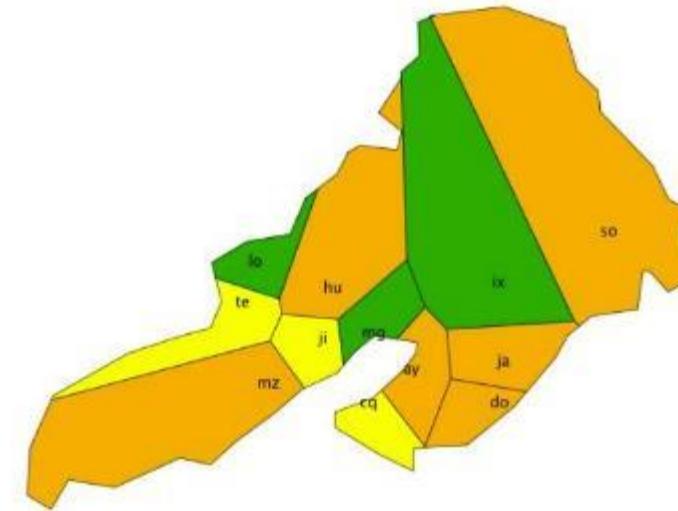
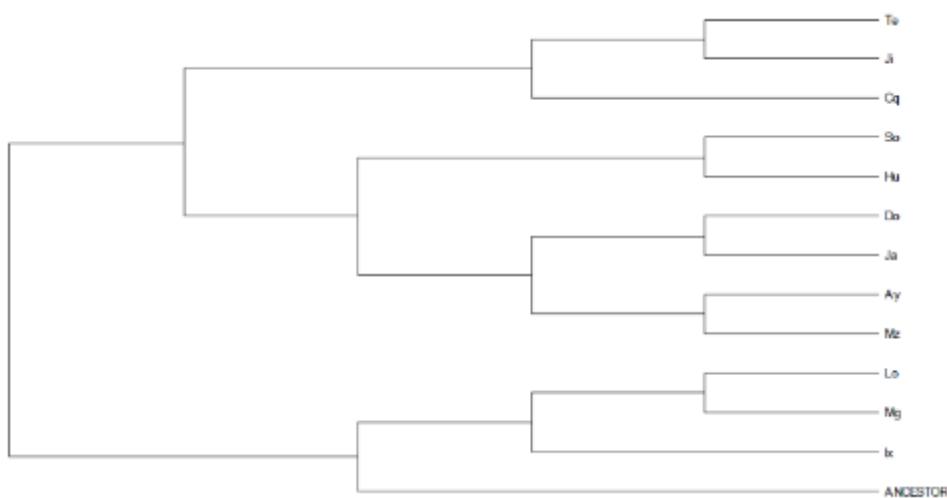


