AMERINDIA 41: 199-240, 2019

From 'complexity' to 'simplexity': A diasystemic approach to Mazatec inflectional classes

Jean Léo LÉONARD
Paris-Sorbonne, EA 4509 (STIH)

Julien FULCRAND
Université Charles-de-Gaulle Lille 3, STL (UMR 8163)

> 'Strikingly, we find that little children seem to have no remarkable difficulty in acquiring languages like Georgian, or Mohawk, or Icelandic along more or less the same time course as children learning English or Mandarin. Of course, it might be that little children are just remarkable geniuses at solving problems that seem impenetrable to scientists. But it seems more likely that morphology, despite the fact that *a priori* it seems like nothing but unmotivated and gratuitous complication, is actually deeply embedded in the nature of language' Stephen R. Anderson (2015).

Abstract: Mazatec provides a good example for the internal variation in inflectional class systems within a large dialectal continuum. This chapter provides first-hand data on a few Mazatec dialects over the Highlands and Midlands dialects, highlighting a number of important issues beyond the specific properties already known about this language in terms of inflectional complexity. The chapter is a first attempt to provide a comprehensive diasystemic description and modeling of this variation. We propose that disentangling this complexity by way of the concept of 'symplexity' is as important as describing intricate patterns within an inflectional system.

Keywords: Mazatec, Oto-Manguean, inflectional classes, inflectional complexity

1. Introduction

This paper studies the emergence of inflectional classes in Mazatec, an Eastern Oto-Manguean language complex of Southwestern Mexico, spoken by roughly 200,000 speakers in the Papaloapam Basin (see map 1). Mazatec provides a good example of why inflectional classes do not work according to an ideal model of morpho-syntax in which each distinct form would correspond to a distinct meaning. Such a model of formal matching between *expression* and *content*, would proceed deductively, instead of a discourse-induced, strongly inductive system. In Mazatec, an inflectional class system has emerged out of simple chains of preverbation consisting of an element which is a semantic correlate of general location verbs (*put, hit, deposit*,

etc.), motion verbs (*come, go, pass*, etc.), causative verbs (*do, make*) and ditransitive verbs (*talk*).

As often, when scrutinizing typological diversity, different causes unexpectedly produce the same effect. In the case of Mazatec, a somewhat broader set of directionals and positionals of the type that Jakobson (1957) called 'shifters' accounts for most of the inflectional class specification patterns. Moreover, the Mazatec inflectional class system shows an alternative way to building and implementing inflectional class. Thus, it makes a compromise between the two extremities of the lexicon/discourse polarity: on the one hand, preverbation turns out to be a powerful tool for inflectional class framing in Mazatec; on the other hand, inflectionally classified lexemes still keep much of their formerly pragmatically induced properties from discourse strategies. They implement tense, aspect, mood and voice (henceforth TAMV) and person subject agreement marking in such a handy way for native speakers that the apparent complexity of the system is actually more a problem for the analyst than for the user –as Anderson's quotation above suggests.

Although we will argue that the core system of Mazatec inflectional classes is rather simple (see Table 5 below), to reach this overall frame, one needs to scrutinize dialect variation, which provides a wide array of mechanisms and processes to building inflectional classes. In other words, the editors' claim that 'the sheer number of classes, their unpredictability, and the layering of cross-classifying systems of affixation, tone and stem alternations present both a descriptive and theoretical challenge', can be countered by a diasystemic approach of structural complexity of inflectional classes. By 'diasystem' (adjective 'diasystemic') (Weinreich 1954), we refer to the resulting metasystem that one obtains when describing the phonological and grammatical patterns of a dialect continuum. In short, a diasystemic approach provides an overview of the inner diversity of a language or a dialectal network, through a structural matrix which is robust enough to cope with geolinguistic and/or sociolinguistic variation. For our purposes, it is a model that can account for inflectional classes in any dialect within the domain in question.

Comparing sets of inflectional class subsystems in dialects helps us to understand how complexity thrives on simple patterns –the 'trick of the trade' of this realm of apparent complexity. Thus, the apparent complexity of Mazatec inflectional class diversification can be explained through a

parsimonious set of processes. The propagation of basic patterns, in particular of a pattern we call 'subconflation' (see section 5), seems to be one of the key mechanisms for further complexification. The San José Independencia variety of Mazatec provides an outstanding example for this (see section 9).

Some processes such as form-meaning reanalysis and 'iterated learning' in the sense of Hruschka *et al.* (2009) may provide another path to explanation. We deal with another family of processes in section 7-10 where we study datasets from several town dialects (Huautla, Jalapa, Mazatlán), sub-dialects (Santa Maria Jiotes, San Lorenzo) and transitional subdialects (San José Independencia, Soyaltitla). Except for the Huautla data, quoted from Pike's (1948: 95-163) seminal study, all data in this chapter are primary data and come from the ALMaz project (*Atlas Lingüístico Mazateco*, IUF, see Léonard *et al*. 2012), which is a recent project on Mazatec geolinguistics with a special focus on collecting first-hand linguistic data on both inflectional classes and tone root classes.

1.1. Complexity *versus* simplexity

Before proceeding to our modeling of Mazatec inflectional classes, we need to briefly define what we mean by the terms 'complexity' and 'simplexity'. Complexity can be understood in two separate ways: 'constitutional complexity' or 'interactional complexity'. Constitutional complexity or 'bit complexity' (*cf.* Mufwene 2013), refers to inventories of functional units or structural features, such as phonemes, morphemes and lexical stems (as in Miestamo *et al.* 2008). Interactional complexity or 'communal' (or 'socio-interactional') complexity involves intricate modules of units and features, or networks of interactive individuals. How these two dimensions of complexity dialogue with each other is still challenging. The complexity of the Mazatec inflectional class diasystem results from communal complexity –but we will not deal with this issue here,[1] as it resorts to sociolinguistics rather than to typological linguistics.

'Simplexity', in turn, means something very different from 'simplicity'. It has indeed much to do with complexity, as an underlying fractal

[1] See Gudschinsky, 1958b for a glimpse at Mazatec ethnohistory, from which one can infer *deterministic complexity* (*cf.* O'Sullivan, 2014) from isoglosses. See Gudschinsky, 1958a for a broader comparative view, in terms of *structural complexity*. For a synthesis between these various types of complexity, see Léonard & al. 2014.

component of surfacing complex sets in e.g. phonology or grammar. Simplexity flows from parsimony and transparency, though such properties may entail complex inner coordinated patterns. These patterns may also involve complex sets of interactions between functional units: below intricate surface patterns, one may find a few and iterative simple patterns embedded at deeper level of the grammar (i.e. *fractal* patterns). What makes the surface patterns look intricate is that they mingle sets of complex interactions (e.g. competing paradigms and models). Simplexity happens to be a central concept for the study of body motion mechanisms (Berthoz, 2009). Here, simplexity will mean *a system of simple constraints making complex outputs predictable*. More explicitly, it can be seen as a 'trick of the trade' that any complex system, such as grammar, may use to achieve its goal for understandability and learnability. In the case of understandability, the goal stands in the realization of linguistic units in speech and discourse, as smoothly and accurately as possible. In the case of learnability, the system should provide simple basic schemes or mechanisms for complex figures and representations. 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. combinatory patterns).

We claim that in Mazatec, the discourse embedding continuum resorts to simplexity, beyond surfacing grammatical complexity. It results in a wide array of inflectional classes, though the latter are transparently coordinated at a diasystemic level. These inflectional classes result mainly from trajectory and motion light verbs, but also partly from causative and active/stative light verbs. As far as complexity is concerned, the diversity of the content of the cells is indeed intricate. But the principles implied in this magic-cube game are actually fairly simple. One of the 'tricks of the trade' qualifies as 'subconflation-driven' (a phenomenon involving split subject agreement marking), see in section 4 –either as forms of neutralizing subconflating paradigms (i.e. *neutralizing* or *avoiding split subject agreement marking* and another process, called *subconflation breaking*) or making them stronger (e.g. through processes such as *string complexification, split subject agreement, stem templatic allomorphy, incompletive overmarking*, etc.: see especially Tables 4 and 5 below).

1.2. Outline of the paper

We present the main dialect areas and the varieties under scrutiny in section 2. In section 3, we introduce the reader to phonological aspects of Mazatec, enhancing the role of modeling in the presentation of linguistic categories. We also justify the use of current Mazatec orthography and spelling conventions, rather than of a phonetic alphabet. Section 4 provides basic templates and information on subject agreement clitics, affixes and person marking. Section 5 presents our model for diasystemic analysis from which all subsequent data will be sorted according to an inflectional class grid. Section 6 enumerates and explains how diasystemic processes can be evaluated and qualified according to three disjunctive blocks: rules of stem selection, rules of exponence, and morphophonological rules, according to basic principles in *Paradigm Function Morphology* (although we will not use formalization currently used in this framework, in order to focus on the descriptive dimension of our modeling, rather than the formal or declarative dimension, as in Stump 2001). The prerequisites of our diasystemic model are implemented in a first comparison between two Highland dialects (Huautla and San Lorenzo). From then on, we describe three types of varieties: town dialects,[2] satellite sub-dialects and transitional dialects. In section 7, a satellite subdialect close to the Huautla variety, Santa Maria Asunción (also called Santa Maria Jiotes), will be studied. Section 8 deals with the second most important town dialect, San Felipe Jalapa de Díaz, which we systematically compare with the Central Highland dialect of Huautla. After this first empirical cluster of dialects, in which we do not lose sight of Huautla data being the most well-known Mazatec variety in the literature,[3] we give an account of the main structural trends in terms of inflectional class taxonomic options in three dialects: two transitional varieties (San José Independencia, in section 9 and Soyaltitla in section 10) and one peripheral dialect from the North-Western Highlands diasystemic segment: San Lorenzo (Mazateco 'poblano', or *an xo'boo*), in section 11. In section 12, we present prospects on five issues underlying our empirical survey: the *empirical*

2 As to the definition of *town dialects* in Mesoamerica, see Suárez (1983: 19-20).

3 Of course, the Chiquihuitlán dialect is actually the most thoroughly described variety as far as IC are concerned, thanks to Jamieson's comprehensive work (Jamieson 1982, 1988, 1996; Jamison & Jamieson 1978). Nevertheless, the Chiquihuitlán variety is one of the most indiosyncratic of the whole Mazatec dialect network (see Kirk 1970), and except the Jamiesons, no one has been publishing more about this specific variety. Instead, Jalapa and Huautla de Jiménez have been studied or documented by a wide array of scholars, and Huautla inflectional patterns have been extremely well described by Kenneth Pike in the Chapter 8 of his essay on tone languages (Kirk *op. cit.*). The Huautla dialect being politically and economically central, we'll often refer to this dialect and to Pike's seminal description.

gap, the *heuristic metasynthesis*, *unavoidable simplexity*, *interaction in 3-D*, and the *model uphill*.

