MODELING SIMPLEXITY VERSUS COMPLEXITY IN MAZATEC INFLECTIONAL CLASSES: A DIA SYSTEMIC 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 Huehuete lán data below), or the Midlands (Jalapa de Díaz, 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 Huautepac (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:

\[
\begin{array}{c|c|c}
\text{TENSE-ASPECT-VOICE} & \text{AGR} \\
\hline
\langle w \ (CV) \rangle & \langle s_t \ CV \rangle & \langle r \ (C)V\rangle V > \\
\end{array}
\]

**Single vs. complex TAMV chains:**

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**), 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.

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 explanatory 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 diysystemic 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*.


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

*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= sji- ská -a    
COMPL IC:3/1SG play 1SG
I played
(3a) ki= ni-ská -i
COMPL IC play 2SG
you played

**Mazatlán**

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

---

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:

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>-'a(n)</td>
<td>-nà, -jnà inclusive</td>
</tr>
<tr>
<td>1</td>
<td>-i, ji</td>
<td>-un, jún, jnú exclusive</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>-a</td>
<td>-a inclusive</td>
</tr>
<tr>
<td>1</td>
<td>-ijin,</td>
<td>exclusive</td>
</tr>
<tr>
<td>2</td>
<td>-i</td>
<td>-o</td>
</tr>
</tbody>
</table>

Table 4.2. HU AGR endings

NB: Pike’s -jn 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).
ely accepted orthography, based on the Spanish Mid-

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

stronger than between Romance languages, reminding us of the d

Table 1.2. Jamieson’s Chiquihuitlán Inflectional Class Prefixes taxonomy revisited, with functional &

semantic/pragmatic reinterpretation. NB: ALFALEIM spelling⁵.

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

such conventions as <

Table 1.1. Inflectional Class Prefixes in Chiquihuitlán, according to Jamieson 1982: 149.

Table 1.1. Inflectional Class Prefixes in Chiquihuitlán, according to Jamieson 1982: 149.

Table 1.2. Jamieson’s Chiquihuitlán Inflectional Class Prefixes taxonomy revisited, with functional &

semantic/pragmatic reinterpretation. NB: ALFALEIM spelling⁵.

NB: Jamieson’s spelling.
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 patterns. Onsets may therefore be complex, and nuclei too (diphthongs), but there is no coda.

<table>
<thead>
<tr>
<th>2.1.1. PIKE’S SYLLABIC CONSTITUENCY MODEL</th>
<th>2.1.2. G&amp;K’S LARYNGEAL FEATURES MODEL (THE VOICE QUALITY &amp; ONSET)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huautla de Jiménez Mazatec consonant system, according to Pike &amp; Pike 1947</td>
<td>Huautla de Jiménez Mazatec consonant system, according to Golston &amp; Kerhein, 1998 : 319</td>
</tr>
<tr>
<td>t, ts, tf, ts, k ; t?, ts?, tf?, ts?, k? ; hts?, htf? ; ht, hts, htf, hts, htk ; nt, nts, ntfs, ntsh, nk ; ?nt, ?nts, ?ntf, ?nts, ?ntk ; hnt, hnts, hntf, hnts, hntk ; nt?, nts, ntfs, ntsh, nk? ; nth, nth, ntsh, ntsh, nk ;</td>
<td>t, ts, tf, ts, k ; (…) ; t, ts, tf, ts, k ; nt, nts, ntfs, ntsh, nk ; nt, nts, ntfs, ntsh, nk ;</td>
</tr>
<tr>
<td>s, f ; sh, fh, s? ; m, n, n ; ?m, ?n, ?n ; m?, n?, n? ; mh, nh ; l, l? ; β, j ; β?, ?j ; hβ, hj ; βh ; ?β, j?</td>
<td>s, f ; m, n, n ; m?, n?, n? ;</td>
</tr>
<tr>
<td></td>
<td>Cj, Cw, Cγ (this Cx parameter eliminates diphthongs)</td>
</tr>
</tbody>
</table>

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

---

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 Lingüístico Mazateco, IUF & LabEx EFL) to the populations.

