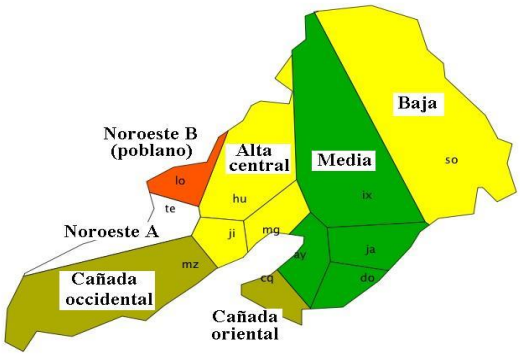
Léonard Jean Léo[1]
IUF & Paris 3, UMR 7018
International Workshop of the Surrey Morphology Group, University of Surrey, UK
Inflectional Morphology and verb classes in Oto-Manguean, 14-03-2014.

MODELING SIMPLEXITY *VERSUS* COMPLEXITY IN MAZATEC INFLECTIONAL CLASSES: A DIASYSTEMIC APPROACH

1. Introduction

The Mazatec diasystem (Popolocan, EO) can be divided into two main zones: the highlands and the lowlands. Other subzones can be further distinguished, such as the Valley (Cañada) and the Puebla area (see San Francisco Huehuetlán data below), or the Midlands (Jalapa de Diaz, Santo Domingo, San Pedro Ixcatlán).



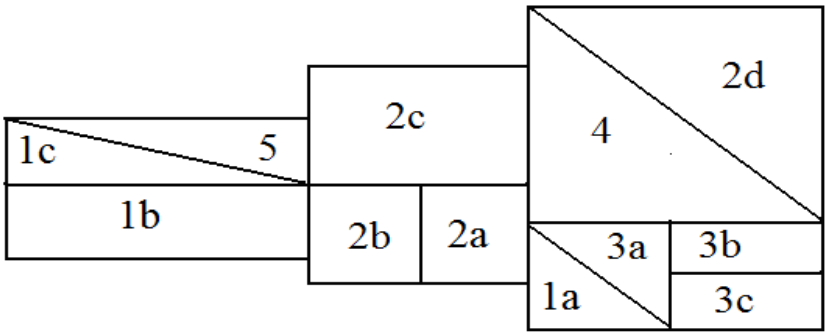


Figure 1. Dialect divisions and a scenario for dialect diversification in 5 periods (Léonard, Gaillard-Corvaglia & Darlu, in progress)[2]

[1] Acknowledgements: this research benefitted extensively from the PFM formalization designed initially by Alain Kihm (cf. Léonard & Kihm, 2010, 2001, 2012), and from substantial help by Julien Fulcran in transcribing data from San Miguel Huautepec (recorded by Léonard & Calderon, ALMaz, August 2012). Errors or inconsistency may only be mine.

[2] The Mazatec dialect areas presented in figures 1.1-2 are the result of a cladistic study with Antonella Gaillard-Corvaglia (Ph. D. Paris 3, postdoc Inalco) and Pierre Darlu (INserm, CNRS), based on Paul L. Kirk's cognate sets for 12 Mazatec varieties

1.1. A preliminary remark

Not only is Mazatec far more regular and easier to describe than what the available data and models from the two or three most famous dialects (Huautla, Jalapa, Chiquihuitlán) suggest since the forties, but it turns out that its very dense dialect network makes the task even easier, as the diasystem varies on a few predictable tricks: preverb selection and class of equipollence, mostly.
Mazatec verb forms can therefore be schematized as in (1), where W, St, and ℜ respectively identify the word, the stem, and the root:

TENSE-ASPECT-VOICE AGR
│ │
(1) $<_{W}$ (CV) $<_{St}$ CV $<_{\Re}$ (C)C$\boxed{\text{VV> V}}$>>

Single vs. complex TAMV chains:
Complexification of IC slot, emerging TAMV (Tense-Aspect-Modality-Voice) specifier strings:

The $_{St}$ CV slot can be expanded into complex strings, according to the ***lexeme*** (Taxinomic Variation, henceforth **TxVr**) and to its ***discursive*** or ***semantic*** & ***pragmatic*** integration (Pragmatic Variation, **PrVr**)[3], and to the ***dialect*** (diasystemic variation, **DsVr**). All these dimensions are involved in the fabric of Simplexity & Complexity in Mazatec (Popolocan, Eastern Otomanguean, Mexico, 220 000 speakers).

- ⇨ Strings on the slot, expanding TAMV chains entails ... Happier, but trickier hunting grounds for complexity of IC prefixal (or proclitic chains): **TxVr**, **PrVr**, **DsVr**. **Integration**

- ⇨ String game hints at Game Theory, Naming Game, and other Complexity Theory (Gribbin, 2004) devices, that would be useful for subsequent modeling or for simulations: preverbs as automata, with a margin of **TxVr**, **PrVr**, **DsVr** complexity.

- ⇨ Towards a CT integration of (socio)linguistics, in the course of the last two decades (see Hruschka & al. 2009, Massip-Bonet & Bastardas-Boada, 2013, esp. Mufwene, 2013).

1.2. Some hints at Complexity Theory – from the outskirts

Complexity Theory (CT) includes three dimensions for the hierarchization and representation of interactive components in holistic sets: 1) *algorithmic* complexity, 2) *deterministic* complexity, 3) *aggregate*, or *communal* (especially in the social field) complexity (O'Sullivan, 2004).

Here, ***algorithmic complexity*** will be understood as the collection of phylogenetic and typological matrices of declarative parameters and variables defining variation in sound systems and grammars of the languages under scrutiny (e.g. rearticulated vowels, voice quality parameters, voice or nasality correlations, preverb selection, patterns of inflectional conflation, inflectional class complexity through mergers or splits, word order, argument alignment, animacy hierarchy, etc.).

(1966). Details on isoglosses are given in Léonard, Dell'Aquila & Gaillard-Corvaglia, 2012. See also Gudschinsky, 1958b. In short, phonological (phonetic laws) and morphological data (on the one hand, synchronic juncture morphophonological processes between preverbs and lexical roots, on the other hand, preverb selection at 3SG of the neuter aspect) have been processed with cladistics, using PAUP 4.0 and FACTOR software (see methodology in Gaillard-Corvaglia, 2012; Gaillard-Corvaglia & al. 2007, 2008, Brun-Trigaud & al. 2014). The figures 1.1-2 are a synthesis of these results, conflating six optimal phylograms (pondered and not pondered).

[3] Here we should include *referential integration* as well, but this would lead us too far.

Deterministic complexity (searching for dynamics and explanative trends rather than claiming for clear-cut causes) will not be handled here. As our intention here is to describe and model paradigms throughout the Mazatec diasystem, the question of the causes and consequences of structural changes in the diasystem, and of the consequences of external factors, lays beyond the parameter of this particular research.

Aggregate complexity here has to do with diasystemic integration: how can morphological areas be delimited? To what extent are inflectional class systems homogeneous or heterogeneous within the diasystem? What accounts for polymorphism, and with what items (lexicon, or broadly, the **TxVr**, **PrVr**, **DsVr** factors)?

Complexity may itself be understood in two separate ways: on the one hand, "constitutional complexity," or "bit complexity" (cf. Mufwene, 2013), i.e. according to inventories of functional units or structural features, such as phonemes, morphemes and lexical stems (as in Miestamo & al. 2008); or on the other hand, "interactional complexity" or "socio-interactional complexity," in other terms, "communal complexity," involving intricate modules of units and features, or networks of interactive individuals and aggregates, resorting to *communal complexity* and *agency*.

Unification between the two dimensions of complexity ("Bit Complexity" *versus* "Interaction Complexity," cf. Mufwene, 2012) is still a challenge.

The complexity of the Mazatec IC diasystem results from communal complexity – but we will not tackle this issue[4].

Simplexity means something very different from *simplicity*. Simplexity flows from ergonomics, parsimony and transparency, though all these properties may entail complex inner structures, or may involve complex sets of interactions between functional units. *Simplexity* is, for instance, a central concept for the study of body motion mechanisms.

Here, *simplexity* will mean *a system of simple constraints making complex outputs predictable*, or more explicitly, a "trick of the trade" that any complex system, such as grammar, may use to achieve its goal. In this case, the goal stands in the realization of linguistic units in speech and discourse, as smoothly and accurately as possible. Inflectional patterns pass through this simplexity filter, so that any complex output should be understood according to simplex inputs in the lexicon and in the inventory of organizing constraints (i.e. grammar).

2. Some hints at the Mazatec verb and embedding of various dimensions of complexity

Segmentation: take note of the categorization into prefixes *versus* proclitics in the TAMV complex and... take note of the the categorization of prefixes *versus* proclitics for agreement suffixes too, whether they are suffixal endings (or *désinences*) or pronominal enclitics (as in 3b). Take note of agreement subject markers' allomorphy as well, as it may prove tricky...