1.3. Dialect clusters

The Mazatec diasystem (Popolocan, Eastern Oto-Manguean) can be divided into two main zones: the Highlands and the Lowlands. Other subzones can be further distinguished, such as the Midlands (Jalapa de Díaz, Santo Domingo, San Pedro Ixcatlán) within the Lowlands, the Cuicatlán Canyon (Chiquihuitlán) and the Puebla area (see San Lorenzo data below). In short, main dialect subdivisions read as follows, slightly modified from Léonard *et al.* (2014):

(1) The Mazatec diasystem: dialects and subdialects

Highlands complex:
-*Central* Highlands (Huautla de Jiménez, Santa Maria Jiotes, San Miguel Huehuetlán)[4]
-*Northwestern* Highlands:
. Central Northwestern Highlands (San Pedro Ocopetatillo, San Jeronimo Tecoatl, San Lucas Zoquiapam, Santa Cruz Acatepec, San Antonio Eloxochitlán)
. Peripheral Northwestern Highlands (San Lorenzo Cuaunecuiltitla, Santa Ana Ateixtlahuaca, San Francisco Huehuetlán)

Lowlands complex:
-*Eastern* Lowlands (San Miguel Soyaltepec)
-*Central* Lowlands (San Pedro Ixcatlán)
-*Piedmont* (Ayautla, San Felipe Jalapa de Díaz, Santo Domingo)

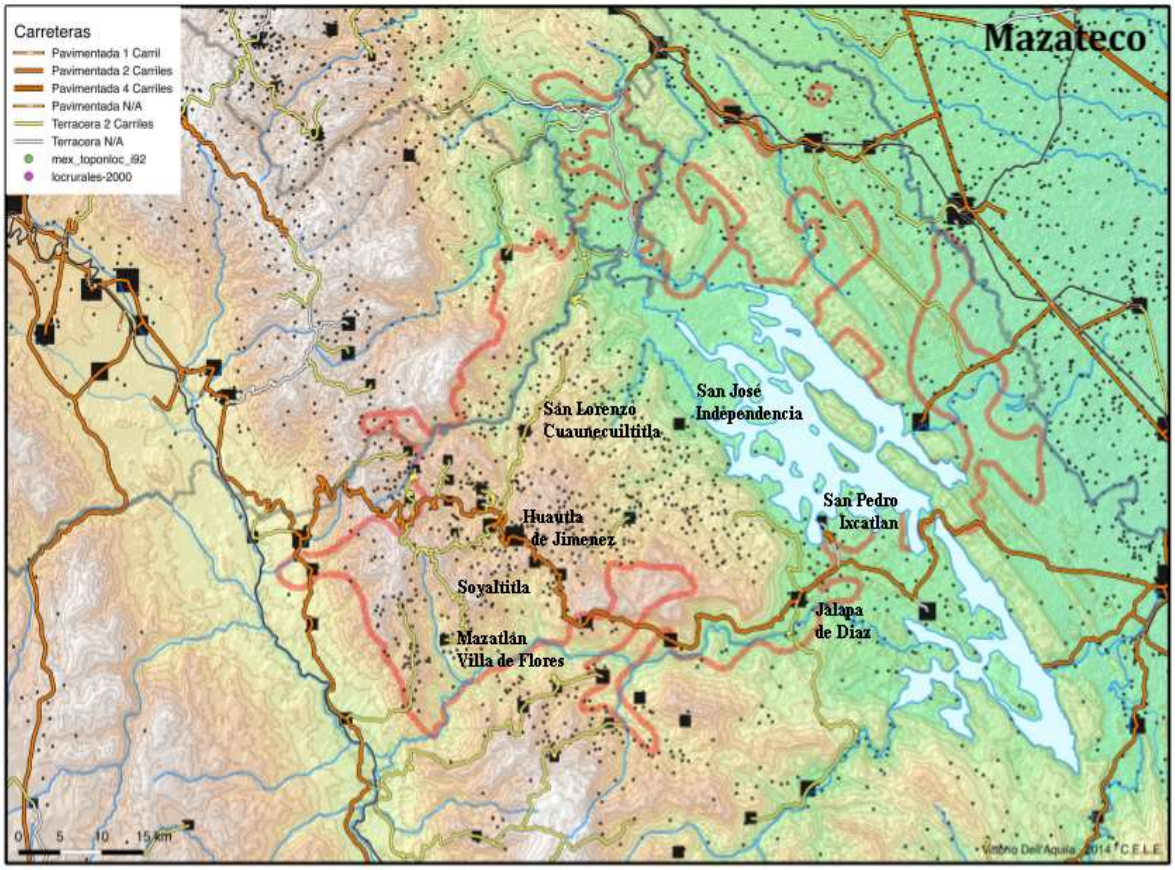
Periphery:
-*Western Cordillera*: Mazatlán Villa de Flores
-*Cuicatlán Canyon*: Chiquihuitlán.

A classification such as (1) provides only a heuristic framework to observe dialectal variation.[5] We have had the opportunity to survey two transitional dialects: on the one hand, the San José Independencia dialect –a variety from the Highlands complex which has recently migrated to the Central Lowlands

4 For more details, see <http://www.inali.gob.mx/clin-inali/html/v_mazateco.html> for a complete list of population centers.

5 Dialect differences may be conspicuous between the main sets, as Central Highlands and Lowlands complex or Periphery, but even within sub areas. This is especially true for peripheral dialects such as Peripheral Northwestern Highlands (the so-called *Poblano* dialect, or *an xo'boo*), as regards Central Northwestern Highlands, and for the two peripheral dialects of Mazatlán and Chiquihuitlán (Cuicatlán Canyon) –the former converges much more with all the other dialects though being phonologically highly idiosyncratic, whereas the latter is radically different from the rest of the network. The Peripheral Northwestern Highlands counts among those which have undergone the most phonological changes through vowel shifts, though it converges strongly with the rest of its sub-area (Northwestern Highlands), and even with the rest of the Central Highlands complex.

area– and on the other hand, the Soyaltitla variety –a South-Western sub-dialect strongly attracted by the Central Highlands type. Map 1 initially took over the localities surveyed by Paul L. Kirk in 1966 (abbreviated on map 1). Black squares hint at the main localities where linguistic data have been gathered for the *ALMaz* inquiry. Circled spots point at the varieties analyzed here. From left to right, clustering groups of two circles: 1) San Lorenzo, 2) Soyaltitla, 3) Huautla, 4) Santa Maria Jíotes, 5) San José Independencia, 6) Jalapa de Díaz.



Map 1. The six Mazatec dialects of the study and localization of Mazatec in Mexico. Map by Vittorio dell'Aquila (CELE)

2. Phonology: inventories and models

As is the case for many Oto-Manguean 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), or on the contrary, its basic consonant inventory may well rank among the average, with only 16 inherited consonants (*cf.*

Table 1). One could even say that, at the segmental level, only consonant cluster constraints in this language are complex (as in Table 2). The basic syllable structure is of the CV type, with $CCCV^N$ patterns, according to Pike & Pike (1947). In this section, some elements will also be given about the ALFALEIM spelling conventions.[6]

	Labial	Coronal			Dorsal	Glottal
		Dental	Palatal	Retroflex		
Stops		t <*t*>			k <*k*>	ʔ <ʾ>
Affricates		ts <*ts*>	tʃ <*ch*>	tʂ <*z*>		
Fricatives		s <*s*>	ʃ <*x*>			/h/ <*j*>
Resonant Nasals	m <*m*>	n <*n*>	ɲ <*ñ*>			
Lateral		l <*l*>				
Approx/Glides	w <*b*>[7]		j <*y*>			

Table 1. The ALFALEIM spelling model (1980-), mostly inherited from P&P 1947

The vowel system is pentavocalic: /i, e, a, u, o/, with restrictions on distribution patterns for the high and mid-high back vowels according to the dialect. All vowels are short, but may be prosodically inflected by contour tone. There are constraints of vowel sandhi in inflection, as we'll soon observe in the data, in the juncture domain between the root vowel and the desinential nucleus. Diphthongs such as *ia, ua, oa, ao, ui, ue, oe* may all be considered as secondary outputs of onset-nuclei interactions, with further vocalization of a contoïd

6 ALFALEIM = 'Alfabeto para la Lengua Indígena Mazateca'. See also Regino, 1993 for basic principles of Mazatec orthography (especially Lowlands Mazatec, whereas the ALFALEIM, contrived initially by Maximino *Cerqueda* García is mostly referred to in the Mazatec Highlands). Both spelling conventions (Regino's codification and *Cerqueda* García ALFALEIM) converge, except for such *ad hoc* graphemic solutions to complex phonemes 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 cultural revival, revitalization and 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 at this time already highly relied on diasystemic description of the Lowlands complex.

7 A strong trend to betacism (*w* > *v, b*) has been spreading swiftly through the diasystem in the late decades. In Kirk's comparative data from 1966 the labial approximant *w* is still abundantly documented for many dialects, especially in the Lowlands and in the Periphery (we also noted [w] ← /w/ instead of [b] ← /w/ in Pueblan Mazatec). It seems the phenomenon has irradiated from centers as Huautla and Jalapa, and it is gaining a stronghold nowadays nearly everywhere in younger speakers' sociolects. Pike & Pike (1947) and Pike (1948) still record an intermediate allophone of the fricative type <v> in their transcription of the Huautla data, i.e. a bilabial spirant [β], whereas most speakers now have plain [b] for /w/. In order to not puzzle the reader with erratic allophones, we have unified data in this paper using <*b*> as the only grapheme for underlying /w/.

element as first component (see Golston & Kehrein 1998, and morphophonological processes explained further in this paper).

Things might get a bit trickier when we consider the syllabic structure. As mentioned before, the complexity of the Mazatec diasystem and its interpretational variation leads to flexible models. Table 2 presents the two main models: Pike & Pike (1947) and Golston & Kehrein (1998):

	P&P's Syllabic Constituency Model	G&K's Laryngeal Features Model	
Stops & Affricates	*t, ts, tʃ, tʂ, k*	*t, ts, tʃ, tʂ, k*	No difference
[spread glottis]	*ht, hts, htʃ, htʂ, hk*	*t̤, ts̤, tʃ̤, tʂ̤, k̤*	Onset Realignment
[constricted glottis]	*tʔ, tsʔ, tʃʔ, tʂʔ, kʔ*	*_V̰*	Nuclear Realignment

Table 2. Two models for consonants in Huautla Mazatec, according to P&P 1947 and G&K 1998 (a fragment)

Mazatec is an archetypical CV language: no codae are allowed, and all complexity concentrates on the onset. In P&P's model, onsets may be complex, with up to three consonants, and nuclei too (diphthongs), with no coda. In Golston & Kehrein's model, onsets may be of the C or the ^{n}C, C̤ or C̰ type, whereas nuclei can be modal (V), spread (V̤) or constricted (V̰). P&P took several things for granted, such as overt consonant clusters, but there are others. There are, for instance, the onsets partially enriched by breathiness or creakiness. There are also the partly breathy or creaky nuclei, i.e terms of a voice quality correlation.

Although these kinds of correlates were known at least since Trubetzkoy's treatise on phonology (Trubetzkoy 1939), P&P missed this opportunity clinging instead to surface patterns. However, from the standpoint of modern phonological theory though, most of Mazatec (and other Oto-Manguean languages) actually resort to this typological parameter of *voice quality*, in close interaction with phonetic implementation (see Silverman 1997; Golston & Kehrein 2004 and more specifically on Jalapa Mazatec, Silverman *et al.* 1987, Kirk *et al*. 1993).

As a consequence, to us, a strong claim in favor of the accuracy of G&K's model is that it better accounts for the complexity/simplexity relation of the glottal and nasal features with pivotal segments, which are basically voiceless occlusives, (except *p*, borrowed from Spanish) and affricates, voiceless fricatives and a set of approximants (*w, y; l* in the Highlands, *r* in the Lowlands).

The ALFALEIM spelling provides a simple and very easy framework to handle any Mazatec dialect. Complex clusters end up in a simple form-spelling matching induced by Spanish orthography, readable to anyone: < ’ > stands as much for a glottal stop as for creakiness, < j > or *jota* stands as much for ‘aspiration’ as for breathiness. < b > accounts for betacism (/w/ > [b]) – a rampant phonological change all over the Mazatec dialect network. < *z* > is a simplex solution to note the retroflex affricate, instead of < *chr* > or < *xr* >, still used by many school teachers in the Highlands. 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 to the populations (see Léonard 2010, for more details on this aspect of the project, and see documents made available online at <http://axe7.labex-efl.org/taxonomy/term/12>).