3. Modeling a diasytemic 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.

Table 3: Highlands Mazatec IC diasytem 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 ba-/fa-, III b’a-/ch’a-, si-/mi-, 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: {baX / biX} > {baX}
  - IC Shift (ICS), ex. IA > IC1: {b’éX} > {biX}.
  - Conflative split (CS), ex. IA > IIC2: {b’éX} > {b’éX / biX}
  - Conflation breaking (CB): more PV stems, more PV...
  - PV string complexification (SCx)
  - Morphophonological conflation transfer (MPCT), ex. San José Independencia, {b’éX} > {b’éX / ‘éX}
  - Incompletive overmarking (IOM)

#### Endings
- **RE**
  - AGRS suffixal endings (SSE)
- AGRS enclitic ending (SEE)
- AGRS indexical (desinential) endings (SIE)

**Tone**

- 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):

```
(aC) ---- (CB) ---- (SCx)
|       |       |
(CS) ---- (SCx) ---- (MPCT)
|       |       |
(IOM) ---- (MPCT)
|       |       |
(SSE) ---- (SIE) ---- (SEE)
```

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):

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
</table>

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'èxá> 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'èR⟩; (b) ⟨L⟩ & σ {AGR 3} ⇔ <b'éR⟩; (c) ⟨L⟩ & σ {AGR 1SG} ⇔ <b'eR⟩; where 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:

**Rule Block II: Rules of Exponence (RE)**

i. RE: Xvb σ {AGR {PERS 3}} ⇒ X^{Ti-gf}$
ii. RE: Xvb σ {AGR {PERS 1, NUM sG}} ⇒ X^{Ti-gf+o}
iii. RE: Xvb σ {AGR {PERS 2, NUM sG}} ⇒ X^{Ti-gf+i}
iv. RE: Xvb σ {AGR {PERS 1, NUM pl, INCL +}} ⇒ X^{Ti-gf+â}
v. RE: Xvb σ {AGR {PERS 1, NUM pl, INCL –}} ⇒ X^{Ti-gf+ê}
vi. RE: Xvb σ {AGR {PERS 2, NUM pl}} ⇒ X^{Ti-gf+ë}

**Rule Block III: Morphophonological Rules (MPR)**

i) MPR <[r] ... Vₐ>Vₐ> → <... Vₐ>
ii) MPR <[r] ... Vₐ{A}>Vₐ{1}]> → <... Vₐ{IA}>

Ex. SMH, SMJ, SLZ, etc. 2SG b'èxá+ë > b'èxé, bixë (San Lucas Zoquiapam)

We extend (5b) to all vowel qualities (Vₐ) since there are empirical reasons to do so, as we shall see presently. And we assume it applies to 1SG b'èxá 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, \(X_Y\) 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. \(\mathcal{RT}\) 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 II\(^{IC\text{-}II}\)

1. RE: \(X_{I\text{-}II} \sigma \{ TA \text{ COMPL} \} \Leftrightarrow tsaki \oplus X^*\)
2. RE: \(X_{I\text{-}II} \sigma \{ TA \text{ INCOMPL} \} \Leftrightarrow <ko \oplus X^*>\)

NB: \(X_i = IC\text{ I \& II} (\text{with labial onsets})\)

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

<table>
<thead>
<tr>
<th>Pike’s Tone</th>
<th>IC</th>
<th>Root</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td>3</td>
<td>H</td>
</tr>
<tr>
<td>1SG</td>
<td>LM</td>
<td>H</td>
</tr>
<tr>
<td>others</td>
<td>h</td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>IC</th>
<th>Root</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>LEX</td>
</tr>
<tr>
<td>1SG</td>
<td>Low-Mid</td>
</tr>
<tr>
<td>others</td>
<td>LEX</td>
</tr>
</tbody>
</table>

Table 5.2. IC and Root tone adjustments or marking, according to Pike’s data (HU).