Huautla

(2) ki= s<u>i</u>i- ská -a (*kis<u>i</u>iskáa*)
COMPL IC.3/1SG play 1SG
I played

(3a) ki= nì-ská -i (*kinìskái*)
COMPL IC play 2SG
you played

Mazatlán

(3b) ki= ní-tsjò=ji (*kinítsjòji)*
COMPL IC toast 2SG
you toasted

[4] See Gudschinsky, 1958b for a glimpse at Mazatec ethnohistory, from which one can infer *deterministic complexity* from isoglosses. See Gudschinsky, 1958a for a broader comparative view, in terms of structural complexity. As for *epigenetic complexity*, see Kirk, 1970, and Casad, 1975.

AGR EXPONENTS

First, we must expose the AGR exponents, as they do not come out quite clearly from the verb forms. As can be readily gathered from Table 4.1, six person-number values are distinguished in Mazatec (all dialects). There is no number contrast in the 3rd person and there is an inclusive vs. exclusive distinction in the 1st person plural. Table 5 below shows the diasystemic paradigm of AGR endings:

	Singular	Plural	
3			
1	*-'a(n)*	*-ñà, -jñà*	inclusive
		-in, jin, jni	exclusive
2	*-i, ji*	*-un, jún, jnú*	

Table 4.1. Mazatec diasystemic paradigm of AGR endings

In Table 4.2. we give the corresponding paradigm for HU, very similar to the diasystemic paradigm:

	Singular	Plural	
3			
1	*-a*	*-à*	inclusive
		-ijin,	exclusive
2	*-i*	*-o*	

Table 4.2. HU AGR endings

NB: Pike's *-in* for 1Pl Excl. is by no means confirmed in today's Huautla Mazatec (HU), which has *-ijin,*instead. The first *-i-* is epenthetic, for the sake of syllabation (concatenation should therefore be analyzed as an enclitic, i.e. *=ijin*, instead of *-ijin*).

2.1. An IC Mendeleev's table : Jamieson's chart (or the *Chiquihuitlán Matrix*)

What language is famous for its 18 IC taxonomy? Chiquihuitlán Mazatec, as described by Carole Jamieson.

The Jamieson (1982) description of Chiquihuitlán Mazatec is accurate; the subdivisions are relevant – they elegantly account for the TAMV/Subject AGR correlation, expressed by conflative paradigms.

This system is most probably relevant for most of the Lowland dialects (the Soyaltepec, and probably the San Pedro Ixcatlán dialects).

C. Jamieson's Inflectional Class	3 DEF. or 1 SG.		Other persons	
	NTR	**INCMPL.**	**NTR**	**INCMPL.**
A	*be-*	*kue-*	*be-*	*kue-*
B	*ba-*	*kua-*	*ba-*	*kua-*
C	*bo-*	*sko-*	*čo-*	*čo-*
D	*bu-*	*sku-*	*ču-*	*ču-*
E	*hu-*	*sku-*	*čhu-*	*čhu-*
F	*hi-*	*ski-*	*čhi-*	*čhi-*
G	*hba-*	*khua-*	*hba-*	*khua-*
H	*tsi-*	*tsi-*	*nį-*	*nį-*
I	*su-*	*su-*	*nų-*	*nų-*
J	*bu-*	*ku-*	*bu-*	*ku-*

K	*ba-*	*kua-*	*ča-*	*ča-*
L	*ka-*	*ska-*	*ča-*	*ča-*
M	*hba-*	*khua-*	*nq-*	*nq-*
N	*ba-*	*kua-*	*nq-*	*nq-*
O	*bi-*	*kui-*	*bi-*	*kui-*
P	*bu-*	*sku-*	*ntu-*	*ntu-*
Q	*hi-*	*si-*	*či-*	*ši-*
R	*hba-*	*khua-*	*čha-*	*čha-*

Table 1.1. Inflectional Class Prefixes in Chiquihuitlán, according to Jamieson 1982: 149. NB: Jamieson's spelling.

	AGREEMENT +3/1SG		**NTR**	**INCMPL.**
Inflectional Class (IC)	**NTR** +3 DEF. or 1 SG.	**INCMPL** +3 DEF. or1 SG.	AGREEMENT -3/1SG	
A	PV *be-*	PV *kue-*	PV *be-*	PV *kue-*
B	PV *ba-*	PV *kua-*	PV *ba-*	PV *kua-*
C	PV *bo-*	PV *sko-*	MV *cho-*	
D	PV *bu-*	PV *sku-*	MV *chu-*	
E	MV *ju-*	PV *sku-*	MV *chju-*	
F	MV *ji-*	PV *ski-*	MV *chji-*	
G	MV *fa-*	MV *kjua-*	MV*fa-*	MV *kjua-*
H	CAUS *tsi-*	CAUS *tsi-*	STAT *nin-*	
I	STAT *su-*	STAT *su-*	STAT *nun-*	
J	PV *bu-*	PV/MV *ku-*	PV *bu-*	PV/MV *ku-*
K	PV *ba-*	PV *kua-*	MV*cha-*	
L	PV/MV *ka-*	*ska-*		
M	MV *fa-*	MV *kjua-*	STAT *nan-*	
N	PV *ba-*	PV *kua-*		
O	MV *bi-*	PV/MV *kui-*	MV *bi-*	PV/MV *kui-*
P	PV *bu-*	PV *sku-*	STAT *ntu-*	
Q	MV *ji-*	STAT *si-*	MV *chi-*	MIDDLE *xi-*
R	MV *fa-*	MV *kjua-*	MV *chja-*	

Table 1.2. Jamieson's Chiquihuitlán Inflectional Class Prefixes taxonomy revisited, with functional & semantic/pragmatic reinterpretation. NB: ALFALEIM spelling[5].

[5] See also Regino, 1993 for basic principles of Mazatec orthography (especially Lowlands Mazatec, whereas the ALFALEIM is mostly referred to in the Mazatec highlands). Both spelling conventions (Regino's and ALFALEIM converge, except for such conventions as < *z* >, for the retroflex unvoiced affricate, in the latter). Although bilingual school teachers are still working hard at learning how to use these conventions properly, especially for tone, we can say that Mazatec is today a language with an adequate and widely accepted orthography, based on the Spanish *Dachsprache* or "roofing language" – see Myer & Benjamin Maldonado, 2010 about the political context of emerging spelling norms in Mexico and the Native Americas. We'll use here the current spelling conventions according to ALFALEIM, which can easily be adapted to any dialect. Why do we use a spelling convention to adapt it to various dialects, instead of aiming at a single standard norm? Juan Gregorio Regino's seminal Mazatec alphabet (1993) already included four Mid- and Lowland dialects. Mazatec codification highly relies on diasystemic description, and must necessarily take it into account, as the inner differences within the Mazatec diasystem are stronger than between Romance languages, reminding us of the degree of differences between Germanic languages (German vs. Dutch vs. Danish vs. Icelandic), or between Baltic languages (Latvian vs. Lithuanian). See Kirk, 1970 for rates of intelligibility between Mazatec dialects (e.g. Chiquihuitlán scores only 10% intelligibility with Jalapa de Diaz, 21% with San Pedro Ixcatlán, 30% with San Miguel Soyaltepec, 39% with Ayautla or Huautla; the San Lorenzo dialect scores 39% with Jalapa (Kirk, *op. cit.*, p. 209).

3.2. Phonological complexity (and simplexity)

As for many Otomanguean languages, the complexity of segmental inventories and phonological properties of categories, leads to flexible models: Mazatec may rank among the most consonantal languages in the world (see table 2.1 below, column 2.1.2), or on the contrary, its basic consonant inventory may well rank among the average, with only 16 inherited consonants (cf. table 2.2 below). One could even say that, at the segmental level, only consonant cluster constraints in this language are complex (as in table 2.1., column 2.1.1.). The basic syllable structure is of the CV type, with $CCCV^N$ patterns. Onsets may therefore be complex, and nuclei too (diphthongs), but there is no coda.