Preverb-Root morphophonological adjustements, pondered

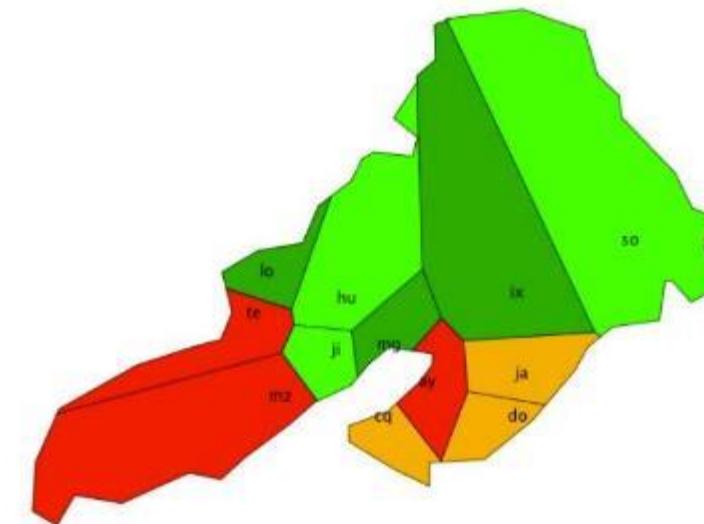
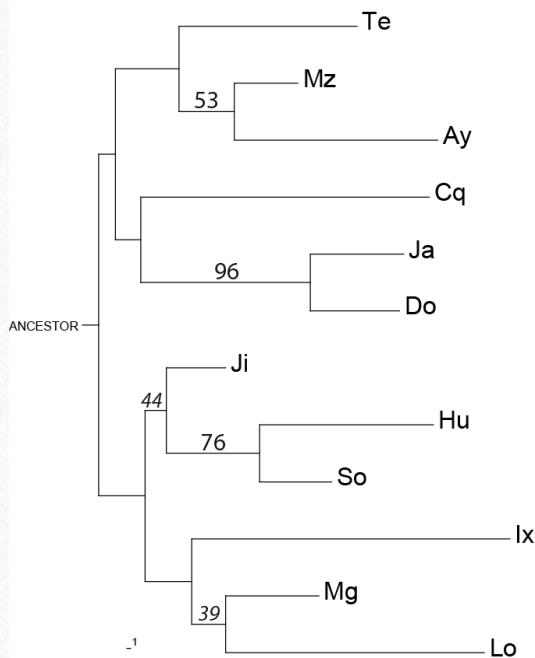


Lexical morphology & Inflection Class taxonomy

Preverbal selection, unpondered



Lexical morphology & Inflection Class taxonomy Preverbal selection, pondered



Why preverbal selection is important?

- Mazatec is a highly head-Marking language: the verb is the « Prince » or the king of sentences.
- Collocation of directional and mouvement light verbs frame up the whole system.
- This system accounts for inflectional taxonomy (inflectional classes).
- As a consequence, preverbal morphemes are densely used in discourse – these items are by no means marginal in speech; instead, they are central.

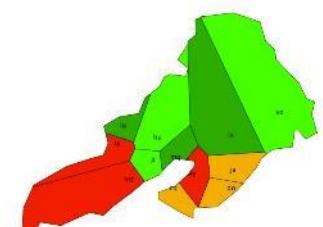
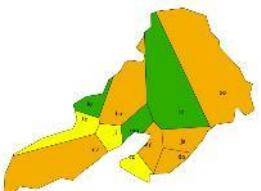
An example of preverbal morphemes in a Mazatec dialect: Chiquihuitlan

Data from: Jamieson Carole Ann 1982: « Conflated Subsystems Marking Person and Aspect in Chiquihuitlán Mazatec Verbs »,
International Journal of American Linguistics, Vol. 48, No. 2 (Apr.), pp. 139-167

Set No.	3def or ls		Other Persons	
	Neutral	Incomp	Neutral	Incomp
1	<i>be-</i>	<i>kue-</i>	<i>be-</i>	<i>kue-</i>
2	<i>ba-</i>	<i>kua-</i>	<i>ba-</i>	<i>kua-</i>
3	<i>bo-</i>	<i>sko-</i>	<i>čo-</i>	<i>čo-</i>
4	<i>bu-</i>	<i>sku-</i>	<i>ču-</i>	<i>ču-</i>
5	<i>hu-</i>	<i>sku-</i>	<i>čhu-</i>	<i>čhu-</i>
6	<i>hi-</i>	<i>ski-</i>	<i>čhi-</i>	<i>čhi-</i>
7	<i>hba-</i>	<i>khua-</i>	<i>hba-</i>	<i>khua-</i>
8	<i>fi-</i>	<i>fi-</i>	<i>nj-</i>	<i>nj-</i>
9	<i>su-</i>	<i>su-</i>	<i>nŋ-</i>	<i>nŋ-</i>
10	<i>bu-</i>	<i>ku-</i>	<i>bu-</i>	<i>ku-</i>
11	<i>ba-</i>	<i>kua-</i>	<i>ča-</i>	<i>ča-</i>
12	<i>ka-</i>	<i>ska-</i>	<i>ča-</i>	<i>ča-</i>
13	<i>hba-</i>	<i>khua-</i>	<i>nq-</i>	<i>nq-</i>
14	<i>ba-</i>	<i>kua-</i>	<i>nq-</i>	<i>nq-</i>
15	<i>bi-</i>	<i>kui-</i>	<i>bi-</i>	<i>kui-</i>
16	<i>bu-</i>	<i>sku-</i>	<i>ntu-</i>	<i>ntu-</i>
17	<i>hi-</i>	<i>si-</i>	<i>či-</i>	<i>ši-</i>
18	<i>hba-</i>	<i>khua-</i>	<i>čha-</i>	<i>čha-</i>