LEX = Lexical tone, Modif = Modified tone (variably neutralized).

(7) RSS: \(<\text{ská}>\) & \(\sigma \{ A \text{ INCOMPL} \} \Leftrightarrow <\text{si}\text{ská}>\)

(8) RSS: \(<\text{ská}>\) & \(\sigma \{ TA \text{ INCOMPL}, AGR 3 \} \Leftrightarrow <\text{si}\text{ská}>\)

(9) RSS: \(<\text{ská}>\) & \(\sigma \{ TA \text{ INCOMPL}, AGR 1SG\}) \Leftrightarrow <\text{si}\text{ská}>\)
12

(10) If a verbal lexeme \( L \) belongs to **Class IV** (according to table 3, above), then (a) \( \langle L \rangle \) & \( \sigma \{ \text{TA INCOMPL} \} \Leftrightarrow \langle \text{si}\rangle \); (b) \( \langle L \rangle \) & \( \sigma \{ \text{TA INCOMPL, AGR 3} \} \Leftrightarrow \langle \text{si}\rangle \); (c) \( \langle L \rangle \) & \( \sigma \{ \text{TA INCOMPL, AGR 1SG} \} \Leftrightarrow \langle \text{si}\rangle \); where \( \mathfrak{R} \) 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).

<table>
<thead>
<tr>
<th></th>
<th>3 Sg. &amp; Pl.</th>
<th>1 Sg. -a</th>
<th>2 Sg. -i</th>
<th>2 Pl. -o</th>
<th>1 Pl. excl. -i</th>
<th>1 Pl. incl. -à</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTR</td>
<td>b’èxá</td>
<td>b’exá</td>
<td>b’èxáí</td>
<td>b’èxóí</td>
<td>b’exáí</td>
<td>b’exáí</td>
</tr>
<tr>
<td>COMPL</td>
<td>tsak’èxá</td>
<td>tsak’ exá</td>
<td>tsak’èxáí</td>
<td>tsak’èxóí</td>
<td>tsak’exáí</td>
<td>tsak’èxá</td>
</tr>
<tr>
<td>INCOMPL</td>
<td>kuèxá</td>
<td>kuexá</td>
<td>kuèxái</td>
<td>kuèxóí</td>
<td>kuèxáí</td>
<td>kuèxá</td>
</tr>
<tr>
<td>IMP 2 SG</td>
<td></td>
<td></td>
<td>tèxáí</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JUSSIVE 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>3 Sg. &amp; Pl.</th>
<th>1 Sg. -a</th>
<th>2 Sg. -i</th>
<th>2 Pl. -o</th>
<th>1 Pl. excl. -i</th>
<th>1 Pl. incl. -à</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTR</td>
<td>beñón</td>
<td>beñóí</td>
<td>beñóí</td>
<td>beñón</td>
<td>beñón</td>
<td>beñóí</td>
</tr>
<tr>
<td>COMPL</td>
<td>tsak’ènón</td>
<td>tsak’ènóí</td>
<td>tsak’ènóí</td>
<td>tsak’ènón</td>
<td>tsak’ènóin</td>
<td>tsak’ènóá</td>
</tr>
<tr>
<td>INCOMPL</td>
<td>kueñón</td>
<td>kueñóá</td>
<td>kueñóí</td>
<td>kueñón</td>
<td>kueñóin</td>
<td>kueñóá</td>
</tr>
</tbody>
</table>

Table 6.2. **Class IA** beñón ‘(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

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>h</th>
<th>Root</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pike’s</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tone Class</td>
<td>I</td>
<td>IC</td>
<td></td>
</tr>
<tr>
<td>3 ISG</td>
<td>M</td>
<td></td>
<td>H</td>
</tr>
</tbody>
</table>

---

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: beñóá > [beñüá].
Table 5.3. Patterns of PV-Root tone adjustment &/or marking in San Miguel Huautpec Mazatec

<table>
<thead>
<tr>
<th>Others</th>
<th>h</th>
</tr>
</thead>
</table>


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

Table 6.1. Class IV síská ‘(s)he plays’, San Miguel Huautpec

Table 6.2. Class IA bëni ‘(s)he hides’, San Miguel Huautpec.