Mazatec has both level and contour tones. Some dialects have three level tones (Jalapa de Díaz) while others have up to four (Huautla, San Miguel Soyaltepec). The ALFALEIM convention for writing tones is the following:

High (/H/)	represented with an acute accent	(*á*)
Mid-high (/h/)	represented with a grave accent	(*à*)
Mid (/M/)	which is not represented	(*a*)
Low (/L/)	represented with an underlining	(*a̲*)

Contour tones (HM, HL, hM, hL, ML, LM, Lh, Mh, hH, LhL, LhM, *cf.* Pike 1967: 313) do not significantly increase duration in nuclei, as already noted by Pike & Pike (1947). Nevertheless, they are noted in ALFALEIM with double vowels: < *áa* > for HM, < *áa̲* > for HL, < *àa* > for hM, etc.

3. Lexical and morphological templates

Not only is Mazatec far more regular and easier to describe than what the available data and models convey, from the two or three most famous dialects, but Mazatec verb forms can be schematized as in (1), where *W, St,* and √ identify the word, the stem, and the root, respectively, as a canonical lexical & morphosyntactic primary or basic template:

TENSE-ASPECT-VOICE agreement markers
(1) < *W* (CV) < *St* CV < √ (C)CV > V >>

The markers for TAM and voice can be either prefixes or proclitics; the same happens to agreement markers. They can be suffixal as in (2-3) or pronominal enclitics (4).

Huautla Mazatec

(2) *ki=si̱i+ská-**a***
CPL=IC.FORM.1SG+play-1SG
'I played.'

(3) *ki=nì+ská-**i***
CPL=IC.FORM[NOT.1SG/3SG]+play-2SG
'You played.'

Mazatlán Mazatec

(4) *ki=ní+tsjò=**ji***
CPL=IC.FORM[NOT.1SG/3SG]+toast=2SG
'You toasted.'

As for subject person agreement markers, there are six person-number values distinguished in Mazatec. There is no number contrast for the 3rd person, at least in the stem, and there is an inclusive *vs* exclusive distinction for the 1st person plural. Table 3.1 shows the diasystemic paradigm of agreement endings:

	SG	PL	
1	*-'a(n)*	INCL	*=ñà, =jñà*
		EXCL	*-in, =ji̱n, =jni̱*
2	*-i, =ji*		*-un, =jún, =jnú, =no*
3	---		

Table 3. Mazatec diasystemic paradigm of subject agreement endings

Agreement markers do not contribute to defining inflectional classes in Mazatec. Nevertheless, as we'll often see in the argumentation, two facts need attention: on the one hand, vowel fusion between the nuclei of lexical roots and light markers, such as affixes and vocalic elements do make up in each dialect or area a taxonomic complex of its own (see 1948: 118-119 to Schram & Pike 1978); and on the other hand, the variability between enclitic and affixal concatenative constraints is often puzzling, within the same dialect, and seems partly lexically driven, partly templatic –i.e. phonological. Nevertheless, as there is no direct incidence of these phenomena on inflectional class shaping and definition, at least from a descriptive standpoint, we'll just keep an eye on them, though we'll not give them the same taxonomic status as preverbs, stems and TAM proclitics.

4. Modeling inflectional classes in the Mazatec diasystem

Focusing on Chiquihuitlán Mazatec, Jamieson (1982) proposes a system of 18 inflectional classes accounting for preverbal stem formatives, that we call

'preverbs' here. We argue that a 'simplex' (i.e. more parsimonious) model of seven inflectional classes can be instead posited to account for most of the complexity available in the rest of the Mazatec diasystem. Table 4 gives the basic grid of our proposal, based on the dialect of Huautla from Pike (1948).[8] The first column gives the preverb, the second enumerates the inflectional class, and the third gives the root tone class of the verb serving as an example for each inflectional class.

Preverb	inflectional class	root tone class	AGR	NTR	CPL	INCPL	
b'é+	I/a	1 /H/	3	***b'é**+xá*	*tsa=k'-**é**+xá*	*k'-**o̲è**+xá*	'send'
			1SG	***b'e̲e**+xáa*	*tsa=k'-**e̲e**+xáa*	*k'-**o̲e**+xáa*	
			1PL.INCL	***b'è**+xá*	*tsa=k'-**è**+xá*	*k'-**oé**+xá*	
ba+	I/b	3 /H/	3	***ba**+né*	*tsa=k-**a**+né*	*k-**o̲a**+nè*	'wash'
			1SG	***ba**+née*	*tsa=k-**a**+ne̲e*	*k-**o̲a**+ne̲e*	
			1PL.INCL	***ba**+nè*	*tsa=k-**a**+nè*	*k-**o̲a**+nè*	
bi+	I.c/1	3 /LM/	3	***bi**+ya̲a*	*tsa=k-**i**+ya̲a*	*k-**o̲i**+yá*	'die'
			1SG	***bi**+ya̲a*	*tsa=k-**i**+ya̲a*	*k-**oi**+yáa*	
			1PL.INCL	***bi**+ya̲à*	*tsa=k-**i**+yaà*	*k-**o̲i**+yá*	
ba+/bi+	II.c/2	3 /M/	3	***ba**+te*	*tsa=k-**a**+te*	*k-**o̲a**+te*	'break'
			1SG	***ba**+te*	*tsa=k-**a**+te*	*k-**o̲a**+te*	
			1PL.INCL	***bi**+chà*	*tsa=k-**i**+chà*	*k-**o̲i**+chá*	
bá+/fa+	II.c/3	3 /M/	3	***báj**+tsa*	*tsa=k-**áj**+tsa*	*k-**o̲àj**+tsa*	'put to cook'
			1SG	***ba̲aj**+tsàa*	*tsa=k-**a̲aj**+tsàa*	*k-**o̲àj**+tsàa*	
			1PL.INCL	***faj**+tsaà*	*tsa=kj-**aj**+tsaà*	*kj-**oáj**+tsaà*	
b'a+/ch'a+	III	1 /H/	3	***b'a**+kjá*	*tsa=k'-**a**+kjá*	*k'-**o̲a**+kjá*	'wear'
			1SG	***b'a**+kjáa*	*tsa=k'-**a**+kjáa*	*k'-**o̲a**+kjáa*	
			1PL.INCL	***ch'à**+kjá*	*ki=**ch'à**+kjá*	***ch'a̲**+kjá*	
sí+/nì+	IV	2 /h/	3	***sí**+tsjò*	*ki=**si̲i**+tsjò*	***si̲i**+tsjò*	'toast'
			1SG	***si̲i**+tsjòa*	*ki=**si̲i**+tsjòa*	***si̲i**+tsjòa*	
			1PL.INCL	***nì**+tsjoà*	*ki=**nì**+tsjoà*	***si̲**+tsjoà*	
tsò+/mi+	V	3 /M/	3	***tsò**+ya*	*ki=**tsò**+ya*	*k'-**ui**+tsò+ya*	'teach'
			1SG	***tsò**+yaa*	*ki=**xi**+nyaa*	***xi̲i**+nya*	
			1PL.INCL	***'mi**+yà*	*tsa=k'-**i**+nyà*	*k'-**u̲i**+nyà*	
kjo+/chjo+	VI	3 /M/	3	***kjo**+ya*	*ki=**sko-**ya*	***skó**+ya*	'cook'
			1SG	***kjo**+yaa*	*ki=**sko-**yaa*	***sko**+ya*	
			1PL.INCL	***chjo**+yà*	*ki=**chjo-**yà*	***chjo**+yà*	
se̲e+/jnta+	VII	5 /BM/	3	***se̲e***	*ki=**se̲e***	*sé*	'sing'
			1SG	***se̲e***	*ki=**se̲e***	*ki=**se̲e** / **se̲e***	
			1PL.INCL	***jntaà***	*ki=**jntaà***	*ko̲i=**jntá***	

Table 4. Huautla root tone classes & inflectional classes. Data from Pike (1945) and ALMaz fieldwork

We take the three person/number forms (i.e. 3rd, 1SG and 1PL.INCL) in the three aspects as illustrative for each class. These forms illustrate best stem alternations involving preverbs as stem formatives. This is because in some

8 The inflection class taxonomy from Class I/a to VII is ours, but was strongly inspired by the discussions with Juan Casimiro Nava (p.c.) –a Mazatec linguist from Huautla.

classes (like, for example, in II.c/2) a given form of the preverb is used for 3rd and 1SG, while another form is used with the rest of the persons (indicated here as '-3/1SG'); and we use the form for the 1PL.INCL to illustrate this. This may refer to a phenomenon we will call 'subconflation' (Jamieson 1982 referred to "*conflated subsystems* marking person and aspect"), and sometimes to stem suppletion as in Class VII.

In Table 5, Roman numerals stand for main inflectional classes, whereas letters and Arabic numbers stand for sub-classes. Preverbs constitute the primary area of dialectal variation. As indicated in Table 3, the classes are thus divided according to the phonological properties of the preverb's onset: with a labial, a coronal or a dorsal consonant.

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

Table 5. Mazatec diasystemic inflectional class taxonomy for verb inflection.

For Huautla Mazatec and for many Highlands varieties, verbal stems are divided into three major morphophonological classes, indicated by the Roman numbers: one with labial onsets (I, II, III), one with coronal onsets (IV, V) and another with dorsal onsets (VI). Subdivisions mainly respond to the existence of some degree of semantic correlations: *general location*, such as 'put', 'get', 'carry', 'hold'; *dynamic* or *motion*, with directionals, such as 'go', 'pass', 'come', etc.; and *causatives* 'make'. But these semantic correlates should be taken as heuristic labels rather than as actual form-meaning mappings in the language. They reflect old compoundings, and most of the time the collocations are lexicalized. Notice for example that the verbs of Classes V and VI have no such discernible correlates.

Also as pointed out above, many such classes are 'subconflative' (i.e. split subject agreement marking), that is, they use one preverb for 3rd and 1SG and a

different one for all other subject agreement cells. The labial onset complex includes three preverbal classes: I, II, and III. Class I/a (*b'é*+), Class I/b (*ba*+), and Class I.c/1 (*bi*+), which all share the property of not being subconflative.

The matrix in Table 5 allows a reference point for comparison across dialects. Some dialects vary more in this aspect than others. For example the Northwestern Highlands dialects and the Eastern Lowlands dialects tend to have more subconflative or split subject agreement classes than others, and they recruit or combine preverbs differently. In all dialects, morphophonological rules targeting the stem-initial labial onsets are valid.

5. Diasystemic variables

We'll first describe stem formation processes, according to TAM marking, before explaining subject agreement marking strategies, and tone patterns.

Table 6 below compares two lexical and inflectional strategies for the verb 'toast', with separate inflectional classes in each dialect for two different lexemes: in San Lorenzo, a Class I/a in monosyllabic lexeme *batsú* ← |***bé***+tsú| resorting to *default* (or *non subconflative*) Class I/a subset versus *sítsjò* ← |***sí***+tsú|, from Class IV in Huautla, with a causative allomorphic preverb ***sí***+/***nì***+, according to the subconflative inflection pattern.