A diasystemic grid for Inflectional classes (Léonard & Kihm 2014, Léonard & Fulcrand, 2015)

Class			Preverb	Semantics	Person split	
Labial onsets	I	a	<i>b'é+</i>	GENERAL LOCATION	No Split	
		b	<i>ba+</i>			
		c/1	<i>bi+</i>			
	II	c/2	<i>ba+/bi+</i>	MOTION		
		c/3	<i>bá+/fa+</i>			
	III		<i>b'a+/ch'a+</i>	ATTACHMENT	Split	
	Coronal onsets	IV		CAUSATIVE		
		V		Variable (unsystematically correlated)		
Dorsal onsets	VI		<i>kjo+/chjo+</i> <i>ko+/cho+</i>	KERNEL PROCESS		
Open set	VII		<i>see+/jnta+</i>			

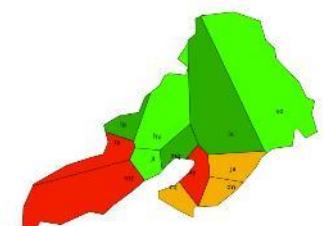
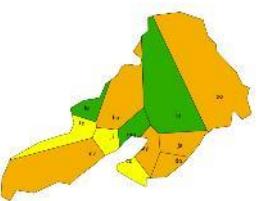