Table 6.3. Class IA bëse ‘(s)he buries’, San Miguel Huautpec.
### Table 10. **Class IA** bèse ‘(s)he whistles’, San Miguel Huautepec.

<table>
<thead>
<tr>
<th>3 sg. &amp; pl.</th>
<th>1 sg. -a</th>
<th>2 sg. -i</th>
<th>2 pl. -o</th>
<th>1 pl. excl. -i</th>
<th>1 pl. incl. -a</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTR</td>
<td>tifia (12)</td>
<td>timii</td>
<td>timankión (14)</td>
<td>timankí</td>
<td>timankià</td>
</tr>
<tr>
<td>COMPL</td>
<td>kia</td>
<td>kiin</td>
<td>tsankiôn</td>
<td>tsankín</td>
<td>tsankià</td>
</tr>
<tr>
<td>INCOMPL</td>
<td>kuei (1)*</td>
<td>kjiò (2)</td>
<td>kuin</td>
<td>kuankion</td>
<td>kuankín</td>
</tr>
<tr>
<td>IMP 2 SG</td>
<td>tin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMP 2 PL</td>
<td>tankión</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 11. **Light Verb Class (LVC)** fì ‘(s)he goes’, San Miguel Huautepec.

<table>
<thead>
<tr>
<th>3 Sg. &amp; Pl.</th>
<th>1 Sg. -a</th>
<th>2 Sg. -i</th>
<th>2 Pl. -o</th>
<th>1 Pl. excl. -i</th>
<th>1 Pl. incl. -à</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTR</td>
<td>fa’a</td>
<td>fa’a</td>
<td>b’itjài</td>
<td>b’itjào</td>
<td>b’itjaa</td>
</tr>
<tr>
<td>COMPL</td>
<td>fà’à</td>
<td>b’itjài</td>
<td>b’itjào</td>
<td>b’itjaa</td>
<td></td>
</tr>
<tr>
<td>INCOMPL</td>
<td>ja’a</td>
<td>ja’a</td>
<td>jitjài</td>
<td>kitjào</td>
<td>kitjai</td>
</tr>
<tr>
<td>INCOMPL</td>
<td>kua’a</td>
<td>kua’a</td>
<td>kuitjài</td>
<td>kuitjào</td>
<td>kuitjái</td>
</tr>
<tr>
<td>IMP 2 SG</td>
<td>titjài</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMP 2 PL</td>
<td></td>
<td>tijjò / titjó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JUSSIVE 3</td>
<td>katafa’a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PASSIVE</td>
<td>a’à</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>3 sg. &amp; pl.</th>
<th>1 sg. -a</th>
<th>2 sg. -i</th>
<th>2 pl. -o</th>
<th>1 pl. excl. -i</th>
<th>1 pl. incl. -a</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTR</td>
<td>b’a</td>
<td>ya / ia</td>
<td>kichai</td>
<td>kichao</td>
<td>kichai</td>
</tr>
<tr>
<td>COMPL</td>
<td>tsaka</td>
<td>tsaka</td>
<td>kichai</td>
<td>kichao</td>
<td>kichai</td>
</tr>
<tr>
<td>INCOMPL</td>
<td>kua</td>
<td>kua</td>
<td>chai</td>
<td>chao</td>
<td>chai</td>
</tr>
<tr>
<td>IMP 2 SG</td>
<td></td>
<td></td>
<td>chaì</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMP 2 PL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JUSSIVE 3</td>
<td>katab’a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PASSIVE</td>
<td>kiko(i)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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)íjn) – 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.