'toast''		SL	HU
		Class I/a	Class IV
NTR	3	*batsú*	*sítsjò*
	1SG	*batsé'e*	*si̲itsjòa*
	2SG	*bats'í*	*nìtsjòi*
	2PL	*batsúu*	*nìtsjòo*
		ba+	***sí***+/***nì***+
CPL	3	*yátsún*	*kisi̲itsjò*
	1SG	*yatsé'e*	*kisi̲itsjòa*
	2SG	*yátsí*	*kinìtsjòi*
	2PL	*yátsúu*	*kinìtsjòo*
		yá+	***ki***=

Table 6. Inflectional Class I/a. SL *batsú*, Class IV HU *sítsjò* 'toast' (Pike, 1948)[9]

Typically, we observe here an inflectional class shift I/a (SL) ↔ IV (HU). Preverb collocation alternates, whereas the lexical polyvalent root remains

[9] This set of San Lorenzo data was recorded during a Mazatec Literacy Workshop in San Lorenzo Cuaunecuiltitla held in August 2013, by Jean Léo Léonard and Jaime Calderón (UNAM), from a young speaker, Isabel Juárez Arciga, aged 15. The idiolect is a mixed SL/Matzazongo de Guerrero variety, closer to San Francisco Huehuetlán than SL proper, as far as completive marking for the I/a Class completive is concerned.

the same (SL) *tsú*, (HU) *tsjò*. As for completive marking, the SL variety neatly shows a completive preverb option: completive implies a preverbal *ba*+ → *yá*+ substitution, whereas HU resorts to a *ki*= proclisis concatenated to the causative preverb: 3 CPL *ki*=*s̲iitsjò* '(s)he toasted'. Indeed, one can observe that the SL dialect undergoes more agreement subject ending coalescence with the root vowel 1SG NTR SL *batsé'**e*** *vs* HU *s̲iitsjò**a***, 2SG NTR SL *bats'í vs* HU *nìtsjòi*, but what happens in this desinential domain does not interfere with inflectional class patterns. Instead, the split subject agreement marking on the suppletive preverb for -3/1SG forms does concern inflectional class assignment: 2SG NTR HU ***nì***+*tsjòi*, 2PL NTR ***nì***+*tsjòo vs* 3 NTR ***sí***+*tsjò* and 1SG ***s̲ii***+*tsjòa*.

In the next set of data in Table 7, SL inflectional class Class I/a *wayo* matches HU Class I.c/1 *biy̲aa* 'die', now as cognates. Here, we included the subject agreement 1PL column at the right of the 1SG column, in order to make SL syncretisms more obvious. The inflectional class shift of the type Class I/a (SL) ↔ Class I.c/1 (HU) accounts for a correspondence between the two dialects: SL *wayo* lexically matches |*weya*|, whereas HU *biy̲aa* resorts to |*wiya*|, i.e. inflectional class Class I/a *vs* Class I.c/1. As before with *batsú* ← |***bé***+tsú| in the previous set of data, although the postlexical realization in SL might wrongly induce an inflectional class I/b classification for the reader at first sight, SL *wayo* does resort to a I/a Class.

Incidentally, notice that a root tone Class shift occurs: SL has a mid tone root *wa**yo***, while HU has a LM contour root tone of the Class pattern *biy**a̲a***. As a result, SL does display an incompletive prosodic lowering process located in the ending, instead of being hosted in the preverbal slot, inflectional class I/a having a labial onset, CPL *k*= and INCPL *lok*= proclisis both trigger the labial onset dropping in SL (noted b-drop in the rightmost cell), so that both dialects (SL & HU) agree for this inflectional class on basic juncture processes within the preverbal domain –they do share some morphophonological constraints on labial onset inflectional class, although not always for the same TAM set (b-voc happens rather in the INCPL in HU than in the CPL). The symbol < & > in the rightmost cell points out to additional processes or constraints, such as a structural correlate for the paradigm.

'die'		SL	HU
NTR	3	*wayo*	*biya̲a*
	1SG	*wayo*	*biya̲a*
	1PL	*wayo*	*biya̲à*
	2SG	*wayi*	*biya̲i*
		Class I/a	Class I.c/1
CPL	3	*kuayo*	*tsakiya̲a*
	1SG	*kuayo*	*tsakiya̲a*
	1PL	*kuayo*	*tsakiyaà*
	2SG	*kuayi*	*tsakiya̲i*
		k= b-voc	*tsak*= b-drop
INCPL	3	*lokayò*	*ko̲iyá*
	1SG	*lokayò*	*koiyáa*
	1PL	*lokayà*	*ko̲iyá*
	2SG	*lokayì*	*ko̲iyái*
		lok= b-drop	*k*= b-voc &, Tone Lowering

Table 7. Class I/a. SL *wayo*, Class I.c/1. HU *biya̲a* 'die' (HU data from Pike, 1948)[10]

Table 8 displays a very interesting set of data, which highlights how inflectional classes may differ from a dialect to another in Mazatec. SL *mònàchòkè* 'run' ← |ma̲ne̲cha̲nki̲| can be further analyzed as |wa-ne=#cha##nki#| –a compound stem combining two polyvalent motion roots *cha* (for 'iterative trajectory motion') and *nki̲* 'go (down)', and a TAMV string with a labial initial onset |***m****ane*=|. The vocalic component in the *ma-/mi-* preverb changes to *-i-* in CPL 1PL & INCPL 2SG, 1PL. The *ma-/mi-* preverb is akin to the *ba-/bi-* type which stands as Class II.c/2 in Table 5. This amounts to a subconflative pattern of the *ma-/mi-* type allowing classification of the SL forms as belonging to this inflectional class.

Moreover, the *mi-* alternation does not follow the canonical subconflative asymmetry constraint (+3 & 1SG *versus* other subject agreement cells, i.e. the α3/1SG person & TAMV marking subsystem constraint), as CPL 2SG has *komonachokìn* instead of expected *kominachokìn*. This suffices for labelling this process as *subconflation breaking*. Even more striking, the subconflative patterns only shows up for CPL 1PL and INCPL 2SG & 1PL, whereas it does not appear in the Neutral paradigm. These asymmetries are confirmed in the complete matrix, not given here in order to make data easier to read. This confirms the inflectional class shift, though with a default variant neutralizing subconflative

10 This set of data, as the following SL material, has been recorded from our main informant, Abraham Cabrera Gabito, mentioned in the list of language consultants at the end of this paper.

asymmetry for a subset of the matrix, while introducing subconflation (i.e. α3/1SG asymmetry) in the INCPL cells –a process avoided in a Central Highlands dialect as HU.

Instead, there is not much to say about the HU data here: all forms classify within inflectional class V (i.e. preverb with a coronal onset), though as a default paradigm (i.e. without asymmetric subsystems α3/1SG). Interestingly enough, HU INCPL preverbal domain shows tone raising, typical of the Central Highlands INCPL tone level shift, according to which a root tone class from the upper prosodic level (H, h) undergoes a preverbal lower register contour (BM) in the INCPL, whereas a root tone class from the lower prosodic level (M, L), on the contrary, undergoes preverbal tone raising. To some extent, a few hints at a similar phenomenon are to be seen in the SL data, as most NTR & CPL cells have a mid-high plateau (corresponding to a lexical low plateau), which undergoes neutralization in the INCPL, except for one mora within the template (on the right margin for INCPL 1SG *lokonachokè* '(s)he'll run' *vs* in the middle of the word form for other cells: INCPL 2SG *lokinàchoki* 'you'll run', 1PL *lokinàchoke* 'we'll run').

'run'		SL	HU
NTR	3	*mònàchòkè*	*tjokaa*
	1SG	*monàchoken*	*tjokaa*
	2SG	*monachokín*	*tjokai*
	1PL	*mònachòkèn*	*tjokaà*
		Class II.c/2 (partially) default	Class V default
CPL	3	*kòmònàchòkè*	*kitjòkaa*
	1SG	*kòmònàchòkè*	*kitjòkaa*
	2SG	*komonachokìn*	*kitjòkai*
	1PL	*ko**mi**nachokè*	*kitjòkaà*
		ko= subconflation breaking	*ki=*
INCPL	3	*lokonachokè*	*tjókaa*
	1SG	*lokonàchoke*	*tjókaa*
	2SG	*lo**ki**nàchoki*	*tjókai*
	1PL	*lo**ki**nàchoke*	*tjókaà*
		lok= b-drop inflectional, Class shift & subcfl	Default stem *tjóka-*, PV Tone Raising

Table 8. Class II.c/2. SL *mònàchòkè*, Class V. HU *tjokaa* 'run' (HU data from Pike 1948)

The discrepancies between SL and HU are conspicuous here, and even more striking when one considers that both are Highlands dialects. As a matter of fact, the Northwestern Highlands dialects (including the central ones, like San Antonio Eloxochitlán) share many inflectional class patterns with the

Lowlands dialects (e.g. the CPL PV *yé-* in inflectional class 1A, but also several preverb string complexification patterns). These facts point in one direction: the strong idiosyncrasy of the Central Highlands dialect (HU) is probably a recent phenomenon due to Huautla's sociocultural and economic hegemony over the Mazatec Highlands.

However, this short survey of a handful of verbs paved the way for us to apply the diasystemic patterns involved in inflectional class diversification. We can now proceed at a slower pace, step by step, from a more Central Highlands variety structurally closer to the HU dialect, before dealing with more different dialects, such as the Lowlands varieties.

6. Santa Maria de la Asunción

The dialect of Santa Maria de la Asunción (called 'Santa Maria Jiotes' in Kirk 1966) also belongs to the Central Highlands network. It is famous among Mazatec highlanders for its 'singing tune'. Huautla speakers find it easy to understand but characterize it as 'somewhat different'. Table 9 compares two verbs of Classes I/a and IV between Santa Maria de la Asunción (SMA) and Huautla (HU), which we take as rendering good diagnostics for diatopic varieties. Class I/a is non subconflative, i.e. a default class, but it has a preverb with a labial onset that regularly produces juncture processes with TAM proclitic strings, such as *b-drop, b-voc*, etc.

These two very geographically close dialects (about 10 kms from each other) differ in various respects. SMA has inflectional class shift to Class I/a. For example, HU has the verb 'plant' in subconflative Class II.c/3 (NTR 3 *bá+ntjè* '(s)he plants', NTR 1SG *ba̲a+ntjèe* 'I plant'; NTR 1PL.INCL *fa+ntjè* 'we plant') but the SMA corresponding lexeme shows up as a verb resorting to Class I/a (NTR 3 *bé+ntjé*, NTR 1SG *bè+ntjé*, NTR 1PL.INCL *bé+ntjé*), compare with Class I/a verb 'send' (also Class I/a in HU).