preverb	inflection class	RT class	AGR	NEUTRAL	CPL	INCPL	
<i>b'é+</i>	I.a	1/H/	3	<i>b'é+xá</i>	<i>tsa=k'-é+xá</i>	<i>k'-<u>o</u>è+xá</i>	'send'
			1SG	<i>b'<u>e</u>e+xá-a</i>	<i>tsa=k'-<u>e</u>e+xá-a</i>	<i>k'-<u>o</u>e+xá-a</i>	
			1PL.INCL	<i>k'-<u>o</u>è+xá</i>	<i>tsa=k'-è+xá</i>	<i>k'-<u>o</u>é+xá</i>	
<i>ba-</i>	I.b	3 /H/	3	<i>ba+né</i>	<i>tsa=k-a+né</i>	<i>k-<u>o</u>a+nè</i>	'wash'
			1SG	<i>ba+né-e</i>	<i>tsa=k-a+ne-e</i>	<i>k-<u>o</u>a+ne-e</i>	
			1PL.INCL	<i>ba+nè</i>	<i>tsa=k-a+nè</i>	<i>k-<u>o</u>a+nè</i>	
<i>bi-</i>	I.c/1	3 /LM/	3	<i>bi+yaa</i>	<i>tsa=k-i+yaa</i>	<i>k-<u>o</u>i+yá</i>	'die'
			1SG	<i>bi+y<u>a</u>-a</i>	<i>tsa=k-i+y<u>a</u>-a</i>	<i>k-<u>o</u>i+yá-a</i>	
			1PL.INCL	<i>bi+y<u>a</u>-à</i>	<i>tsa=k-i+y<u>a</u>-à</i>	<i>k-<u>o</u>i+yá</i>	
<i>ba-/bi-</i>	II.c/2	1 /M/	3	<i>ba+te</i>	<i>tsa=k-a+te</i>	<i>k-<u>o</u>a+te</i>	'break'
			1SG	<i>ba+te</i>	<i>tsa=k-a+te</i>	<i>k-<u>o</u>a+te</i>	
			1PL.INCL	<i>bi+chà</i>	<i>tsa=k-i+chà</i>	<i>k-<u>o</u>i+chá</i>	
<i>bá-/fa-</i>	II.c/3	3 /M/	3	<i>báj+tsa</i>	<i>tsa=k-áj+tsa</i>	<i>k-<u>o</u>áj+tsa</i>	'put to'
			1SG	<i>b<u>a</u>j+ts<u>a</u>-a</i>	<i>tsa=k-<u>a</u>j+ts<u>a</u>-a</i>	<i>k-<u>o</u>áj+ts<u>a</u>-a</i>	'cook'
			1PL.INCL	<i>faj+tsa-à</i>	<i>tsa=kj-aj+tsa-à</i>	<i>kj-<u>o</u>áj+tsa-à</i>	
<i>b'a-/ch'a-</i>	III	2 /H/	3	<i>b'a+kjá</i>	<i>tsa=k'-a+kjá</i>	<i>k'-<u>o</u>a+kjá</i>	'wear'
			1SG	<i>b'a+kjá-a</i>	<i>tsa=k'-a+kjá-a</i>	<i>k'-<u>o</u>a+kjá-a</i>	
			1PL.INCL	<i>ch'á+kjá</i>	<i>ki=ch'á+kjá</i>	<i>ch'<u>a</u>+kjá</i>	
<i>sí-/ni-</i>	IV	5 /h/	3	<i>sí+tsjò</i>	<i>ki=s<u>í</u>i+tsjò</i>	<i>s<u>í</u>i+tsjò</i>	'toast'
			1SG	<i>s<u>í</u>i+tsjò-a</i>	<i>ki=s<u>í</u>i+tsjò-a</i>	<i>s<u>í</u>i+tsjò-a</i>	

3	<i>b'é+xá</i>	<i>tsa=k'-é+xá</i>	<i>k'-<u>oè</u>+xá</i>	'send'
1SG	<i>b'<u>e</u>e+xá-a</i>	<i>tsa=k'-<u>ee</u>+xá-a</i>	<i>k'-<u>oe</u>+xá-a</i>	
1PL.INCL	<i>k'-<u>oè</u>+xá</i>	<i>tsa=k'-è+xá</i>	<i>k'-<u>oé</u>+xá</i>	
3	<i>ba+né</i>	<i>tsa=k-a+né</i>	<i>k-<u>oa</u>+nè</i>	'wash'
1SG	<i>ba+né-e</i>	<i>tsa=k-a+ne<u>e</u>-e</i>	<i>k-<u>oa</u>+ne<u>e</u>-e</i>	
1PL.INCL	<i>ba+nè</i>	<i>tsa=k-a+nè</i>	<i>k-<u>oa</u>+nè</i>	
3	<i>bi+y<u>aa</u></i>	<i>tsa=k-i+y<u>aa</u></i>	<i>k-<u>oi</u>+yá</i>	'die'
1SG	<i>bi+y<u>a</u>-a</i>	<i>tsa=k-i+y<u>a</u>-a</i>	<i>k-<u>oi</u>+yá-a</i>	
1PL.INCL	<i>bi+y<u>a</u>-à</i>	<i>tsa=k-i+y<u>a</u>-à</i>	<i>k-<u>oi</u>+yá</i>	
3	<i>ba+te</i>	<i>tsa=k-a+te</i>	<i>k-<u>oa</u>+te</i>	'break'
1SG	<i>ba+te</i>	<i>tsa=k-a+te</i>	<i>k-<u>oa</u>+te</i>	
1PL.INCL	<i>bi+chà</i>	<i>tsa=k-i+chà</i>	<i>k-<u>oi</u>+chá</i>	