<table>
<thead>
<tr>
<th></th>
<th>3 sg. &amp; pl.</th>
<th>1 sg. -a</th>
<th>2 sg. -i</th>
<th>2 pl. -o</th>
<th>1 pl. excl. -i</th>
<th>1 pl. incl. -a</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTR</td>
<td>síská</td>
<td>síská</td>
<td>sískái</td>
<td>sískáón</td>
<td>sískájn</td>
<td>síská</td>
</tr>
<tr>
<td>COMPL</td>
<td>kísíská</td>
<td>kísíská</td>
<td>kínískái</td>
<td>kíníská(o)</td>
<td>nísíká(j)ín</td>
<td>nísíká</td>
</tr>
<tr>
<td>INCOMPL</td>
<td>kúesíská</td>
<td>kuísíská</td>
<td>kuínískái</td>
<td>kuáníská(o)n</td>
<td>kuáníská(j)ín</td>
<td>kuíníská</td>
</tr>
</tbody>
</table>

Table 12.1. Class IV síská ‘(s)he plays’, Santa Maria Jiotes.

Rule Subblock IIV
(11) RSS: 〈SKÁ〉 & σ {A INCOMPL} ⇔ HU 〈síská〉, SMJ 〈síská〉
(12) RSS: 〈SKÁ〉 & σ {TA INCOMPL, AGR 3} ⇔ HU 〈síská〉, SMJ 〈síská〉
(13) RSS: 〈SKÁ〉 & σ {TA INCOMPL, AGR 1SG} ⇔ HU 〈síská〉, SMJ 〈síská〉

NB: only one form in SMJ: 〈síská〉, neutralization of the INCOMPL downstep.

(14) Rule Subblock IIIV
i. RE: XIIV σ {TA COMPL} ⇔ HU, SMJ kí_REPO
ii. RE: XIIV σ {TA INCOMPL} ⇔ HUXI*, SMJ kue-, kuí-, kuá_REPO*

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:

<table>
<thead>
<tr>
<th></th>
<th>3 sg. &amp; pl.</th>
<th>1 sg. -a</th>
<th>2 sg. -i</th>
<th>2 pl. -o</th>
<th>1 pl. excl. -i</th>
<th>1 pl. incl. -a</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTR</td>
<td>sítšjó</td>
<td>sítšjúà</td>
<td>sítšjoí</td>
<td>sítšjó</td>
<td>sítšjoí(j)ín</td>
<td>sítšjoá</td>
</tr>
<tr>
<td>COMPL</td>
<td>kísítsjó</td>
<td>kísítsjoá</td>
<td>kínútsjoí</td>
<td>kínútsjó</td>
<td>kínútsjoí(j)ín</td>
<td>kínútsjoá</td>
</tr>
</tbody>
</table>
Table 12.2. **Class IV** *sitsjó* ‘(s)he crisps’, Santa Maria Jiotes.