Preverbal tonal contrasts in SMA are fine-grained: the preverbal mid-high tone in NTR 1SG *bè+ntjé* matches HU preverbal tone lowering contour in NTR 1SG *ba̲a+ntjèe*. Both varieties have the clitic *tsak=* for the completive and elision of the /b/ in the onset of the preverb (SMA CPL 3 *tsá-k=(b)éntjé*, HU *tsa-k=(b)ántjè*). For the INCPL, HU has *k=* < **ku* with labial onset vocalization INCPL 3 *k=o̲àntjè*), while SMA has the complex string *kua-k=* with *b-drop*: (INCPL 3 *kua-k=*(*b*)*èntjé*) which alternates with 1SG *kuí-k=*(*b*)*èntjé*. The forms *kua=/kui=* are added to the INCPL form in a phenomenon of proclitic chain

reanalysis that we treat here as 'incompletive overmarking'. The forms stem from a reanalysis of old INCPL formations of Class II.c/2 such as **ku=(b)a+* and **ku=(b)i+* as exponents for the INCPL, resulting in *kua=/kui=* secondary proclitics, with the vocalic elements being a splinter of the vocalized INCPL stem allomorph: *k=(b)è+ntjé* > ***k'uè***+*ntjé* > *k'uè-**k***=*è+ntjé.* > SMA *ku**í**-k=èntjé.*

'plant'		SMA	HU	*'play'*		SMA	HU
NTR	3	*bé+ntjé*	*bá+ntjè*	NTR	3	*sí+ská*	*sí+ská*
	1SG	*bè+ntjé*	*ba̲a+ntjèe*		1SG	*sì+ská*	*si̲i+skáa*
	1PL.INCL	*bé+ntjé*	*fa+ntjè*		1PL.INCL	*si+ská*	*nì+ská*
CPL	3	*tsá-k=é+ntjé*	*tsa-k=á+ntjè*	CPL	3	*kì=sí+ská*	*ki=si̲i+ská*
	1SG	*tsà-k=è+ntjé*	*tsa-k=a̲a+ntjèe*		1SG	*kì=sì+ská*	*ki=si̲i+skáa*
	1PL.INCL	*tsà-k=é+ntjé*	*tsa-k=a+ntjè*		1PL.INCL	*ní+ská*	*ki=nì+ská*
INCPL	3	*kua-k=è+ntjé*	*k=o̲àntjè*	INCPL	3	*kue=sí+ská*	*si̲ì+ská*
	1SG	*kuí-k=èntjé*	*k=o̲antjèe*		1SG	*kuí=sí+ská*	*si̲i+skáa*
	1PL.INCL	*kuá-k=èntjé*	*kj=oántjè*		1PL.INCL	*kuì=ni+ská*	*si̲+ská*
IC		Class I/a	Class II.c/3	IC		Class IV	Class IV
Preverb		*bé+*	*bá+/ fa+*	Preverb		*sí+/nì+*	*sí+/nì+*

Table 9. SMA *vs* HU inflectional class classes. Pike's (1948) and ALMaz data

As shown in Table 9, dialect contrasts involving inflectional class IV, with subconflative stem formative allomorphs *sí+/nì+* for HU and SMA *sí+ská* 'play', are less intricate, but there are deviations in SMA. First, SMA does not use TAMV & person asymmetry in the neutral paradigm, as HU does: NTR 3 *sí+ská*, 1SG *sì+ská*, 1PL.INCL *si+ská*. In contrast, in SMA a stem *ni+ská* is used in the subconflative subset. Absence of asymmetric TAMV/person marking happens in HU in the INCPL (3 *si̲ì+ská*, 1SG *si̲i+skáa* and 1PL.INCL *si̲+ská*). As with the other cases, all these phenomena do not apply to the verbs in question only. They are found generalized throughout inflectional class IV. These are highly regular patterns for the SMA subdialect as a tiny segment of the Central Highlands dialect.

7. San Felipe Jalapa de Díaz: a Piedmont town dialect

The San Felipe Jalapa de Díaz (JD) dialect is a Piedmont variety, resorting to the Western Lowlands segment of the Mazatec diasystem. It is typically a three level tone dialect, without mid-high tone. As we'll see, it shares to some extent several important inflectional tone processes with Huautla, as for example the preverbal tone lowering contour in NTR 1SG and all INCPL cells. Table 10 shows a sample of phenomena differentiating inflectional

class patterns between these two town dialects, pointing at a few processes still uncovered by the previous comparison.

First, Jalapa de Díaz Mazatec has the *y'e+* preverb allomorph in inflectional class Class I/a for CPL –a process we indicate as 'CPL.PV' (i.e. completive preverb collocation): NTR 3 ***b'é****+xá* '(s)he sends' *vs* CPL 3 ***y'e****+xá* '(s)he sent'. As most preverbs, the *y'e+* prefixe is a light verb on its own, which conveys a trajectory meaning, namely 'take away'. It appears autonomously in CPL 3, but it is bound in other subject agreement cells: CPL 1SG *kik'**ie**xá* ← *kik='**ye**+xá*, 1PL.INCL *kik'**ie**xáa* ← *kik='**ye**+xáa*, with proclitic *kik=* proclisis and glottal leftward realignment as a juncture surface phenomenon. INCPL marking is fairly similar to what happens in Huautla. These processes are consistently repeated for *b'é+ntje* 'plant', with two remarkable facts: i) the inflectional class shift in Jalapa de Díaz to Class I/a; ii) the root tone class shift, with mid-high tone neutralization in Jalapa de Díaz, as *b'é+ntj**e*** has a mid root tone, whereas in Huautla the cognate *bántj**è*** has a mid-high tone.

Class I/a *'send'*		JD	HU	*'plant'*		JD	HU
NTR	3	*b'é+xá*	*b'é+xá*	NTR	3	*b'é+ntje*	*bá+ntjè*
	1SG	*b'e+xá*	*b'ee+xáa*		1SG	*b'e+ntje*	*baa+ntjèe*
	1PL.INCL	*b'e+xáa*	*b'è+xá*		1PL.INCL	*b'e+ntjée*	*fa+ntjè*
CPL	3	*y'e+xá*	*tsak='é+xá*	CPL	3	*y'e+ntje*	*tsak=á+ntjè*
	1SG	*kik='ie +xá*	*tsak='ee+xáa*		1SG	*y'e+ntje*	*tsak=aa+ntjèe*
	1PL.INCL	*kik='ie+xáa*	*tsak='è+xá*		1PL.INCL	*y'e+ntjée*	*tsak=ja+ntjè*
INCPL	3	*k='ue+xá*	*k='oè+xá*	INCPL	3	*k='ue+ntje*	*k=oà+ntjè*
	1SG	*k='ue+xá*	*k='oe+xáa*		1SG	*k=oà+ntjè*	*k=oa+ntjèe*
	1PL.INCL	*k='ue+xáa*	*k='oé+xá*		1PL.INCL	*k='ue+ntjée*	*k=joá+ntjè*
Inflectional class		Class I/a		Inflectional class		Class I/a	Class II.c/3
Preverb		*b'é+*	*b'é+*	Preverb		*b'é+*	*bá+/fa+*
Processes		[NTR] Non subclf PV [CPL] CPL PV & kik'=y'é- [INCPL] k'=b VOC	[NTR] Non subclf PV [INCPL] k'=b VOC	Processes		[NTR] IC Shift [CPL] CPL PV [INCPL] k(')=b VOC	[NTR] Non subclf PV [INCPL] k(')=b VOC

Table 10. Sampling of basic paradigms for the survey of Mazatec verbal inflection: Class I-II.

The next matrix in Table 11 shows striking discrepancies. Jalapa de Díaz NTR 3 *b'é+jña'ma* '(s)he hides' matches HU *b'é+ma*, with further stem template extension: NTR 1SG *ba+t'ejña+'ma*, and 1PL.INCL *b'i+ntsjuba+'ma*, becoming unpredictable from a HU standpoint, as compared to the HU smooth default set based merely on *b'é+* preverbation. The JD preverb string complexification (*b'é+jña'ma* → *ba+**t'e**jña+'ma*) and the increase in conflation entail

something utterly new in the Mazatec inflectional class system as far as we have been able to see up to now. What is a single default class in one dialect may split into a threefold preverbal paradigm (*b'é+/ba+/b'i+*) –a situation pointed at with *b'é+/&* under the table, in the row describing preverb collocations. Nonetheless, the process stem template allomorphy is not legitimated only by preverb string complexification, but also by the inner compounding structure of the stems: NTR 1SG *bat'ejña'ma vs* 1PL.INCL *b'intsjuba'ma* imply a *jña* positional 'sit, stay' formative root in the former, and a *ntsjuba* 'mouth' formative root in the latter –now combined with the formative *'ma*,[11] a polysemic root with intricate Mode & Voice values such as 'can' and 'oneself' (median, or reflexive).

'hide'		JD	HU
NTR	3	*b'éjña'ma*	*b'éma*
	1SG	*bat'ejña'ma*	*b'ee'màa*
	1PL.INCL	*b'intsjuba'ma*	*b'è'maà*
CPL	3	*y'ejña'ma*	*tsak'éma*
	1SG	*tsat'ejña'ma*	*tsak'eemàa*
	1PL.INCL	*tsini'ma'an*	*tsak'èmaà*
INCPL	3	*k'uejña'ma*	*k'oè'ma*
	1SG	*k'uát'éjña'ma*	*k'oe'màa*
	1PL.INCL	*k'uintsjuba'ma*	*k'oémaà*
Inflectional class		Class II.c/2'	Class I/a
Preverb		*b'é+/&*	*b'é+*

Table 11. Comparative analysis of the Huautla and Jalapa de Díaz dialects.

In JD, not only do the completive forms entail conflation breaking: CPL 3 ***y'e****jña'ma*, 1SG ***tsa****t'ejña'ma*, 1PL.INCL ***tsi****ni'ma'an*. They also induce stem template allomorphy: CPL 3 *y'e**jña'**ma*, 1SG *tsa**t'ejña**'ma*, 1PL.INCL *tsi**ni'ma'**an*. From this quick survey of a few basic inflectional classes in JD and HU, we can draw a few conclusions. First, diasystemic complexity is more local than pervasive. As compared to Huautla, it may even turn out to be less conspicuous than in a subvariety of Central Highlands, as SMA. Outstanding processes are mostly preverbal allomorphy in the CPL paradigm for Class I/a (NTR *b'é+* ↔ CPL *y'é+*), inflectional class shifts (Class I/a ↔ II.c/3, or with further modification of a Class II.c/2' ↔ I/a), and more locally in the system, a striking entropy of complexification processes, such as *subconflative breaking, string complexification* and *stem templatic*

[11] As in Spanish, one (suppletive) stem allomorph for *esconder(se)* 'hide (oneself)' would be *embocarse* for 1PL, whereas another would be analogous to *acurrucarse* or *agacharse*, for 3 SG & PL, etc.

allomorphy, induced by suppletive roots for some lexemes, as for the verb ‘hide’ in Table 11.

8. San José Independencia: A Central Highlands/Western Lowlands transitional dialect.

With the next variety, San José Independencia, we will face mostly one new process, which blurs the disjunction between rules of stem selection (inducing default or subconflative sets of stems) and morphophonological rules (activated especially on labial onsets in preverbal domains, as can already be seen), namely *subconflative morphophonological transfer*. This is a mechanism by which b-dropping extends from the CPL to the Neutral, without any concatenation licencing of the labial onset deletion, as exemplified in Table 12. This phenomenon occurs only for the cells -3/1SG, according to the split subject agreement marking constraint (i.e. *subconflation*). For convenience, we have also included information about 1PL.EXCL. In the INCPL paradigm, a fine-grained morphophonological rule on vowel quality points at incompletive overmarking, with stem initial domain INCPL 3 *k=**'oe***+ ← |*ku=b’é*+| *vs* other subject agreement *k=**'ue***+.

	3	1SG	1PL.INCL	1PL.EXCL
NTR	***b'è***+*t'añòn*	***b'e***+*t'añòn*	***'e***+*t'añòn*	***'e***+*t'añù=jin*
	Default subset		**Subconflation morphophon. transfer**	
CPL	*tsik=**'e***+*tañòn*	*tsik=**'e***+*ñòn*	*tsik=**'e***+*t'añòn*	*tsik=**'e***+*tañon=jin*
	tsik'= b-drop & stem allomorphy			
INCPL	*k'=oe+tañòn*	*k'=ue+t'añòn*	*k'=uè+t'añòn*	*k'=uè+t'àñòn=jin*
	k'oe/k'uè= **incompletive overmarking**			

Table 12. Class I/a. bètañòn ‘braid’ in San José Independencia

Subconflation transfer seems to be a recent phenomenon in this dialect, embedded in the Western Lowlands, on the shores of the Miguel Alemán dam, but which most probably migrated from the Central Highlands less than a century ago. Several items will be necessary to check the robustness of the process –e.g. Table 13 shows evidence that some verbs do not show it: *bèxkia* ‘write’ and *b’èsee* ‘whistle’, while others regularly do.