2.1.1. PIKE'S SYLLABIC CONSTITUENCY MODEL Huautla de Jiménez Mazatec consonant system, according to Pike & Pike 1947	**2.1.2. G&K'S LARYNGEAL FEATURES MODEL (THE VOICE QUALITY & ONSET** Huautla de Jiménez Mazatec consonant system, according to Golston & Kerhein, 1998 : 319[6]
t, ts, tʃ, tʂ, k ; *tʔ, tsʔ, tʃʔ, tʂʔ, kʔ ;* *htsʔ, htʃʔ ;* *ht, hts, htʃ, htʂ, hk ;* *nt, nts, ntʃ, ntʂ, nk ;* *ʔnt, ʔnts, ʔntʃ, ʔntʂ, ʔnk ;* *hnt, hnts, hntʃ, hntʂ, hnk ;* *ntʔ, nts, ntʃʔ, ntʂʔ, nkʔ ;* *nth, ntsh, ntʃh, ntʂh, nkh ;*	*t, ts, tʃ, tʂ, k ;* *(…)* *t̪, t̪s̪, t̪ʃ, t̪ʂ, k̪ ;* *nt, nts, ntʃ, ntʂ, nk ;* *n̪t, n̪ts, n̪tʃ, n̪tʂ, n̪k ;* *n̪t̪, n̪t̪s, n̪t̪ʃ, n̪t̪ʂ, n̪k̪ ;*
s, ʃ ; *sh, ʃh, sʔ ;* *m, n, ɲ ;* *ʔm, ʔn, ʔɲ ;* *mʔ, nʔ, ɲʔ ;* *mh, nh ;* *l, lʔ ; β, j ;* *βʔ, ʔj ;* *hβ, hj ;* *βh ; ʔβ, jʔ*	*s, ʃ ;* *m, n, ɲ ;* *m̪, n̪, ɲ̪ ;* *m̪, n̪, ɲ̪ ;* *β, j ; β̪, j̪ ;* *β̪, j̪* *Cj, Cw, Cγ* *(this* C_x *parameter eliminates diphthongs)*

Table 2.1. Two models for onset inventories in Huautla Mazatec, according to P&P 1947 and G&K 1998

[6] See also Golston & Kerhein, 2004, Kirk et al. 1993.

	LABIAL	CORONAL			DORSAL	GLOTTAL
		[alveolar]	[high]	[retroflex]		
STOPS	*(p)*	*t*			*k*	*’*
AFFRICATES		*ts*	*ch*	*z*		
FRICATIVES	*f*	*s*	*x*			*j*
RESONANTS NASALS	*m*	*n*	*ñ*			
LATERAL		*l*				
APPROX/GLIDES	*b (= /w/)*		*y*			

Table 2.2. The ALFALEIM (Alfabeto para la Lengua Indigena Mazateca) spelling model (1980-), mostly inherited from P&P 1947.

From now on, we’ll use the ALFALEIM spelling conventions for all data.

Notwithstanding readability criteria, this choice is motivated by a strong concern for availability of the ALMaz data (*Atlas Lingüistico Mazateco*, IUF & LabEx EFL) to the populations.

3. Modeling a diasystemic IC matrix

3.1. Diasystem

Question: how many diasystems are there in Mazatec? Two? => Highlands, Lowlands; Three? => Highlands, Midlands, Lowlands; Four? => Highlands, Midlands, Lowlands, Canyon (Mazatlán, Chiquihuitlán), etc.

We will assume that the IC taxonomy proposed in table 3 is adequate for most Highlands dialects. Instead, we suggest that Lowland dialects have a fairly more complex system, similar to the one we have already seen in table 1 above (Chiquihuitlán, as described by Jamieson, 1982).

In table 3 Roman numerals stand for main inflectional classes, whereas capital letters (with or without Arabic numbers) stand for sub-classes. As we saw, preverbs constitute the primary area of dialectal variation. In Table 3 they are first divided into three morphophonological classes according to whether the preverbal onset is a labial, a coronal, or a dorsal consonant.

MORPHOPHONOLOGICAL CLASS	PREVERBAL CLASS		PREVERB	SEMANTIC CORRELATES	CLASS OF EQUIPOLLENCE
Labial Onsets	I	A	*b’é-*	Collocative	No Split
		B	*ba-*		
		C1	*bi-*	Dynamic	
	II	C2	*ba-/bi-*		Split
		C3	*bá-/fa-*		
	III		*b’a-/ch’a-*	Collocative	
Coronal Onsets	IV		*sí-/ni-*	Causative	
	V		*tsjoa-*	Dation	No Split
Dorsal Onsets	VI		*kjoa-*	Movement	

Table 3 : Highlands Mazatec IC diasystem for verb inflection (ALMaz model)

In a few words, IC Huautla Mazatec and many highland varieties, divide verbal stems into three major morphophonological classes, indicated by Roman numbers: one with labial onsets (I, II, III), one with coronal onsets (IV, V) and another with dorsal onsets (VI). Subdivisions are due to semantic correlates: *collocative*, such as 'put', 'get' 'carry', 'hold'; *dynamic*, with directionals, such as 'go', 'pass', etc.), *causative* 'makes', dation 'gives', movement "comes', and major morphological subclasses of equipollence, i.e. conflation. In the conflative paradigms, the preverb formative selection makes a split between the 3 SG & PL and the 1SG versus all other persons, which conflate in a stem selection item obtained by suppletion. In table 6, classes IA-C1 do not split, whereas IC IIC2-IV do split: IC IIC2 *ba-/bi*, IIC3 *bá-/fa-*, III *b'a-/ch'a-*, *sí-/ni-*, IV are conflative.

The labial onset class includes three preverbal classes: I, II, and III. IA (*b'é-*), IB (*ba-*), and IC1 (*bi-*), which share the property of not being split, that is of being constant across all Person-Number values in the paradigm.

In terms of semantic correlates, IA and B are collocative, as the preverbs' meaning *qua* verbal lexemes is roughly "to put," while IC1 has the same dynamic correlate as Class II preverbs. It is certainly useful here to recall the already given warning: not all verbs sharing a preverb share the preverb's "basic" meaning in an obvious or even retrievable way. This is why we speak of "correlate," that is, a fluid connection that may help us to understand the diachronic origin of the preverb class, but should not be used to characterize it synchronically as a whole.

Class II of the labial onset class consists of IIC2 and IIC3, both split – i.e. contrasting one form for 3 and 1SG and another form for the remaining cells – and semantically dynamic. Class IIID, also split, has the same collocative semantic correlate as IA-B. The coronal onset class includes two preverbal classes, both split: IV and V. The former is semantically causative, and it is certainly the one where the form-meaning connection is the most regular. We call class V "donative" as *tsjoa* by itself means 'to give'. Finally, there is only one diasystemic class with a dorsal onset: G, split, with a semantic correlate that may be defined as "movement".

As such, class VI is a compromise with the former morphophonological classes proposed by Brian Bull (1984), but it is so underrepresented in the lexicon that we could end this taxonomic matrix with class V, or class F. Table 3 condensates most of the strategic information about inflectional classes in Mazatec: Kenneth Pike's main light verbs (Pike, 1948), Jamieson's preverb paradigms (1982), Brian Bull's stem initial onsets (*op. cit.*) – labial, coronal and dorsal. It holds preverbal classes, semantic correlates and classes of equipollence of the "split type" or not.

This multidimensional matrix makes the system more understandable as a TAMV grid applied to natural language, according to directional, causative and donative preverbs. Dialects vary more on some points of these patterns than others: e.g., Northwestern Highland dialects of Mazatec tend to have more split classes of equipollence, and to combine preverbs differently. In all dialects, morphophonological rules on the stem initial labial onsets are valid.

3.2. Diasystemic variables

Stems

RSS

- Anticonflative (aC), ex. IIC2 > IB, IC1: {ba*X* / bi*X*} > {ba*X*}
- IC Shift (ICS), ex. IA > IC1: {b'é*X* } > {bi*X*}: .
- Conflative split (CS), ex. IA > IIC2: {b'é*X*} > {b'é*X* / bi*X*}
- Conflation breaking (CB): more PV stems, more PV…
- PV string complexification (SCx)
- Morphophonological conflation transfer (MPCT), ex. San José Independencia, {b'é*X*} > {b'é*X* / 'é*X*}
- Incompletive overmarking (IOM)

Endings

RE

- AGRS suffixal endings (SSE)

- AGRS enclitic ending (SEE)

MPR

- AGRS indexical (desinential) endings (SIE)

Tone

RSS

- Four tones level system (4TL) & two contour tone (BM, Bh) at Root Tone Class (RTC) level
- Lexical (i.e. Root) tone robustness (LTR)
- Preverb lexical tone (PvT) and its integration in PvT-Root Tone Strings (**PvRTS**) at stem level
- AGRS 1SG & incompletive downstep (S1IDS)[7]

RSS = Rules of Stem Selection
RE = Rules of Exponence
MPR = Morphophonological Rules

Mazatec dialects differ in their way of associating IC parameters, as suggested in the graph in figure 2 below (links here are only abstract: they will interact differently for each variety):

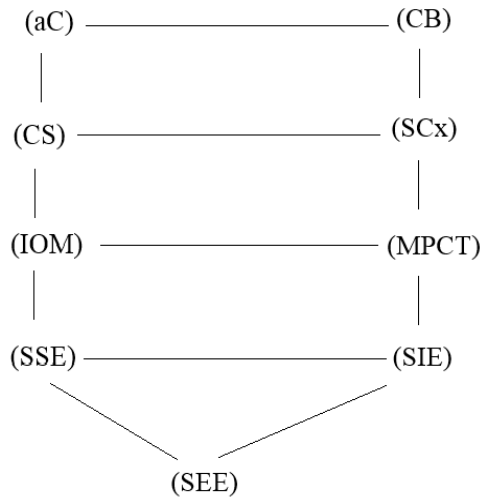


Figure 2. IC diasystemic parameters as a graph of potential links (tone not considered here)

4. A PFM approach at diasystemic complexity

In terms of Complexity Theory (henceforth, CT), we will now handle some aspects of *algorithmic complexity* for the description of the Mazatec verb. We will start first with Pike's data on Huautla Mazatec (Pike, *op.cit*.), before extending the model to other dialects. Then we'll handle *algorithmic complexity* through PFM (Paradigm Function Morphology, see Stump, 2001).