### 6.2. Tone simplicity in the SMJ subdialect

<table>
<thead>
<tr>
<th></th>
<th>3 sg. &amp; pl.</th>
<th>1 sg. -a</th>
<th>2 sg. -i</th>
<th>2 pl. -o</th>
<th>1 pl. excl. -i</th>
<th>1 pl. incl. -a</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NTR</strong></td>
<td>béntjé</td>
<td>bêntjé</td>
<td>bêntjái</td>
<td>bêntjó</td>
<td>bêntjáit(j)in</td>
<td>bêntjái</td>
</tr>
<tr>
<td><strong>COMPL</strong></td>
<td>tsâkêntjé</td>
<td>tsâkêntjé</td>
<td>tsâkêntjái</td>
<td>tsâkêntj(a)ó</td>
<td>tsâkêntjáit(j)in</td>
<td>tsâkêntjé</td>
</tr>
<tr>
<td><strong>INCOMPL</strong></td>
<td>kuakêntjé</td>
<td>kuakêntjé</td>
<td>kuakêntjái</td>
<td>kuakêntj(a)ó</td>
<td>kuakêntjáit(j)in</td>
<td>kuakêntjé</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>3 sg. &amp; pl.</th>
<th>1 sg. -a</th>
<th>2 sg. -i</th>
<th>2 pl. -o</th>
<th>1 pl. excl. -i</th>
<th>1 pl. incl. -a</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NTR</strong></td>
<td>bêchúá</td>
<td>bêchúá</td>
<td>bêchúái</td>
<td>bêchúó</td>
<td>bêchúáit(j)in</td>
<td>bêchúá</td>
</tr>
<tr>
<td><strong>COMPL</strong></td>
<td>k'abéchúá</td>
<td>k'abéchúá</td>
<td>k'abéchúái</td>
<td>k'abéchú(a)ó</td>
<td>k'abéchúáit(j)in</td>
<td>k'abéchúá</td>
</tr>
<tr>
<td><strong>INCOMPL</strong></td>
<td>kuakéchúá</td>
<td>kuakéchúá</td>
<td>kuakéchúái</td>
<td>kuakéchú(a)ó</td>
<td>kuakéchúáit(j)in</td>
<td>kuakéchúá</td>
</tr>
</tbody>
</table>

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 Huautec Mazatec (tables 5.5, 5.6 below):

<table>
<thead>
<tr>
<th>Pike’s Tone</th>
<th>IC</th>
<th>Root</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>1SG</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td>others</td>
<td>H</td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>IC</th>
<th>Root</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>LEX</td>
</tr>
<tr>
<td>1SG</td>
<td>OCP</td>
</tr>
<tr>
<td>others</td>
<td>LEX</td>
</tr>
</tbody>
</table>

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 > u.*

NB: /u/ 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 /u/).

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’

<table>
<thead>
<tr>
<th>AGRS</th>
<th>3 SG</th>
<th>1 SG</th>
<th>2 SG</th>
<th>2 PL</th>
<th>1 PL</th>
<th>3 PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTR</td>
<td>ba'ín</td>
<td>tába'niön</td>
<td>téba'niì(n)</td>
<td>t'nìon</td>
<td>tásũba'niìn</td>
<td>inyoba'ín</td>
</tr>
<tr>
<td>COMPL</td>
<td>ya'ún</td>
<td>ya'niön</td>
<td>kika'ún</td>
<td>kika'niójón</td>
<td>kika'ñújín</td>
<td>ya'ín</td>
</tr>
<tr>
<td>INCOMPL</td>
<td>ku'ín</td>
<td>ku'niön</td>
<td>ta'ín</td>
<td>ku'ñòjón</td>
<td>ku'ñújín</td>
<td>ku'ín</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>AGRS</th>
<th>3 SG</th>
<th>1 SG</th>
<th>2 SG</th>
<th>2 PL</th>
<th>1 PL</th>
<th>3 PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTR</td>
<td>be'iún</td>
<td>tébe'niön=an</td>
<td>tibe'niùn-i</td>
<td>té'niön=ùn</td>
<td>tétsũbe'nì=ñín</td>
<td>inyabe'ín</td>
</tr>
<tr>
<td>COMPL</td>
<td>ye'iún</td>
<td>ye'niön=an</td>
<td>kike'niùn-i</td>
<td>kike'niön=ùn</td>
<td>kike'nú=ñín</td>
<td>ye'in</td>
</tr>
<tr>
<td>INCOMPL</td>
<td>ku'e'iún</td>
<td>ku'e'niön=an</td>
<td>te'niùn-i</td>
<td>ku'e'nìon=ùn</td>
<td>ku'e'nú=ñín</td>
<td>ku'e'ín</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>AGRS</th>
<th>3 SG</th>
<th>1 SG</th>
<th>2 SG</th>
<th>2 PL</th>
<th>1 PL</th>
<th>3 PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTR</td>
<td>ba'</td>
<td>tába'</td>
<td>téba'</td>
<td>t'à</td>
<td>tásũba'</td>
<td>inyoba'</td>
</tr>
<tr>
<td>COMPL</td>
<td>ya'</td>
<td>ya'</td>
<td>kika'</td>
<td>kika'</td>
<td>kika'</td>
<td>ya'</td>
</tr>
</tbody>
</table>
| INCOMPL | ku'á' | ku'á' | ta' | ku'á' | ku'á' | ku'á'