This process is easier to account for when one considers that the San José Independencia transitional dialect is geographically embedded in the San Pedro Ixcatlán area. In this dialect, onset dropping in the CPL cells is strong: NTR 3 *wí+tiñon* ‘s/he braids’, 1SG *wi+tiñon*, 1PL.INCL *wì+tìñon* compared with CPL 3

ì+tìñon, 1SG *i+tiñon*, 1PL *a+kìtiñon,* which points to an inversion of the preverbs of Class II.c/2: *wi-/wa-* instead of *ba-/bi-*. Although the *i+* reflex resorts to a reduction of the Lowlands completive preverb *y'é+* (> *yi+* > *i+*), this trend still suggests that some kind of morphophonological conflation transfer started first in the CPL in the regional centre, before being extended later to the neutral aspect subject to subconflation rules. A further incentivate may be the b-dropping after CPL *tsik'=* proclisis. But as such, it might not be sufficient as many others dialect do have b-drop in the completive without developing the *subconflative morphophonological transfer option* in their inflectional system.

In Table 13 we have chosen verbs of Class I/a, which is the most regular. It is represented here by six verbs: *b'èxoan* 'boil'*, b'èts'a* 'beg', *b'èntjo* 'fan', *bènìji* 'bury', *b'èes'e* 'whistle', and *b'èxkia* 'read'. In these verbs, we observe that the CPL also has conflation by means of the clitic *tsik= vs jètsik=* (*jè* is a resultative light verb meaning 'finished', see Kirk 1966). The verb *b'èt'añòn* in Table 12 does not have it because it is a compound with a polysyllabic root, consisting of *t'a* ← **ta'a* 'aside' plus *ñòn*, an adverbial root of intensity meaning 'strongly', 'much'. The templatic constraint also accounts for the presence of a complex proclitic for the INCPL preverb in the form of *sa'ak'oe=/sa'ak'ue=x'o.* We see a similar constraint in operation with the verb *b'è+nìji* 'bury' which has *jetsik=* for the CPL in all persons. From now on, we only segment the leftward domain, whereas AGR will not be further analyzed in order to focus on the leftmost domain of the inflected lexemes.

For verbs of subconflative Class IV, in the CPL we observe a process of double subconflation (or twofold asymmetry), as not only does the preverb alternate, but also the TAMV/subject agreement clitic chain (3 *kì=**sì**+ská*, 1SG *kì=**si**+skà vs* other ***jè**ki=**ni**+skà*). In the INCPL, a proclitic *sa'a=* concatenates to a default (i.e. non subconflative) stem. This clitic widely occurs in most Mazatec dialects with a meaning of proximate future.

The San José Independencia sub-dialect shares much with the Central Highlands dialect, as an enclave embedded in the Western Lowlands. But it is remarkable for the resilience it shows concerning patterns of subconflation by extending them to new structures, such as for example the new differential marking occurring in the Neutral aspect from the extension of the reduced form of the preverb resulting from b-dropping in the CPL. The next set of data stems from another transitional dialect (Soyaltitla), though from a different source: a

Southwestern Highlands dialect (Mazatlán) under influence of the Central Highlands dialect (Huautla) area.

Class I/a	ASP	3	1SG	1PL.INCL	1PL.EXCL
bèxoan	NTR	*b'è+xoan*	*b'e+xò*	*'e+x'o*	*'e+x'ojin*
'boil'	CPL	*tsik='è+xo*	*tsik='e+xo*	*jètsik='e+xo*	*jètsik='e+xojin*
	INCPL	*sa'ak='oe+xo*	*sa'ak='ue+x'o*	*sa'ak='ue+x'o*	*sa'ak='ue+x'ojin*
bèts'a	NTR	*b'è+ts'a*	*b'e+ts'a*	*'e+ts'a*	*'e+ts'ajin*
'beg'	CPL	*tsik='e+ts'a*	*tsik='e+ts'a*	*jetsik='e+ts'a*	*jetsik='e+ts'ajin*
	INCPL	*(sa'a)k='oe+ts'a*	*sa'ak='ue+ts'a*	*sa'ak='ue+ts'a*	*sa'ak='ue+ts'ajin*
bèntjo	NTR	*b'è+ntjo*	*b'e+ntjo*	*'e+ntjo*	*'e+ntjojin*
'fan'	CPL	*tsik='e+ntjo*	*tsik='e+ntjo*	*jetsik='e+ntjo*	*jètsik='e+ntjojin*
	INCPL	*sa'a-k='oe+ntjo*	*sa'ak='ue+ntjo*	*sa'ak='ue+ntjo*	*sa'ak='ue+ntjojin*
bènìji	NTR	*b'è+nìji*	*b'e+nijian*	*'e+nijiàn*	*'e+nijìjin*
'bury'	CPL	*jetsik='e+niji*	*jetsik='e+nijian*	*jetsik='e+nijiàn*	*jetsik='e+nijijin*
	INCPL	*sa'ak='oe+niji*	*saak='ue+nijian*	*sa'ak='ue+nìjiàn*	*sa'ak='uè+nìjijin*
bèes'e	NTR	*b'è+es'e*	*b'e+es'e*	*b'e+es'e*	*b'e+es'ejin*
'whistle'	CPL	*jek'a=be+es'e*	*jetsik='e+s'e*	*jetsik='e+s'e*	*jètsik='e+s'ejin*
	INCPL	*sa'ak='oe+es'e*	*sa'ak='ue+es'e*	*sa'ak='ue+es'e*	*sa'ak='ue+s'ejin*
bèxkia	NTR	*b'è+xkia*	*b'e+xkia*	*b'e+xkia*	*b'e+xkiajin*
'read'	CPL	*ka=b'e+xkia*	*jetsik='e+xkia*	*jetsik='e+xkia*	*jètsik='e+xkiajin*
	INCPL	*sa'ak='oe+xkia*	*sa'ak='ue+xkia*	*sa'ak='ue+xkia*	*sa'ak='ue+xkiajin*

Table 13. Class I/a verbs in San José Independencia

Class IV	ASP	3	1SG	1PL.INCL	1PL.EXCL
s'ìsk'à	NTR	*sì+ská*	*si+skà*	*ní+skájin*	*nì+ská*
'play'	CPL	*kì=sì+ská*	*kì=si+skà*	*jèki=ni+skàjin*	*jèki=ni+skà*
	INCPL	*sa'a=sii+skà*	*sa'a=si+ska*	*sa'a=si+skajin*	*sa'a=si+ska*
sítsjó	NTR	*sí+tsjó*[12]	*sì+tsjò*	*nì+tsjòjin*	*nì+tsjó*
'toast'	CPL	*kì=sì+tsjò*	*kì=si+tsjò*	*kì=nì+tsjòjin*	*ki=ni+tsjò*
	INCPL	*sa'a=sii+tsjò*	*sa'a=si+tsjò*	*sa'a=si+tsjòjin*	*sa'a=si+tsjò*
sichìkon	NTR	*si+chìkon*	*si+chikon*	*nì+chìk'onjin*	*nì+chik'on*
'bless'	CPL	*jeki=sì+chikon*	*jeki=si+chikon*	*jèki=ni+chikonjin*	*jeki=ni+chikon*
	INCPL	*sa'a=sii+chikon*	*sa'a=si+chikon*	*sa'a=si+chikonjin*	*sa'a=si+chik'on*

Table 14. San José Independencia: inflectional class IV

[12] The back mid vowel /o/ in lexical roots is realized most of the time closer to a high back vowel /u/, as in Lowland dialects.

9. Soyaltitla: A transitional Southwestern/Central Highlands subdialect

In Table 15, Soyaltitla *wè+xtè* 'wrap' (Mazatlán Villa de Flores (henceforth MzVF, data from Léonard 2013) *bí+xtié*, HU *b'é+tjé*) suggests inner variation, as our informant hesitates in many cells of the Neutral paradigm between a preverb *wè+* and a *wì+*.[13] As far as rules of exponence are concerned for agreement suffixes, this variety matches enclitic concatenation constraints as in Mazatlán: NTR 1SG *wì+xtè=**'àn***, 2SG *wè+xtè=**jì***, 1PL.EXCL *wì+xtè=**jìn***, with minimal vowel fusion –except for 2PL *wì+xtè* ← *wì+xtè-o* ← |*wì+xtè=jon*|. This deletion process of desinential 2PL *-o* is widely spread at the periphery of the Central Highlands (e.g. San Andrés Hidalgo, where a subdialect of HU is spoken) and in the Northwestern Highlands (e.g. Santa Cruz Acatepec), according to our fieldwork observations. Nonetheless, the main constraint seems to preserve the root stem vowel from any juncture modification, either through enclisis or through desinential deletion.[14] Here, for the sake of conciseness, we will not deal with the completive paradigm.

'wrap'		Soyal.	MzVF
NTR	3	*wè+xtè, wì+xtè*	*bí+xtié*
	1SG	*wì+xtè'àn*	*bi+xtia*
	2SG	*wè+xtèjì*	*bí+xtièji*
	2PL	*wì+xtè*	*bí+xtièjón*
	1PL.EXCL	*wì+xtèjìn*	*bí+xtièjin*
INCPL	3	*sè=kù+xtè*	*kui+xtiè*
	1SG	*sè=kù+xtèà*	*kui+xtia*
	2SG	*sè=kù+xtèjì*	*kui+xtieji*
	2PL	*sè=kù+xtéjón*	*kui+xtiejon*
	1PL.EXCL	*sè=kù+xtèjìn*	*kui+xtiejin*

Table 15. Soyaltitla *wèxtè*: 'wrap' (MzVF *bíxtié*)

As compared to MzVF, the town dialect proper, of which Soyaltitla can be considered as a rural satellite, the following trends are to be seen:

i) No betacism (*w* > *b*) in Soyaltitla, unlike in the other main Mazatec town dialects (HU, JD, MzVF, etc.), at least in this data, provided by a young informant.

ii) Diphthongization of the mid front vowel (*e* > *ie*), as in MzVF and San Pedro Ixcatlán, and unlike HU and JD.

13 Here rules of stem selection therefore appear somewhat unstable, though the CPL paradigm and the comparison with the including dialect (Mazatlán Villa de Flore, hence MzVF) leave no doubt about the basic inflection class categorization of this lexeme: as *wì+* prevails in the CPL as a default paradigm, it classifies as a Class I.c/1 verb in our model (see Table 5)

14 As in Mazatlán, tone patterns tend to follow propagative plateau constraints (see Léonard & Fulcrand 2016: 184-191), i.e. the stem dominant level tone spreads in both directions, leftward and rightward, as can be seen in most of the cells in Table 16, except for INCPL 2PL *sèkù**xtéjón***, where a h-H contour shows up.

iii) In MzVF, the structural option *k*= & *b-voc* (INCPL 3 *kuixtiè*) occurs in the incompletive as in Huautla, instead of the proximal future proclisis in Soyaltitla, with preverb string complexification (*sè-k=ù+xtè*). Nevertheless, the *b-voc* is active in both varieties for this inflectional class I.c/1 lexeme.

iv) The Soyaltitla rules of exponence component seem dominantly akin to MzVF enclisis strategy, although Soyaltitla does have enclitic agreement strategies hinting at Central Highlands patterns. In Soyaltitla, agreement desinential dropping also happens in NTR 2PL *wì+xtè vs* MzVF *bí+xtiè=jón* –a convergence with other rural Central Highlands and Northwestern Highlands varieties, as already pointed out. More exemplary paradigms of Class I/a are given in Table 16.

In the following set of data (the 'control set' for inflectional class I/a), other inflectional class I/a lexemes show a variegated array of the same rules of stem selection and rules of exponence processes and fine-grained contrasts, especially in tone inflection. However this idiolect might seem to cling to a simplex tone implementation strategy through mere plateauing of the root tone class's main prosodeme –as also observed in MzVF (Léonard 2013).