Verbs with <b'é-> stems (**Class IA** in table 3 above)

4.1. Neutral paradigms (i. e. neutral aspect)

From Table 9a in Pike (1948:111) we extract the following paradigm of a verb with High (H) root tone, for the neutral paradigm (i. e. neutral aspect):

	Singular	Plural	

[7] Cf. Jamieson, 1982, Baerman, 2014.

3	*b'éxá*		
1	*b'eexáa*	*b'èxá*	inclusive
		b'èxái	exclusive
2	*b'èxái*	*b'èxáo*	

Table 4. Neutral paradigm of HU XÁ 'to order /send to fetch' (Pike 1948:111).

Two things immediately come out from table 4: (i) the H tone on the root is stable across cells; (ii) the stem includes a <b'é-> prefix at all values of AGR, diversely toned according to these values. From (i) we conclude that H is inherent to the root; from (ii), that the lexeme XÁ belongs to Class IA. We therefore write the following *rules of stem selection* (RSS):

Rule Block I: Stem Selection (RSS)
(1) RSS: ⟨XÁ⟩ & σ { } ⇔ <b'èxá>
(2) RSS: ⟨XÁ⟩ & σ {AGR *3*} ⇔ <b'éxá>
(3) RSS: ⟨XÁ⟩ & σ {AGR *1SG*} ⇔ <b'eexá>

Rule (1) is the default rule: when no specified value of AGR is associated to the lexeme XÁ and no tense value is specified, as none need be in the Neutral (see above). The lexeme then shows up as the stem <b'èxá>, with h tone on the IC prefix. It is that stem which appears at all persons but 3 and 1SG. Rules (2) and (3) are more specific as they return stems for 3 and 1SG, namely <b'éxá> with H on the prefix for 3SG/PL and <b'eexá> with an LM contour on the prefix for 1SG.

Given the *a priori* unpredictability of IC assignment (except when it has semantic content, which is not the case here), RSC's (Root Stem Choice) are better written lexeme by lexeme. Insofar as all verbs that inflect like XÁ build up an inflection class arbitrarily labelled A, we can then introduce a metarule such as (4):

(4) If a verbal lexeme L belongs to Class A, then (a) ⟨L⟩ & σ { } ⇔ <b'è$\mathfrak{R}$>; (b) ⟨L⟩ & σ {AGR *3*} ⇔ <b'é$\mathfrak{R}$>; (c) ⟨L⟩ & σ {AGR *1SG*} ⇔ <b'ee$\mathfrak{R}$>; where $\mathfrak{R}$ is the root manifesting L.

Once the stems have been selected, we must account for the word-forms including the AGR endings. This we do by means of the following block of rules of exponence that we assume to be valid for all inflection classes, to the exception of so-called uninflected verbs:

(5a) Rule Block II: Rules of Exponence (RE)

i. RE: X_{Vb} σ {AGR {PERS *3*}} ⇒ $X^{Ti\text{-}\mathfrak{R}T}$
ii. RE: X_{Vb} σ {AGR {PERS *1*, NUM *sg*}} ⇒ $X^{Ti\text{-}\mathfrak{R}T}\oplus a$
iii. RE: X_{Vb} σ {AGR {PERS *2*, NUM *sg*}} ⇒ $X^{Ti\text{-}\mathfrak{R}T}\oplus i$
iv. RE: X_{Vb} σ {AGR {PERS *1*, NUM *pl*, INCL +}} ⇒ $X^{Ti\text{-}\mathfrak{R}T}\oplus \grave{a}$
v. RE: X_{Vb} σ {AGR {PERS *1*, NUM *pl*, INCL –}} ⇒ $X^{Ti\text{-}\mathfrak{R}T}\oplus \underline{i}$
vi. RE: X_{Vb} σ {AGR {PERS *2*, NUM *pl*}} ⇒ $X^{Ti\text{-}\mathfrak{R}T}\oplus o$

(5b) Rule Block III: Morphophonological Rules (MPR)
i) MPR <$_{\mathfrak{R}}$ ... V_i>V_i> → <... V_i>
ii) MPR <$_{\mathfrak{R}}$ ... V_i{A} >V_j{I}> → <... V_i{IA} >
Ex. SMH, SMJ, SLZ, etc. 2SG *b'èxá⊕i* > *b'èxé*, *bixë* (San Lucas Zoquiapam)

We extend (5b) to all vowel qualities (V_i) since there are empirical reasons to do so, as we shall see presently. And we assume it applies to 1SG *beexáa* as well, where <áa> actually transcribes the fusion of root H /á/ and M 1SG /a/ resulting in an HM contour.

These rules constitute a Rule Block, insofar as they apply disjunctively, i.e. applying one excludes applying any of the others (Stump 2001:33). In the input, X_V is a verb stem belonging to any class. The outputs are actual word-forms. T_i is the tone or tone contour borne by the prefix depending on the verb's inflection class and relevant RSS. $\mathcal{R}T$ is the lexically specified root tone. Recall that no diacritic above or under a vowel means mid tone.

Tonal "perturbations," to use Pike's term, at the juncture of the stem and the AGR endings – entailing that the latter do not always show up with their "basic" tones – can be accounted for by morphophonological rules (MPR's) for which the rules of exponence are indexed. The root tone is a constant, and there is almost no variation in the tonology of the IC prefix at the various values of AGR. We shall therefore assume, as just mentioned (and subject to possible revision), that Rule Block I accounts for AGR endings across the board.

4.2. Rule block for Tense-Aspect (TA) exponence

Rule Block I does not mention any TA value, thereby implying that the forms it describes bear the default for this feature, i.e. Neutral. We therefore posit another block, Rule Block II, for nondefault values of TA. Unlike Rule Block I, which applies across the board, Rule Block II must be broken down into disjunctive subblocks according to inflection class since, as we shall see, TA exponence varies along this parameter. In (6) we give the subblock for Classes I & II verbs:

(6) Rule Subblock $\text{II}^{\text{IC I-II}}$

i. RE: $X_{\text{I-II}}$ σ {TA *COMPL*} ⇔ *tsaki*⊕*X**
ii. RE: $X_{\text{I-II}}$ σ {TA *INCOMPL*} ⇔ <ko⊕*X**>

NB: X_I = IC I & II (with labial onsets)

A simplex pattern for PV-Root tone adjustment &/or marking in Huautla Mazatec (HU):

Pike's Tone Class I	IC	Root
3	H	H
1SG	LM	
others	h	

Table 5.1. PV-Root tone preservation vs. sandhi in Huautla Mazatec (according to Pike, *op. cit*.)

	IC	Root
3	LEX	LEX
1SG	Low-Mid	
others	LEX Modif.	

Table 5.2. IC and Root tone adjustments or marking, acording to Pike's data (HU).
LEX = Lexical tone, Modif = Modified tone (variably neutralized).

(7) RSS: ⟨SKÁ⟩ & σ {A *INCOMPL*} ⇔ <s<u>i</u>ská>
(8) RSS: ⟨SKÁ⟩ & σ {TA *INCOMPL*, AGR *3*} ⇔ <s<u>i</u>ìská>
(9) RSS: ⟨SKÁ⟩ & σ {TA *INCOMPL*, AGR *1SG*}⟩) ⇔ <s<u>ii</u>ská>

(10) If a verbal lexeme L belongs to **Class IV** (according to table 3, above), then (a) ⟨L⟩ & σ {TA *INCOMPL*} ⇔ <s̲iℜ>; (b) ⟨L⟩ & σ {TA *INCOMPL*, AGR *3*} ⇔ <s̲iìℜ>; (c) ⟨L⟩ & σ {TA *INCOMPL*, AGR *1SG*} ⇔ <s̲iiℜ>; where ℜ is the root manifesting L.

5. A first contrastive case study: San Miguel Huautepec

A very interesting dialect, as a subdialect of the central highland area, i.e. a subdialect of Huautla, easily comparable to Pike's 1948 data.

Except sets of cognates in Kirk (1966), the San Miguel Huautepec has never been documented.