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

<table>
<thead>
<tr>
<th>AGRS</th>
<th>3 SG</th>
<th>1 SG</th>
<th>2 SG</th>
<th>2 PL</th>
<th>1 PL</th>
<th>3 PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTR</td>
<td>tinósínxó</td>
<td>tassínxó</td>
<td>tenióxé, tantènioxé</td>
<td>tatsenióxó</td>
<td>tatsenióxéen</td>
<td>ndiosínxó</td>
</tr>
<tr>
<td>COMPL</td>
<td>kisínxó</td>
<td>kisínxó</td>
<td>kinióxé</td>
<td>kiníoxó</td>
<td>kiníoxéen</td>
<td>kínxó</td>
</tr>
<tr>
<td>INCOMPL</td>
<td>bisínxó</td>
<td>bisínxó</td>
<td>tiki'ínxé</td>
<td>tettió(n)xó</td>
<td>tettió(n)xéen</td>
<td>bisínxó</td>
</tr>
</tbody>
</table>

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

**Rule Block I**

(14) RSS: \(\{XÁ\} \& \{σ\} \Leftrightarrow \text{< nioxó} \)

(15) RSS: \(\{XÁ\} \& \{\text{AGR 3}\} \Leftrightarrow \text{< sínxó} \)

(16) RSS: \(\{XÁ\} \& \{\text{AGR 1SG}\} \Leftrightarrow \text{< sínxó} \)

(17) Rule Block II

i. RE: \(X\text{vb }\sigma\{\text{AGR }\{\text{PERS 3}\}\} \Rightarrow X^{\text{ft-}\sigma}\)

ii. RE: \(X\text{vb }\sigma\{\text{AGR }\{\text{AGR 1, NUM sg}\}\} \Rightarrow X^{\text{ft-}\sigma\otimes}\)
iii. \( \text{RE: } X_{\text{vb}} \sigma \{ \text{AGR} \{ \text{PERS 2, NUM } \text{sg} \} } \Rightarrow X^{\text{Ti-} \text{ℜ} \text{T} \text{v}} \text{σ} \{ \text{AGR} \{ \text{PERS 2, NUM } \text{sg} \} } \}

iv. \( \text{RE: } X_{\text{vb}} \sigma \{ \text{AGR} \{ \text{PERS 1, NUM } \text{pl}, \text{INCL} \} } \Rightarrow X^{\text{Ti-} \text{ℜ} \text{T} \text{v}} \text{σ} \{ \text{AGR} \{ \text{PERS 1, NUM } \text{pl}, \text{INCL} \} } \}

v. \( \text{RE: } X_{\text{vb}} \sigma \{ \text{AGR} \{ \text{PERS 2, NUM } \text{pl} \} } \Rightarrow X^{\text{Ti-} \text{ℜ} \text{T} \text{v}} \text{σ} \{ \text{AGR} \{ \text{PERS 2, NUM } \text{pl} \} } \}

\begin{center}
\begin{tabular}{|c|c|c|c|c|c|}
\hline
\text{NTR} & \text{tinósín-} & \text{tasín-} & \text{tenio-, tanténio-} & \text{tatsenió-} & \text{tatsenió-} & \text{ndiosín-} \\
\hline
\text{COMPL} & \text{kisín-} & \text{kisín-} & \text{kinio-} & \text{kinio-} & \text{kinio-} & \text{kinio-} \\
\hline
\text{INCOMPL} & \text{bisín-} & \text{bisín-} & \text{tiki'-} & \text{tetió-} & \text{tetión-} & \text{bisín-} \\
\hline
\end{tabular}
\end{center}

Table 15.2. Postlexical forms, PV strings, Class IVE: \textit{tinósínxó} ‘(s)he works’, Huehuetlán.