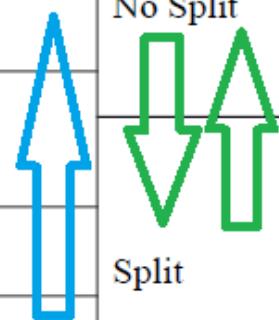
Some basic Inflectional Class Shifts

Class		Preverb	Semantics	Person split
Labial onsets	I	a <i>b'é+</i> b <i>ba+</i> c/1 <i>bi+</i>	GENERAL LOCATION	No Split
	II	c/2 <i>ba+/bi+</i> c/3 <i>bá+/fa+</i>	MOTION	Split
	III	<i>b'a+ / ch'a+</i>	ATTACHMENT	
	IV	<i>sí+ / ni+</i>	CAUSATIVE	
	V	<i>tsò+ / mi+</i>	Variable (unsystematically correlated)	
Dorsal onsets	VI	<i>kjo+ / chjo+</i> <i>ko+ / cho+</i>	KERNEL PROCESS	
Open set	VII	<i>see+ / jnta+</i>		

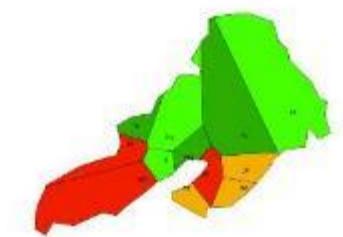


Some endemic trends correlated to IC Shifts

Class			Preverb	Semantics	Person split	
Labial onsets	I	a	<i>b'é+</i>	GENERAL LOCATION	No Split	
		b	<i>ba+</i>			
		c/1	<i>bi+</i>			
	II	c/2	<i>ba+/bi+</i>	MOTION		
		c/3	<i>bá+/fa+</i>			
	III		<i>b'a+/ch'a+</i>	ATTACHMENT		
Coronal onsets	IV		<i>sí+ / ni+</i>	CAUSATIVE	Split	
	V		<i>tsò+ / mi+</i>	Variable (unsystematically correlated)		
Dorsal onsets	VI		<i>kjo+ / chjo+</i> <i>ko+ / cho+</i>	KERNEL PROCESS		
Open set	VII		<i>see+ / jnta+</i>			

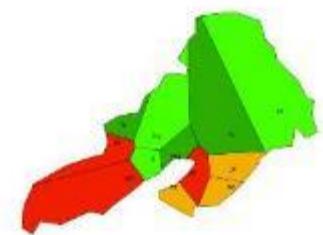
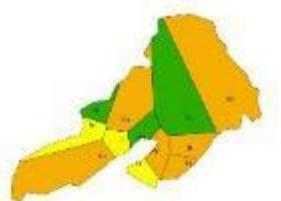