The diatopic and diastratic situation of this dialect is that of a peripheral variety, though strongly influenced by the Huautla dialect, which is endowed with much prestige throughout the Mazatec highlands. It is not easy to find reliable (bilingual) speakers of the local variety, nowadays. Monolinguals' speech is still to be systematically recorded and documented. Here we will use materials from a formal elicitation of verbal paradigmes, with a middle-aged school teacher, whose idiolect can be considered as a reliable sample of average San Miguel Huautepec speech. Data recorded in August 2013 by Jean Léo Léonard (Paris 3) & Jaime Calderon (UNAM), in San Miguel Huautepec, transcribed by Julien Fulcran (Lille 3).

	3 Sg. & Pl.	1 Sg. -a	2 Sg. -i	2 Pl. -o	1 Pl. excl. -i	1 Pl. incl. -à
NTR	*b'èxá*	*b'exá*	*b'èxái*[8]	*b'èxó(i)*	*b'exài*	*b'exá*
COMPL	*tsak'èxá*	*tsak'exá*	*tsak'èxái*	*tsak'èxói*	*tsak'exái*	*tsak'èxá*
INCOMPL	*kuèxá*	*kuexá*	*kuéxái*	*kuéxói*	*kuéxái*	*kuéxá*
IMP 2 SG			*tèxái*			
JUSSIVE 3	*katabèxá*					

Table 6.1. **Class IA** *b'èxá* '(s)he sends to fetch, (s)he orders (someone to do something)', San Miguel Huautepec

	3 Sg. & Pl.	1 Sg. -a	2 Sg. -i	2 Pl. -o	1 Pl. excl. -i	1 Pl. incl. -à
NTR	*bèñón*	*beñòá*[9]	*beñòí*	*beñón*	*beñóin*	*bèñòá*
COMPL	*tsak'èñón*	*tsak'èñòá*	*tsak'èñòí*	*tsak'èñón*	*tsak'eñóin*	*tsak'eñòá*
INCOMPL	*kuèñón*	*kueñòá*	*kuèñòí*	*kuèñón*	*kuèñóin*	*kuèñòá*

Table 6.2. **Class IA** *bèñón* '(s)he braids (so), ties up', San Miguel Huautepec

A ***simplex pattern*** accounting for PV-Root tone adjustment &/or marking in San Miguel Huautepec Mazatec

Pike's Tone Class I	IC	Root
3	h	H
1SG	M	

[8] With an allophonic trend towards vowel fusion: *ai* > *e*, with few or no morphological consequences (i.e. *allophony* proper, without *allomorphy*).
[9] With *allophonic* gliding, and a trend to fronting of the mid back nucleus: *beñòá* > [beñüá].

others	h	

Table 5.3. Patterns of PV-Root tone adjustment &/or marking in San Miguel Huautepec Mazatec

	IC	Root
3	OCP	LEX
1SG	Neutral	
others	OCP	

Table 5.4. Constraints on PV-Root tone adjustment &/or marking in San Miguel Huautepec Mazatec. NB: OCP = Obligatory Contour Principle (cf. McCarthy, 1986).

NB: neutralization of the **PV Incompletive General Downstep**, following constraints sketched in Table 5.4. Unlike Huautla, the incompletive cells follow a conflative pattern, as the NTR. In Huautla, incompletive stems are anticonflative (e.g. *siskái* instead of *nìskái*)

	3 Sg. & Pl.	1 Sg. -a	2 Sg. -i	2 Pl. -o	1 Pl. excl. -i	1 Pl. incl. -à
NTR	*síská*	*sìská*	*nìskái*	*nìskó*	*nìskái*	*níská*
COMPL	*kisìská*	*kisìská*	*kinìskái*	*kinìskó(i)*	*kinìskái*	*kinìská*
INCOMPL	*síská*	*sìská*	*niskái*	*niskó*	*niskái*	*niská*

Table 7. **Class IV** *síská* ‘(s)he plays’, San Miguel Huautepec

	3 SG. & PL.	**1 SG. -A**	**2 SG. -I**	**2 PL. -O**	**1 PL. EXCL. -I**	**1 PL. INCL. -A**
NTR	*b'ènama*	*b'ènamaa*	*b'enamai*	*beyomo*	*beyomai*	*beyomà*
COMPL	*tsakenama*	*tsakàtenama*	*tsakenamai*	*tsakeyomò*	*tsakeyomai*	*tsakeyomà*
INCOMPL	*kuénòma*	*kuàtenama*	*kuenamai*	*kueyomò*	*kueyomai*	*kueyomà*
IMP 2 SG			*tenamai*			
JUSSIVE 1		*tiyomòàn*				
JUSSIVE 3	*katab'enama'a*					
	tínàma'a					

Table 8. **Class IA** modified *b'ènama* ‘(s)he hides’, San Miguel Huautepec.

	3 Sg. & Pl.	1 Sg. -a	2 Sg. -i	2 Pl. -o	1 Pl. excl. -i	1 Pl. incl. -à
NTR	*bèñii*	*beñoa*	*beñoi*	*beñon*	*beñoi*	*beñoa*
COMPL	*tsakèñi*	*tsakeñoa*	*tsakeñi*	*tsakeñon*	*tsakeñi*	*tsakeña*
INCOMPL	*kueñi*	*kueña*	*kueñoi*	*kuèñon*	*kuéñin*	*kuèña*
IMP 2 SG			*tèñe*			
JUSSIVE 1				*tèñon*		
JUSSIVE 3	*katab'èñi*					
PASSIVE	*kiseñi*					

Table 9. **Class IA** *bèñii* ‘(s)he buries’, San Miguel Huautepec.

	3 Sg. & Pl.	1 Sg. -a	2 Sg. -i	2 Pl. -o	1 Pl. excl. -i	1 Pl. incl. -à
NTR	*b'èse*	*b'èse*	*b'esai (1)*	*b'eso*	*b'esai*	*b'esee*
COMPL	*tsakèse*	*tsakese*	*tsakesai*	*tsakeso*	*tsakesai*	*tsakesee*
COMPL		*tsakese*	*tsakesai*	*tsakeso*		

INCOMPL	*kuese*	*kuese*	*kuésai*	*kuéso*	*kuèsai*	*kuésee*
IMP 2 SG			*tèsai*			
IMP 2 PL				*tèso*		
JUSSIVE 3	*katab'èse*					
PASSIVE	*sese*					

Table 10. **Class IA** *bèse* ‘(s)he whistles’, San Miguel Huautepec.

	3 sg. & pl.	1 sg. -a	2 sg. -i	2 pl. -o	1 pl. excl. -i	1 pl. incl. -a
NTR	fì	tifia (12)	timii	timankión (14)	timankì	timankià
COMPL	ki	kia	kiin	tsankiòn	tsankìn	tsankià
INCOMPL	kuei (1)*	kjioá (2)	kuín	kuankíon	kuankìn	kuankià
IMP 2 SG			tin			
IMP 2 PL				tankíon		
JUSSIVE 3	katafì					
PASSIVE	kafì					

Table 11.1. **Light Verb Class (LVC)** *fi* ‘s(he) goes)’, San Miguel Huautepec.

	3 Sg. & Pl.	1 Sg. -a	2 Sg. -i	2 Pl. -o	1 Pl. excl. -i	1 Pl. incl. -à
NTR	*fa'a*	*fa'a*	*b'itjai*	*b'itjao*	*b'itjai*	*b'itjaa*
		fà'a	*b'itjai*	*b'itjao*		*b'itjaa*
COMPL	*ja'a*	*jà'a*	*jitjai*	*kitjao*	*kitjai*	*kitjoa*
INCOMPL	*kua'a*	*kua'a*	*kuitjài*	*kuitjào*	*kuitjài*	*kuitjà*
INCOMPL			*kuitjái*			
IMP 2 SG			*titjài*			
IMP 2 PL				*titjò / titjó*		
JUSSIVE 3	*katafa'a*					
PASSIVE	*a'a*					

Table 11.2. **Light Verb Class** *fa'a* ‘s(he) passes)’, San Miguel Huautepec.

	3 sg. & pl.	1 sg. -a	2 sg. -i	2 pl. -o	1 pl. excl. -i	1 pl. incl. -a
NTR	*b'a*	*ya / ia*	*kichai*	*kichao*	*kichai*	*kichà*
COMPL	*tsaka*	*tsaka*	*kichai*	*kichao*	*kichai*	*kichà*
INCOMPL	*kua*	*kua*	*chai*	*chao*	*chai*	*chaà*
IMP 2 SG			*chai*			
IMP 2 PL				*chao*		
JUSSIVE 3	*katab'a*					
PASSIVE	*kiko(i)*					

Table 11.3. **Light Verb Class** *b'a* ‘s(he) carries, takes away’, San Miguel Huautepec.

NB: This LVC, in table 11.3., matches IC IIID from table 3 above.