	LEXEME 1	3	1SG	1PL EXCL
'close'	NTR	*wè+xchjá*	*wè+xchjà*	*wè+xchjàjìn*
	INCPL	*kuè+xchà*	*kuè+xchàà*	*kuè+xchájìn*
	LEXEME 2	3	1SG	1PL EXCL
plant'	NTR	*wè+ntjè*	*wè+ntjèà*	*wè+ntjéjìn*
	INCPL	*se=kuè+ntjè*	*sè=kuè+ntjèà*	*sè=kuè+ntjèjìn*
	LEXEME 3	3	1SG	1PL EXCL
'beg'	NTR	*wè+tsuà*	*we+tsuà*	*wè+tsuàjìn*
	INCPL	*kuè+tsà*	*kuè+tsàà*	*kuè+tsàjìn*
	LEXEME 4	3	1SG	1PL EXCL
'hide'	NTR	*wè+ñama*	*wà+tèxñama*	*wì+xñàmàjin*
	INCPL	*kuè+xñama*	*kuà+tèxñama*	*kui+ñamajin*

Table 16. Inflectional class Class I/a in Soyaltitla Mazatec

In lexeme 1, NTR 1SG *wèxchj**à*** opposes lemmatic NTR 3 *wèxchj**á*** with a root tone lowering (lowered h versus H on the root mora in *-chj**á***). In lexeme 2, the so-called 'Obligatory contour principle' or just OCP (see McCarthy 1986) makes a difference in NTR 1PL.EXCL *wèntj**é**jìn*, as opposed to plateauing NTR 3 *wèntj**è*** and 1SG *wèntj**è**à.* In lexeme 3, 1SG tone lowering occurs in the preverbal domain: NTR 3 *w**è**tsuà vs* 1SG *w**e**tsuà*. In the INCPL, the pseudo diphthong in the lemmatic form NTR 3 *wèts**uà*** changes to a monophthong for AGR3 & Other, with a *kuèts**à*** stem. As far as stem formation processes are concerned, the same proclitic strings observed as in Table 16 still hold; *subconflation breaking* also prevails in the INCPL.

Instead, *preverb string complexification* and *incompletive overmarking* are less favored than in SMA or SJ.

As for inflectional class IV for the lexeme *sì+stsjù*: 'toast' (MzVF *tsì+tsjò*: see comparative data in Table 17),[15] the same processes repeat just like for inflectional class I/a: *sè=* proclisis in the INCPL, which is subconflative, unlike in Huautla.

'toast'		Soyal.	MzVF
NTR	3	*sìstsjù*	*tsìtsjò*
	1SG	*sìstsjòà*	*tsitsjòa*
	2SG	*nìstsjùjì*	*nítsjòji*
	2PL	*nìstsjójón, nìstsjújón*	*nítsjùjòn*
	1PL.EXCL	*nìstsjùjìn*	*nitsjòjin*
INCPL	3	*sèsìtsjù*	*kuititsjù*
	1SG	*sèsìtsjóà*	*titsjoa*
	2SG	*sènistsjùjì*	*nitsjoji*
	2PL	*sènìstsùjón*	*nitsjujon*
	1PL.EXCL	*sènistsùjìn*	*nitsjojin*

Table 17. Soyaltitla *sìstsjù*: 'toast', MzVF *tsìtsjò* (HU *sítsjò*)

Table 18 provides the control set for inflectional class IV, with verbs such as *síská* 'play' *síxí* 'dry' *sìxkua* 'break up', showing how much most of the stem formation processes hold, with regular outputs.

	3	1SG	1PL EXCL	PROCESSES
NTR	*síská*	*sìskàà*	*nìskàjìn*	subcfl
INCPL	*sèsúsìskà*	*sèsùsìskà*	*sènìskàjìn*	subcfl & preverb string complexification
NTR	*síxí*	*sìxìà*	*nìxíjìn*	subcfl
INCPL	*sèsìxí*	*sèsìxíà*	*nìxíjìn*	*sè=* & subcfl
NTR	*sìxkua*	*sìxkuaa*	*nìxkuajin*	subcfl
INCPL	*sixkua*	*sìxkua'a*	*nixkuajin*	subcfl
NTR	*síchàjà*	*sìchàjà*	*nìchàjajin*	subcfl
INCPL	*sichaja*	*sìchájà*	*nichajajin*	subcfl

Table 18. Soyaltitla inflectional class IV: *síská* 'play', *síxí* 'dry', *sìxkua* 'break up', *sích'àjà* 'lose'

As to INCPL stem formation, in Soyaltitla, Class I/b undergoes both preverb string complexification and subconflation breaking: compare 3 & 1SG ***sèk****uatèxá* with 1PL ***kàtsìnìji****kuatèxàjìn* on the one hand, and the inner diversity of stem formation processes in the -3/1SG sector of the split subject

[15] The complex intervocalic onset *-sts-* in is a phonolexical hapax, limited to lexemes with an intervocalic coronal affricate.

agreement marking: INCPL 2SG ***kuà****tèxà*, 2PL, ***sèkua****tèxàjòn* 1PL.EXCL ***kàtsìnìjìkua****tèxàjìn*.

	3	1SG	2SG	2PL	1PL EXCL
NTR	*wàtèxá*	*wàtèxà*	*wàtèx'á*	*wàtèx'àjòn*	*wàtèx'àjìn*
INCPL	*sèkuatèxá*	*sèkuatèxá*	*kuàtèxà*	*sèkuatèxàjòn*	*kàtsìnìjìkuatèxàjìn*

Table 19. Class I/b. *wàtèxá* 'send'

The opacity of rules of exponence for agreement markers 2SG and PL in Soyaltitla is also intriguing. In 2SG no traces of a *=ji* enclitic, or of any lighter exponent (such as a desinential *-i* as in many other dialects) is to be seen (2SG NTR *wàtèx'á*, INCPL *kuàtèxà*), whereas in 2PL NTR *wàtèx'à=jòn*, INCPL *sèkuatèxà=jòn*, at least two forms do have a *=jòn* enclitic marker.

Once more, the hypothesis of *templatic complexity* may be useful: unlike HU *b'éxá*, which has the simplex form of a couplet (PV+Root), *wàtèxá* resorts to a compound *wà+#tè##xá#* (PV+ *tè* 'surface' + *xá* 'work, task'). As such, it has a heavy stem template, inducing more *templatic allomorphy*. This might also explain the role here of distributed creakiness of the main root nucleus. It is absent in NTR 3 & 1SG, but present in trisyllabic stem allomorphs implying either phonological reduction of desinential elements, as in NTR 2SG *wàtèx'á* (for expected *wàtèxá=(j)i*), CPL 1PL.EXCL *ki=wàtèx'à=jìn* (not included in the chart). This might also explain agreement marker enclisis, such as 2PL *wàtèx'à=jìn*, while heavier templates or INCPL stems are preserved from this postlexical process. Here, one has to rely on manifold hypotheses to explain surface phenomena deeply embedded in both phonological and morphological structures –and confirming the G&K hypothesis more than P&P's.

10. *An xo'boo*: a peripheral Northwestern Highlands dialect

The next and last variety of our survey is the innovative San Lorenzo subdialect of the peripheral Northwestern Highlands, on the brink of the state of Puebla –this dialect is therefore often called *Mazateco poblano*. Its inflectional class system converges strongly with the Piedmont and Lowlands dialects, in spite of geographic distance. The vowel system of this dialect undergoes a pull & drag chain (vowel shift), consisting in high and mid vowel lowering and high back vowel retraction:

The Mazateco poblano Vowel Shift:
**i > i, e; *e > a; *a > o; u > ʉ*

In this Vowel Shift chain, /u/ → [ʉ] ↔ orthographic < ö > (although not used here: we will simply give preference to < u >, as the high back vowel reduction seems to us more allophonic than phonemic in most of the localities we have visited in the Poblano area) is a high back unrounded vowel. We take this drift to Vowel Shift as a model for deducing lexical forms out of surfacing postlexical forms. Table 20 presents both forms (lexical and postlexical, i.e. inputs and outputs): e.g. *watòo* 'write' derives from |*wetà'a*|; CPL 3 *koatòo* from |*ka=wetà'a*|; and INCPL 3 *lokatòo* from |*lak=wetà'a*|.[16] We'll first analyze lexemes from inflectional classes with a canonical labial, such as Class I/a *watòo* and Class II.c/2 *wota* in Table 20, then Class I/b *wokian* |bakien| 'eat' and a less canonical lexeme from Class II.c/2' *fikè* 'take away' in Table 21. In Table 22, we study the inflection of two lexemes from our Class IV: *sikè* 'make, do' and *simàn* 'love'. After this we then deal with less canonical inflectional classes in Table 23: Class V *sòkò* 'fall' and Class VII *sà* 'sing'. We end this brief survey of the San Lorenzo dialect with three motion verbs in Table 24: Class VII *fi* 'go', *fìì* 'comes', *fòwò* 'pass', in order to get a more complete picture of irregular stems.

The first impression given by the San Lorenzo (SL) data is that the main phenomena affecting rules of exponence are *neutralization* and *ending fusion*. Although AGR endings do not correlate with inflectional class assignement, it does nonetheless contribute to the overall complexity or simplexity of realizational forms. In SL, ending fusion (hence, ef) can be considered as a phenomenon increasing the weight of morphophonological rules applied on the lexical inputs, albeit producing shorter or lighter outputs. A tone shift in AGR endings also seems to happen in this dialect: 1SG *-a* > *-a̲*, 1PL.INCL *-à* > *a*. Ending fusion deletes the root vowel, whose slot is filled by the desinential vowel. Modalisation of onsets and nuclei prevails: *watòo* ← |*weta̲'a*| 'write' with *we+* (a general locative preverb, the same as *b'é+* in HU) plus *ta̲'a* (directional root for 'aside'), as English 'put down' (here literally 'put aside') ↔ 'write'. As in many peripheral dialects, the 1PL.INCL/1PL.EXCL opposition is weakened. The San Lorenzo data already pointed at this phenomenon. In this idiolect, we suspect the 1PL.INCPL *lokatàa̲* to stem from |*lak=wet**a̲**+**i̲**n*| - otherwise, the expected output of |*lak=wet**a̲**+**a***| would be *lokatòo*, not *lokatàa̲*.

16 For more details on the geolinguistic emergence and sociolinguistic patterns of this dialect, see Léonard (2014).

All these processes concern rules of exponence and juncture processes which resort to morphophonological rules, such as *w-drop* and *agreement markers merging* in the rightmost column, essentially a *V-drop* (vowel dropping) variable. Beyond these processes, the inflectional class formation processes are fairly simple. The *wota* 'cut' lexeme only undergoes an inflectional class shift (II.C/2 *ba/bi-* ↔ II.C/2' *wa/mi-*), with postlexical variation in the labial onsets of the preverbal allomorphs, as compared to types enumerated in Table 5. Rules of stem selection are nevertheless fairly predictable and regular, and although this inflectional system may seem baffling for a Central Highlands speaker, it actually attains a high degree of simplexity, through the systematic application of simple and regular rules of exponence and morphophonological constraints of ending fusion (or coalescence).