No Split
Split

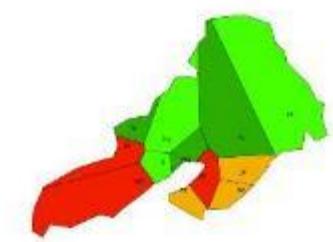
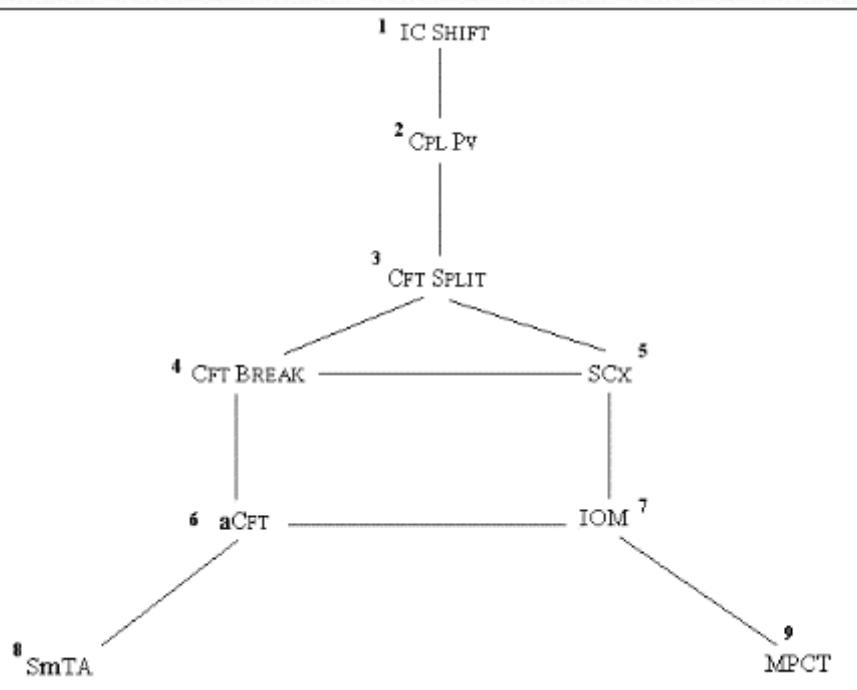


Preverbal extension

Class			Preverb	Semantics	Person split
Labial onsets	I	a	<i>b'é+</i>	GENERAL LOCATION	No Split
		b	<i>ba+</i>		
		c/1	<i>bi+</i>	MOTION	
	II	c/2	<i>ba+/ bi+</i>	Split	
		c/3	<i>bá+/ fa+</i>		
	III		<i>b'a+/ ch'a+</i>		ATTACHMENT
Coronal onsets	IV		<i>si+ / ni+</i>	CAUSATIVE	
	V		<i>tsò+ / mi+</i>	Variable	
Dorsal onsets	VI		<i>kjo+ / chjo+ ko+ / cho+</i>	(unsystematically correlated)	
Open set	VII		<i>see+ / jnta-</i>	KERNEL PROCESS	



Modeling inflexional class shifts in the Mazatec diasystem in a (partly implicational) graph (cf. Haspelmath 2004) (Léonard & Fulcrand, 2015)



Stems (RSS: Rules of Stem Selection)

Léonard & Fulcrand (2015)

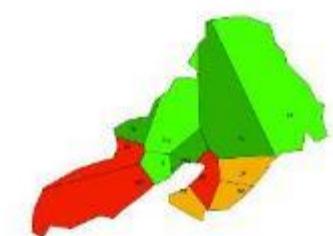
- Stem formation processes:

- Anticonflative (**aCfl**), ex. IIC2 > IB, IC1: {ba X / bi X } > {ba X }
- IC Shift (**ICS**), ex. IA > IC1: {b'é X } > {bi X }
- Conflation split (**CflS**), ex. IA > IIC2: {b'é X } > {b'é X / bi X }
- Specific Completive preverb (**CPL PV**): Neutral {b'é X } vs. CPL {y'é X }
- Conflation breaking (**Cfl.Br**): ex. +3 vs. -3, instead of +3/1SG vs. other AGRS (e.g. including 1 SG, or opposing 1 SG to 3 vs. -3/-1)
- Stem Template Allomorphy (**SmTA**), ex. b'é X vs. b'ét'a X

Clitics & Affixes (RE = Rules of Exponence)

Léonard & Fulcrand (2015)

- The next two processes have to do with Rules of Exponence (SCx & IOM), while the last one (MPCT) has much to do with Morphophonological Rules.
- PV string complexification (**SCx**). NB: the string may be made of clitics or prefixes.
- Incomplete overmarking (**IOM**).
- Morphophonological conflation transfer (**MPCT**), ex. San José Independencia, {b'éX} > {b'éX / 'éX}.



Modeling inflectional class shifts in the Mazatec diasystem in a (partly implicational) graph (Léonard & Fulcrand, 2015)