Simplexity trends:

In the San Miguel Huautepec subdialect of the Highlands Central dialect (HU), the voice quality correlation is weak, with a strong trend towards modalization. Inflectional tonal contrasts are very subtle, without clear-cut downstep, relying on neutralization of marked lexical tones, in the “marking contour paradigms” (1SG & INCOMPL). Tonal levelling (e.g. 1PLexcl. =*jin* is no more Low) seems also to be active. The PV system is the same as HU, except for a few quasiderivative stem alternations.

Complexity:

In the SMH variety, the stem vowel allophony with Palatal subject agreement endings (2Sg *-i*, 1PlExcl. *-(j)in*) – undergoes a few morphophonological rules – MPR – (or constraints) more. Fine-grained tonal contrasts in stems may be cognitively difficult to acquire and to follow up.

6. A second contrastive case study: Santa Maria de la Asuncion (Santa Maria Jiotes)

6.1. Asymmetries in IC class IV conflation

The subdialect of Santa Maria de la Asuncion (Santa Maria Jiotes) also belongs to the Central Highland network (HU): it is famous among Mazatec highlanders for its "singing tone," for which it enjoys some prestige. HU speakers find it easy to understand, but "somewhat different".

Data: recording by JLL & JC (data from two speakers in their thirties), August 2013, transcription by Julien Fulcran.

	3 sg. & pl.	1 sg. -a	2 sg. -i	2 pl. -o	1 pl. excl. -i	1 pl. incl. -a
NTR	*síská*	*sìská*	*siskái*	*siskáòn*	*siskájin*	*siská*
COMPL	*kìsíská*	*kìsìská*	*kìnískái*	*kìníská(o)*	*níská(j)in*	*níská*
				kìnískán		
INCOMPL	*kuesíská*	*kuísíská*	*kuínìskái*	*kuániská(o)n*	*kuániská(j)in*	*kuìniská*

Table 12.1. **Class IV** *síská* '(s)he plays', Santa Maria Jiotes.

Rule Subblock I^{IV}

(11) RSS: ⟨SKÁ⟩ & σ {A *INCOMPL*} ⇔ HU <siská>, SMJ < *síská* >

(12) RSS: ⟨SKÁ⟩ & σ {TA *INCOMPL*, AGR *3*} ⇔ HU <sìská>, SMJ < *síská* >

(13) RSS: ⟨SKÁ⟩ & σ {TA *INCOMPL*, AGR *1SG*}⟩) ⇔ HU <siiská>, SMJ < *síská* >

NB: only one form in SMJ: < *síská* >, neutralization of the INCOMPL downstep.

(14) Rule Subblock II^{IV}

i. RE: X_{IV} σ {TA *COMPL*} ⇔ HU, SMJ *ki*⊕*X**

ii. RE: X_{IV} σ {TA *INCOMPL*} ⇔ HU *X**, SMJ *kue-*, *kuí-*, *kuá*⊕*X**>

Striking phenomena:

- The causative IC (IV) is not conflative in the neutral aspect, but it is still conflative in the completive: should we have doubts about the data? What is the speaker's proficiency in Mazatec exactly?
- The intricate allomorphy of the incompletive – and the fact that it also conflates, as in the NTR aspect paradigm. In HU, the NTR and completive aspects are conflative, whereas the incompletive is not. Here it works the other way round, with the NTR being non conflative.

Though, both phenomena repeat in table 12.2. below:

	3 sg. & pl.	1 sg. -a	2 sg. -i	2 pl. -o	1 pl. excl. -i	1 pl. incl. -a
NTR	*sítsjò*	*sìtsjuà*	*sitsjoí*	*sitsjò*	*sitsjoí(j)in*	*sitsjoá*
					nítsjoí(j)in	*sitsjuá*
COMPL	*kìsítsjó*	*kìsìtsjoá*	*kìnítsjoí*	*kìnítsjó*	*kínítsjoí(j)in*	*kìnítsjoá*
						kìnítsjuá

INCOMPL	*kuasítsjó*	*kuísìtsjoá*	*kuínitsjoí*	*kuánitsjó*	*kuánitsjoí(j)in*	*kuánitsjoá*
						kuánitsjuá

Table 12.2. **Class IV** *sítsjò* ‘(s)he crisps’, Santa Maria Jiotes.

6.2. Tone simplexity in the SMJ subdialect

	3 sg. & pl.	1 sg. -a	2 sg. -i	2 pl. -o	1 pl. excl. -i	1 pl. incl. -a
NTR	*béntjé*	*bèntjé*	*béntjái*	*béntjó*	*béntjái(j)in*	*béntjái*
						béntjé
COMPL	*tsákéntjé*	*tsàkèntjé*	*tsákéntjái*	*tsàkéntj(a)ó*	*tsàkéntjái(j)in*	*tsàkéntjé*
INCOMPL	*kuakèntjé*	*kuíkèntjé*	*kuíkèntjài*	*kuákèntj(a)ó*	*kuákèntjái(j)in*	*kuákèntjé*

Table 13.1. **Class IA** *béntjé* ‘(s)he wraps’, Santa Maria Jiotes.

	3 sg. & pl.	1 sg. -a	2 sg. -i	2 pl. -o	1 pl. excl. -i	1 pl. incl. -a
NTR	*béchuá*	*bèchuá*	*béchuái*	*kuéch(a)ó*	*kuéchuá(j)in*	*kuéchuá*
COMPL	*k'ábéchuá*	*k'ábèchuá*	*k'ábèchuái*	*k'àbéchu(a)ó*	*k'ábéchuá(j)in*	*k'ábéchuá*
INCOMPL	*kuakéchuá*	*kuíkéchuá*	*kuíkéchuái*	*kuákéch(a)ó*	*kuákèchuá(j)in*	*kuákèchuá*

Table 13.2. **Class IA** *b’etsoa* ‘(s)he closes’, Santa Maria Jiotes.

A ***simplex pattern*** accounting for PV-Root tone adjustment &/or marking in San Miguel Huautepec Mazatec (tables 5.5, 5.6 below):

Pike’s Tone Class I	IC	Root
3	H	H
1SG	h	
others	H	

Table 5.5. Patterns of PV-Root tone adjustment &/or marking in Santa Maria Jiotes Mazatec

	IC	Root
3	LEX	LEX
1SG	OCP	
others	LEX	

Table 5.6. Constraints on PV-Root tone adjustment &/or marking in Santa Maria Jiotes Mazatec
OCP = Obligatory Contour Principle.

NB: neutralization of the PV Incompletive General Downstep, with additive phenomena of variation (PV from M to H).

7. A third case study: San Francisco Huehuetlán (*an xo’boo*: Mazateco “Poblano”) San Francisco Huehuetlán

This is a very innovative peripheral Highland dialect. Its IC system converges strongly with Mazatlán and Midland and Lowland dialects, in spite of geographic distance. *Deterministic* and *epistemic complexity* are here at their highest, but we will cling to *algorithmic complexity* only, observing only paradigms and IC diversification mechanisms.

Data from an elicitation carried out in San Francisco Huehuetlán, section 2, October 2012, by Jean Léo Léonard (ALMaz), with a native speaker, farmer, aged 35.

Phonology: Mazateco Poblano undergoes a Pull & Drag chain (vowel shift), such as:

The San Lorenzo Vowel Shift:

**i > i, e; *e > a; *a > o; u > ʉ.*

NB: /ʉ/ is a high back unrounded vowel (spelled < ö > in ALFALEIM, but we will treat it here as more allophonic than phonemic, and we will consider that in the Huehuetlán subdialect, the input is /u/ instead of /ʉ/).

This vowel shift makes morphemes transparent when applied to postlexical forms. We will therefore present both forms (lexical and postlexical, i.e. inputs and outputs) when necessary.

/be'ín / *ba'ín* '(s)he braids, ties up'

AGRS	3 SG	1 SG	2 SG	2 PL	1 PL	3 PL
PRONOUNS	já	ngo'on	xi	ngo'ón	nguijin	niòn
NTR	*ba'ín*	*tába'nióon*	*téba'níi(n)*	*tá'níon*	*tátsuba'nìin*	*ínyoba'ín*
COMPL	*ya'ún*	*ya'nión*	*kika'ún*	*kika'niójon*	*kika'nújin*	*ya'ín*
INCOMPL	*kuà'ín*	*kuà'nióon*	*ta'íin*	*kuà' níojon*	*kuà'(n)ujin*	*kuà'ín*

Table 14.1. **Postlexical forms**. Class IA *ba'ín* '(s)he braids, ties up', San Francisco Huehuetlán.