//ba’atjà/ ‘(s)he plants’ (HU b’entjè)

\textbf{Table 16. IC IA: ba’atjà ‘(s)he plants’: sets of forms, San Francisco Huehuetlán}

16.1. Tokens

\begin{center}
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
\text{3 SG} & \text{1 SG} & \text{2 SG} & \text{2 PL} & \text{1 PL} & \text{3 PL} \\
\hline
\text{NEUTRE} & \text{kinoba’atjà} & \text{ba’atjà} & \text{teba’itja} & \text{tatsub’a’itja} & \text{tatsub’e’tja} & \text{yoba’atjà} \\
\hline
\text{COMPLETIVE} & \text{yàtig} & \text{yàtig} & \text{kitìkòj} & \text{kitìkòj} & \text{kitìkòj} & \text{yàtig} \\
\hline
\text{INCOMPLETIVE} & \text{kitìkòj} & \text{kitìkòj} & \text{tìkàtìg} & \text{tìkàtìg} & \text{tìkàtìg} & \text{kitìkòj} \\
\hline
\end{tabular}
\end{center}

16.1. PV string output

\begin{center}
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
\text{3 SG} & \text{1 SG} & \text{2 SG} & \text{2 PL} & \text{1 PL} & \text{3 PL} \\
\hline
\text{NEUTRE} & \text{kino-} & - & \text{te-} & \text{tatsu-} & \text{tatsu-} & \text{yo-} \\
\hline
\text{COMPLETIVE} & \text{yà-} & \text{yà-} & \text{kik-} & \text{kik-} & \text{kik-} & \text{yà-} \\
\hline
\text{INCOMPLETIVE} & \text{kitìkòj} & \text{kitìkòj} & \text{tìkìkòj} & \text{tìkìkòj} & \text{tìkìkòj} & \text{kitìkòj} \\
\hline
\end{tabular}
\end{center}

16.1. PV string inputs

\begin{center}
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
\text{3 SG} & \text{1 SG} & \text{2 SG} & \text{2 PL} & \text{1 PL} & \text{3 PL} \\
\hline
\text{NEUTRE} & \text{kina-} & - & \text{ti-} & \text{tetsu-} & \text{tetsu-} & \text{ya-} \\
\hline
\text{COMPLETIVE} & \text{yè-} & \text{yè-} & \text{kik-} & \text{kik-} & \text{kik-} & \text{yè-} \\
\hline
\text{INCOMPLETIVE} & \text{katìkòj} & \text{katìkòj} & \text{tìkìkatìkòj} & \text{tìkìkatìkòj} & \text{tìkìkatìkòj} & \text{kati-} \\
\hline
\end{tabular}
\end{center}

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

\textbf{Stems}

IC Shift \textbf{YES}

Anticonflative (aC) \textbf{NO}

Conflative split (CS) \textbf{YES}

Conflation breaking (CB) \textbf{YES}

PV string complexification (SCx) \textbf{YES}

Morphophonological conflation transfer (MPCT) \textbf{NO}

Incomplete overmarking (IOM) \textbf{YES}

\textbf{Endings}

AGRS suffixal endings (SSE) \textbf{YES}

AGRS enclitic ending (SEE) \textbf{YES}

AGRS indexical (desinential) endings (SIE) \textbf{YES}

\textbf{Tone}

Four tones level system (4TL) \textbf{NO}

AGRS1 & incompletive downstep (SID) \textbf{NO}

Lexical (i.e. Root) tone robustness (LTR) \textbf{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 Neogrammarnarian 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 Otomangan 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. 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:**
Casad, Eugene 1974. *Dialect intelligibility testing*, University of Oklahoma, the Summer Institute of Linguistics.


Regino, Juan Gregorio 1993. *Alfabeto mazateco*, Oaxaca, Ciesas, IOC, CDCNC.