AGRS	3 SG	1 SG	2 SG	2 PL	1 PL	3 PL
NTR	*be'iún*	*tébe'nión=an*	*tíbe'níún-i*	*té' nión=jún*	*tétsube'nù=jin*	*ínyabe'ín*
COMPL	*ye'iún*	*ye'nión=an*	*kike'níún-i*	*kike'nión=jún*	*kike' nù=jin*	*ye'ín*
INCOMPL	*kuè'iún*	*kuè'nión=an*	*te'níún-i*	*kuè' níon=jún*	*kuè' nù=jin*	*kuè'ín*

Table 14.2. **Lexical forms** *ba'ín* '(s)he braids, ties up', San Francisco Huehuetlán.

	3 SG	1 SG	2 SG	2 PL	1 PL	3 PL
NTR	*ba'*	*tába'*	*téba'*	*tá'*	*tátsuba'*	*ínyoba'*
COMPL	*ya'*	*ya'*	*kika'*	*kika'*	*kika'*	*ya'*
INCOMPL	*kuà'*	*kuà'*	*ta'*	*kuà'*	*kuà'*	*kuà'*

Table 14.3. **PV selection** *ba'ín* '(s)he braids, ties up', San Francisco Huehuetlán.

	3 SG	1 SG	2 SG	2 PL	1 PL	3 PL
NTR	*tinósínxó*	*tasínxóo*	*tenioxé, tantènioxé*	*tatseniòxón*	*tatseniòxéen*	*ndiosínxó*
COMPL	*kisínxó*	*kisínxóo*	*kinioxé*	*kiniòxón*	*kinioxéen*	*kisínxó*
INCOMPL	*bisínxó*	*bisínxòo*	*tiki'anxe*	*tetiò(n)xón*	*tetió(n)xéen*	*bisínxó*

Table 15.1**. Class IV** (causative): *tinósínxó* '(s)he works', San Francisco Huehuetlán.

Rule Block I

(14) RSS: ⟨XÁ⟩ & σ { } ⇔ < niòxó >

(15) RSS: ⟨XÁ⟩ & σ {AGR *3*} ⇔ < sínxó >

(16) RSS: ⟨XÁ⟩ & σ {AGR *1SG*} ⇔ < sínxó >

(17) Rule Block II

i. RE: X_{Vb} σ {AGR {PERS *3*}} ⇒ $X^{Ti\text{-}\mathfrak{H}T}$

ii. RE: X_{Vb} σ {AGR {PERS *1*, NUM *sg*}} ⇒ $X^{Ti\text{-}\mathfrak{H}T}\oplus o$

iii. RE: X_{Vb} σ {AGR {PERS *2*, NUM *sg*}} ⇒ $X^{Ti\text{-}\mathfrak{H}T}\oplus i$
iv. RE: X_{Vb} σ {AGR {PERS *1*, NUM *pl*, INCL –}} ⇒ $X^{Ti\text{-}\mathfrak{H}T}\oplus in$
v. RE: X_{Vb} σ {AGR {PERS *2*, NUM *pl*}} ⇒ $X^{Ti\text{-}\mathfrak{H}T}\oplus ún$

	3 SG	1 SG	2 SG	2 PL	1 PL	3 PL
NTR	*tinósín-*	*tasín-*	*tenio-, tantènio-*	*tatseniò-*	*tatseniò-*	*ndiosín-*
COMPL	*kisín-*	*kisín-*	*kinio-*	*kiniò-*	*kinio-*	*kisín-*
INCOMPL	*bisín-*	*bisín-*	*tiki'-*	*tetió-*	*tetión-*	*bisín-*

Table 15.2**.** Postlexical forms, PV strings, Class IVE: *tinósínxó* '(s)he works', Huehuetlán.

/ba'atjà/ '(s)he plants » (HU *b'éntjè*)

16. 1. **Tokens**	3 SG	1 SG	2 SG	2 PL	1 PL	3 PL
NEUTRE	*kinoba'atjà*	*ba'atja*	*teba'itja*	*tatsuba'atja*	*tatsube'etja*	*yoba'atjà*
COMPLETIVE	*yàtja*	*yàtjaa*	*kikatje*	*kikatjaja*	*kikatjejin*	*yàtja*
INCOMPLETIVE	*kotikatja*	*bíkatjà*	*tikátje*	*tonguikátja*	*(tonguikátjejin)*	*kotikatja*
16. 1. **PV string output**						
NEUTRE	*kino-*	-	*te-*	*tatsu-*	*tatsu-*	*yo-*
COMPLETIVE	*yà-*	*yà-*	*kik-*	*kik-*	*kik-*	*yà-*
INCOMPLETIVE	*kotik-*	*bík-*	*tik-*	*tonguik-*	*(tonguik-)*	*koti-*
16. 1. **PV string inputs**						
NEUTRE	*kina-*	-	*ti-*	*tetsu-*	*tetsu-*	*ya-*
COMPLETIVE	*yè-*	*yè-*	*kik-*	*kik-*	*kik-*	*yè-*
INCOMPLETIVE	*katik-*	*bík-*	*tik-*	*tanguik-*	*(tanguik-)*	*kati-*

Table 16. IC IA: *ba'atjà* 's(he) plants': sets of forms, San Francisco Huehuetlán

(18) Indexation of San Francisco Huehuetlán IC parameters, on the basis of currently available data (ALMaz, 2012):

Stems
IC Shift **YES**
Anticonflative (aC) **NO**
Conflative split (CS) **YES**
Conflation breaking (CB) **YES**
PV string complexification (SCx) **YES**
Morphophonological conflation transfer (MPCT) **NO**
Incompletive overmarking (IOM) **YES**
Endings
AGRS suffixal endings (SSE) **YES**
AGRS enclitic ending (SEE) **YES**
AGRS indexical (desinential) endings (SIE) **YES**
Tone
Four tones level system (4TL) **NO**
AGRS1 & incompletive downstep (SID) **NO**
Lexical (i.e. Root) tone robustness (LTR) **NO**

NB: as to SFH, we should add a **Root Allomorphy (RA)** parameter to the list, in the **Stem** section (cf. tables 14.1-2. and 16.1. above).

8. Conclusion & prospects

This sketch with first hand data on three Mazatec highlands dialects highlights the following points:

The Empirical Gap: Mazatec **IC systems** are still waiting *diasystemic description and modeling*. Since Pike's (1948) and Jamieson's (1982, 1988) seminal work, most scholars have relied on second-hand data. Though, a rapid glimpse at some paradigms from two subdialects close to HU, and at a dialect still totally undescribed, like SFH, show that it is worth exploring more dialects.

Interacting Models: Mazatec IC systems might display an array of IC matrices, whose complexity varies from the Jamieson's 'Mendeleev table' (table 1 above) to the ALMaz model (table 3). The ALMaz model draws a synthesis between many competing models, such as Bull's (1984) morphophonological (cyclical rules) model (MPRM), Pike's (1948) Light Verb+Polyvalent Root Model (LVM) and Jamieson's (1982) IC Stems Model (ICSM).

Complexity Theory: All models are equally true and falsify each other. In Popperian terms: Bull's MPRM, Pike's LVM and Jamieson's ICSM, as conflated in table 3, make up a holographic model for IC building from the structural complexity standpoint ('bit complexity'). While complexity should be less intended as a metaphor or as qualifying intricate inflectional patterns, most phenomena resorting to stem building (PV, PV strings, broken and *conflation Split*) could be also simulated with the help of Game Theory (*Naming Game*) or *inflectional automata*. Paradigm interaction (at PF level) meets further PFM building blocks (RSS, RE, MPR), resulting in more than mere Neogrammarian analogy. As any system is at the same time less and more than the sum of its components, interaction between building blocks, morphophonological constraints, the lexicon and the discursive-referential factors make of the Mazatec IC diasystem much more – and also less than – the sum of its parts.

Algorithmic Complexity: PFM turns out to be a very handy tool in checking the degree of structural complexity and building block interaction. Some dialects put a heavy burden on RSS, others on MPR, others on RE. A few dialects, such as SFH, but also SMJ, challenge the model on its limits, though, PFM still works as a very heuristic tool for handling *bit complexity* through *algorithmic complexity*. It allows the identification of mechanisms or zones of simplexity too, as in syncretic RSS, or TAMV PV tone marking.

Simplexity & Ergonomy: Many components of the Mazatec IC system, which had been described or considered as complex, are actually simplex, as the Lexical Tone Taxonomy, which belongs to Pike's LVM. For Pike, Mazatec roots are polyvalent, and give the lexical tone of the stem. PV can only have a high or a mid-tone at lexical level (because of their highly lexical origin). Whatever the morphosyntactic concatenation may be with any PV or AGRS ending, they cling to their lexical tone pattern. Most changes happen elsewhere, especially in the PV complex, resorting to two simple phenomena: either an OCP adjustment (merely allotonic), or 1SG and INCOMPL marking by a contour of the downstep type. Further phenomena are *simplex* rather than *complex* too, such as Stem Tone Leveling (Mazatlán), according to which all tones assimilate partially or totally with the Root Tone.

Further Research: Encouraging results from the ALMaz project, such as the points we just raised in this conclusion, suggest that the development of Geolinguistic Documentation Procedure (GDP) of Otomanguean languages should be fostered in the years to come. As a subcomponent of the MAmP project, whose goal was to explore GDP and create tools in this direction, the ALMaz project telescoped

its initial goals. The drastic turn was when we serendipitously realized, in late 2011, that the Kirk's (1966) cognate list gave far less interesting results, even for isogloss mapping, than could lists of inflectional paradigms from Pike (1948). Verbal and Nominal Possession Paradigms (**VNPP**) provided more relevant phenomena than trivial diachronic variables. Moreover, from 140 verbs taken into account in Kirk's cognate sets, all were given in NTR 3SG, which tells us very little about IC variation. A whole continent was to discover underneath.

Cross-dimensional Complexity: VNPP *algorithmic complexity* may be even more efficient as hints to explore *deterministic* and *epigenetic complexity*, or make both domains of complexity interact. In other words, the SF Huehuetlán IC system tells us more on how this dialect emerged than standard diachronic variables from cognate sets. Mazatec VNPP geolinguistic mapping is, therefore, a genuine new horizon for Otomanguean linguistics. We should plan more VNPP linguistic atlases or at least diasystemic databases.

9. References:

Baerman, Matthews, 2014. "Inflectional class interactions in Oto-Manguean", *Patterns in Mesoamerican Morphology, Proceedings of the Meso-American Inflectional Morphology Workshop*, Paris 2011 (forthcoming).

Berthoz, Alain 2009. *La simplexité*, Paris, Odile Jacob.

Bull Brian 1984. "Aspect Formation of San Jeronimo Mazatec Verb", SIL-Mexico Workpapers, n° 6: 93-117.

Brun-Trigaud Guylaine; Pierre Darlu, Antonella Gaillard-Corvaglia & Patric Sauzet, 2014. « Exploration cladistique de l'ALLoc » au 10ème CIEO (Congrès de l'Associacion internacionala d'Estudis Occitans), 12-19 juin 2011, Béziers, in Alén Garabato Carmen, Claire Torreilles & Marie-Jeanne Verny (eds.), 2014 : *Actes du Xe Congrès de l'AIEO, Association Internationale d'Etudes Occitanes*, Limoges, Lambert Lucas, vol. 2: 489-500.

Casad, Eugene 1974. *Dialect intelligibility testing*, University of Oklahoma, the Summer Institute of Linguistics.

Gaillard-Corvaglia, Antonella ; Jean Léo Léonard ; Pierre Darlu, 2007. « Testing cladistics on dialect networks and phyla (gallo-romance and southern italo-romance) », Proceedings of *Ninth Meeting of the ACL Special Interest Group in Computational Morphology and Phonology*, pages 23–30, Prague, Juin 2007. Association for Computational Linguistics : http://www.aclweb.org/anthology-new/W/W07/W07-1304.pdf

Gaillard-Corvaglia, Antonella ; Pierre Darlu, Jean Léo Léonard, 2008. « Approche cladistique des vocalismes d'oïl à partir des données de l'A.L.F. », *Bollettino dell'Atlante Linguistico Italiano* III Serie - Dispensa N. 32 2008: 36-78.

Gaillard-Corvaglia, Antonella, 2012. *De la cladistique à la linguistique. Une étude appliquée aux dialectes italo-romans méridionaux et salentins*, Alessandria, Edizioni dell'Orso. Golston, Chris & Kehrein, Wolfgang 1998. "Mazatec onsets and nuclei", *International Journal of American Linguistics*, 64.4: 311-337.

Golston, Chris & Kehrein, Wolfgang 2004. "A prosodic Theory of laryngeal contrasts", *Phonology* 21, 1-33.< http://zimmer.csufresno.edu/~chrisg/>

Gribbin, John, 2004. *Deep Simplicity, Chaos, Complexity and the Emergence of Life*, London, Penguin.

Gudschinsky, Sarah 1958a. *Proto-Popotecan. A Comparative Study of Popolocan and Mixtecan*, *IJAL*, 25(2).

Gudschinsky, Sarah C. 1958b. "Mazatec dialect history: a Study in Miniature", *Language* 34, 469-481.

Jamieson, Allan & Jamieson, Carole 1978. *Mazateco de Chiquihuitlán*, Archivo de Lenguas Indígenas de México, 5, México, D.F., Centro de Investigación para la integración social.

Jamieson, Carole Ann 1982. "Conflated subsystems marking person and aspect in Chiquihuitlán Mazatec verb", *IJAL* 48(2), 139-167.

Jamieson, Carole Ann 1988. *Gramática mazateca del Municipio de Chuiquihuitlán, Oaxaca*. Mexico: Instituto Lingüístico de Verano, A.C.

Massip-Bonet, Àngels & Albert Bastardas-Boada (eds.), 2013. *Complexity Perspectives on Language, Communication and Society*, Springer.
Hruschka, Daniel & al. 2009. « Building social cognitive models of language change », *Trends in Cognitive Sciences*, 13-11: 464-469.
Kirk, Paul Livingston 1966. *Proto-Mazatec phonology*. PhD dissertation, University of Washington.
Kirk, Paul Livingston 1970. "Dialect Intelligibility Testing: The Mazatec Study", *International Journal of American Linguistics*, Vol. 36, 3 : 205-211.
Kirk P. L., Ladefoged J. & Ladefoged P., 1993. "Quantifying acoustic properties of modal, breathy and creaky vowels in Jalapa Mazatec", *American Indian Linguistics and Ethnography in Honor of Laurence C. Thompson*, Mattina A. & Montler T. (eds.), University of Montana, 435-40.
Léonard, Jean Léo ; dell'Aquila, Vittorio & Gaillard-Corvaglia, Antonella, 2012. "The ALMaz (*Atlas Lingüístico Mazateco*): from geolinguistic data processing to typological traits", *STUF*, Akademie Verlag, 65-1, 78-94.
Léonard, Jean Léo 2011. "A quoi reconnaît-on la tagmémique ? Entre structuralisme périphérique et grammaire de texte : essai de modélisation épistémologique", *Histoire, Epistémologie, Langage*, 32-3 (sous presse).
Léonard, J.L. 2010. "Enquêtes exploratoires pour l'ALMaz (*Atlas Lingüístico Mazateco*). Elicitation croisée, entre typologie et codification d'une langue otomangue", Grenoble, *Géolinguistique* 11 : 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 2014. "Mazatec verb inflection: A revisiting of Pike (1948) and a comparison of six dialects", *Patterns in Mesoamerican Morphology, Proceedings of the Meso-American Inflectional Morphology Workshop*, Paris 2011 (forthcoming).
McCarthy, John, 1986. "OCP Effects: Gemination and Antigemination", *Linguistic Inquiry* 17, 207-263.
Miestamo, Matti, Kaius Sinnemaki & Fred Karlsson (eds.), 2008. *Language Complexity: Typology, Contact, Change*. Amsterdam, John Benjamins.
Myer, Lois & Benjamin Maldonado, Alvarado (eds.), 2010. *New World of Indigenous Resistance. Noam Chomsky and Voices from North, South and Central America*, San Francisco, City Light Books.
Mufwene, Salikoko, 2013. *Complexity perspectives on language, communication, and society*, in Massip-Bonet & Bastardas-Boada: 197-218.
O'Sullivan, David, 2004. « Complexity Science and Human Geography », *Transactions of the Institute of British Geography*: 282-295.
Patriarca, M., X.Castello, Uriarte, J., Eguiluz, V. & Miguel, M. S., 2012. « Modeling two language competition dynamics », *Advances in Complex Systems*, 15: 125-48.
Mufwene, Salikoko, 2001. *The Ecology of Language Evolution*, Cambridge Approaches to Language Contact, Cambridge, CUP.
Pike, Kenneth 1948. *Tone Languages. A Technique for Determining the Number and Types of Pitch Contrasts in a Language, with Studies in Tonemic Substitution and Fusion*, Ann Arbor, University of Michigan Press. NB: Further editions: 1949, 56, 57, 61.
Pike, Eunice, 1956. "Tonally Differentiated Allomorphs in Soyaltepec Mazatec", *IJAL* 22: 57-71.
Pike, Eunice, 1967. "Huautla de Jiménez Mazatec", in McQuown, Norman (ed.) *Handbook of Middle American Indians*, 5 *Linguistics*, Austin, University of Texas Press : 311-330.
Pike, Kenneth & Pike; Eunice 1947. "Immediate constituents of Mazatec Syllables", *IJAL* 13:78-91.
Regino, Juan Gregorio 1993. *Alfabeto mazateco*, Oaxaca, Ciesas, IOC, CDCNC.
Stump, Gregory 2001. *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge, Cambridge University Press.