	3	1SG	2SG	1PL	PROCESSES
NTR	*watòo*	*watò*	*watì*	*watò*	
	|*weta'a*|	|*wet**a**+**a***|	|*wet**a**+**i***|	|*wet**a**+**a***|	agreement markers ending fusion (ef)
CPL	*koatòo*	*koatò*	*koatì*	*koatò*	*ka*= w drop
	|*ka=weta'a*|	|*ka=wet**a**+**a***|	|*ka=wet**a**+**i***|	|*ka=wet**a**+**a***|	agreement markers ef
INCPL	*lokatòo*	*lokatò*	*lokatì*	*lokatàa*	*lok*= w drop
	|*lak=weta'a*|	|*lak=wet**a**+**a***|	|*lak=wet**a**+**i***|	|*lak=wet**a**+**in***|	agreement markers ef & 1PL scambling
NTR	*wota*	*wotà*	*michi*	*micho*	inflectional class shift *ba/bi-* ↔ *wa/mi-*
	|*wate*|	|*wate+**a***|	|*wicha+**i***|	|*wicha+**a***|	agreement markers ef
CPL	*kowota*	*kòwòtà*	*komichi*	*komicho*	*ka*= w drop
	|*ka=wate*|	|*ka=wate+**a***|	|*ka=wicha+**i***|	|*ka=wicha+**a***|	agreement markers ef
INCPL	*lokota*	*lokotà*	*lokichì*	*lokichò*	*lok*= w drop
	|*lak=wate*|	|*lak=wate+**a***|	|*lak=wicha+**i***|	|*lak=wicha+**a***|	agreement markers ef

Table 20. SL: Class I/a. *watòo* 'write' and Class II.c/2. *wota* 'cut'

The next set of data in Table 21 confirms these trends: Class I/b *wokian* |bakien| 'eat' has a mechanism opposing a *wo*+ ← |wa+| preverb in the +3/1SG zone to a *wa*+ ← |we+| preverbal allomorph for -3/1SG stems, such as in *wa+chi* for NTR 2SG as much as for 1PL. This conflative pattern breaks up in the CPL 1PL cell with *ko=wo+kian*, instead of expected *ko=wa+chi*. The INCPL -3/1SG cells confirm the conflative split. As the stem for 'eat' tends to belong elsewhere to inflectional class VII (i.e. HU *kji+ne/ chji+ne*), this stem has been reanalyzed and shaped anew with cyclic preverbation of the split Class I/b inflectional class type. According to our fieldwork data,

this split subconflation widely occurs in other Northwestern Highlands varieties as well as in the Eastern Lowlands dialect. But the Northwestern Highlands Peripheric subdialect specifically uses it as a cyclical process for converting inflectional class VII stems into more canonical subconflative stems belonging to Classes I-II, as in below. The next item, *fi+kè* 'take away' exists elsewhere as well, with a light motion verb *fi+* as an inflectional class preverb, but it undergoes a variegated set of subconflation breaking processes, with a handful of preverbal allomorphs, as NTR *mi+, ma+*, CPL *ki+, sa+*, INCPL *wo+, wa-* in 3 *lok**o**kian* ← |lak=**wa**-kien| and 2SG *lok**a**chi* ← |lak=**we**-chi|.

	3	1SG	2SG	1PL	PROCESSES
NTR	*wo+kian*	*wo+kiàn*	*wa+chi*	*wa+chi*	subcfl *wo-/wa-*
CPL	*ko=wo+kian*	*kò=wò+kiàn*	*ko=wa+chi*	*ko=wo+kian*	*ko=* & subcfl+conflation breaking
INCPL	*lo=ko+kian*	*lo=ko+kiàn*	*lo=ka+chi*	*lo=ka+chi*	*lok=* b-drop & cfl
NTR	*fi+kè*	*fi+kè*	*mi+ki*	*ma+kè*	conflation breaking
CPL	*kì=kè*	*ki=kè*	*ki=kì*	*sa=kè*	*ki=*& conflation breaking
INCPL	*lòki=ke*	*loki=kè*	*loki=kì*	*lòka=kè*	*loki/-ka=* & conflation breaking

Table 21. SL: Class I/b Modified (I/b') *wokian* |bakien| 'eat' and II/c Modif. *fikè* 'take away'

Conflative split and breaking are also to be seen in Class IV for causative verbs, as in Table 22: *si+kè* 'make, do', *si+màn* 'love'. Here a mixed strategy of *proclisis vs bare stem* and appears for *si+kè* 'make, do', when the -3/1SG zone of the field of subconflation has a proclitic exponent in CPL 2SG *ko=na+kì*, whereas 1PL *na+kè* shows up as a mere suppletive stem, without TAMV proclisis – i.e. a bare stem. The subconflative patterns in INCPL 3 *si+ke* and 1SG *si+kè vs* 2SG *sa+kì*, 1PL *sa+kè* are distinct from the non-conflative, default pattern in Huautla, as seen above. Strikingly enough, the whole game seems a compromise, as lexical faithfulness of the *sa+* ← |*se+*| preverbal allomorph (as far as the onset is concerned) points to the application of a default strategy, whereas the *si+*/*sa+* vowel alternation points at split subject agreement marking. Nonetheless, proclitic chains differ for the verb *si+màn* 'love', which look more regular and less auxiliary-like: *ko=* is used for CPL except for 1PL, whereas *lo=* concatenates with all forms for the INCPL. For verbs of this class, rules of stem selection contrast the NTR 1PL *na+man* with CPL 1PL *sa+man*, a contrast which hardly occurs in other dialects.

	3	1SG	2SG	1PL	PROCESSES
NTR	*si+kè*	*si+kè*	*na+kì*	*na+kè*	subcfl
CPL	*ki=si+ke*	*ki=sì+kè*	*ko=na+kì*	*na+kè*	*ki/ko= vs* bare stem
INCPL	*si+ke*	*si+kè*	*sa+kì*	*sa+kè*	subcfl
NTR	*si+màn*	*si+màn*	*na+mi*	*na+man*	subcfl
CPL	*kò=sì+màn*	*ko=si+màn*	*ko=na+mi*	*sa+man*	*ko= vs* bare stem
INCPL	*lo=si+man*	*lo=si+màn*	*lo=sa+mi*	*lo=sa+man*	*lo=* & subcfl

Table 22. SL: Class IV *sikè* |siki| 'make, do', *simàn* |simen| 'love'

The data in Table 23 show verbs of coronal onset type classes. Here, the simplex trends observed above are no longer relevant. On the contrary, complexity prevails, as the enumeration of processes suggests, in the rightmost column of Table 23. Here we have SL: V. *sòkò* 'fall', VII. *sà* 'sing' (*cf.* HU *kaa* 'fall', *see* 'sing'). The former shows an inflectional class shift as seen before: an item belonging to inflectional class VII is promoted to a prefixal inflectional class, higher in the hierarchy designed in Table 5 above: *sòkò* ← |**sa**+ka| ← |sa+kaa|.

Moreover, conflation breaking, conflative split and stem template allomorphy all happen between the cells in the NTR: conflation breaking, as each cell has a specific preverb or proclitic string of its own (*cf.* light grey cells), stem template allomorphy as NTR 1PL suppletive *wè**xè*** (dark grey cell) opposes all the others with a ***kò*** root. In the INCPL enclisis of an oblique object marker occurs in 1 & 2SG: *kowoko=lò* ← |*ka=waka=le-a*| 'I'll fall' and *kokò=lì* 'you'll fall', instead of expected *lo=woko* & *lo=ko*. Here, CPL *ko=* seems to invade the INCPL cells, as in 1SG *ko=wokò* 'he'll fall', instead of *lo=wokò*.

	3	1SG	2SG	1PL
NTR	*sòkò*	*kowokò*	*sokì*	*wèxè*
	IC shift & subconflation breaking + stem template allomorphy (suppletive)			
CPL	*kòwòkò*		*kowòkì*	*kokìxè*
	ko/-ki= & stem template allomorphy			
INCPL	*kowokò*	*kowokolò*	*kokòlì*	*lokixè*
	ko(k)/loki=/=lo & subconflation breaking + stem template allomorphy			
NTR	*sà*	*sà*	*mìntì*	*mìnto*
	inflectional class shift & stem template allomorphy			
CPL	*kòsà*	*kosà*	*komìntì*	*komintò*
	ko= & stem template allomorphy			
INCPL	*losà*	*losà*	*losì*	*lokintò*
	lo(k)= (b-drop) Cfl.br & stem template allomorphy			

Table 23. San Lorenzo: Class V. *sòkò* 'fall', Class VII. *sà* 'sing'

As for *sà* 'sing', conflation goes with a motion preverb *mì-* and a suppletive root *ntò* (grey cells). In the INCPL, conflation breaking though, is obvious, as 2SG *losì* follows the same rules of stem selection constraint, with a *sà* lexical allomorph. On the other hand, 1PL *lok**intò*** follows the suppletive allomorphic pattern with root *nto* as in the other cells belonging to the subconflative zone of the α3/1SG equipollence. Moreover, *lo(k)=* (b-drop) is used, since the suppletive stem is of the *mintò* type, which suggests an inflectional class shift of the type VII > Class II.c/2 modified, i.e. II.c/2' (the *ba-/bi-* or *ba-/mi-* inflectional class, much used in the Northwestern Highlands segment of the diasystem).

We will conclude this sketch on inflectional class in the SL dialect with a few motion verbs in their free form as lexical verbs instead of light verbs. In this case, they are bound to a lexical root within a stem as prefixes. Table 24 shows how much this subclass of verbs is sensitive to suppletion and also provides evidence of their high level of entropy, as far as sets of processes are concerned. We would also suggest that a great part of the intricacy of inflectional patterns to be seen in the Mazatec diasystem flows from the *complexity* –instead of *simplexity*– models implemented in motion verbs. They make up nests, or matrixes of paradigmatic complexification: the *fi* 'go' lexeme shows subconflative split with the NTR cells already: 3 ***fi***, 2SG ***mì***, 1PL *mo-**nkì-i***. The *fìì* 'come' paradigm alternates with NTR 1SG *nzowò*, 1PL *nzowà vs* 2SG *nzofi*, and |ki| allomorphs in the INCPL. Nevertheless, syncretism may interfere to make patterns more simple, as for *fòwò* 'pass', whose suppletive patterns are to some extent compensated by syncretic trends.

	3	1SG	2SG	1PL	PROCESSES
NTR	*fi*	*fè*	*mì*	*monkìi*	Subconflation
CPL	*kii*	*kofè*	*kìi*	*sònkì*	*ko/so=* & Subconflation breaking & Stem allomorphy
INCPL	*ki*	*lòki*	*nokè*	*lòkònkì*	*lò(kò)/no=* & Stem allomorphy
NTR	*fìì*	*nzowò*	*nzofi*	*nzowà*	Distributed Conflation shift & Breaking
CPL	*kòfi*	*kòfè*	*kòfi*		*kò=* & Subconflation
INCPL	*lòkìì*	*lok'è*	*lokìi*	*lok'ì*	*lò=* & Default stem
NTR	*fòwò*		*tsìtì*	*metò*	Subconflation
CPL				*sìtò*	Subconflation
INCPL	*lòko*		*lokitìi*	*lokità*	*lòk/-o-=*

Table 24. San Lorenzo: VII. *fi* 'go', *fìì* 'comes', *fòwò* 'pass'

These different phenomena in Table 24 give an overall final view of the issues in Mazatec inflectional class complexification and simplexification processes. It reminds us how strongly the system is determined by motion verbs, especially *trajectory* verbs (directional, itive and so forth). The most simplex inflectional classes are more of the general location type (inflectional classes I & II). The most complex classes resort to motion light verbs of the trajectory and locative type. As for rules of stem selection, Mazatec dialects converge towards conjugating the preverb for TAMV categories and marking the agreement in the preverb allomorphs. The same trick applies for rules of exponence and morphophonological rules in the root vowel ending. All in all, Mazatec dialects have contrived ingenuous ways of completing or specifying TAMV or even person agreement – according to the split subject agreement optional constraint– through a full house of TAMV clitics.

11. Distribution patterns

We can now summarize the most important processes or mechanisms met in this survey, as in Table 25. This matrix displays an array of eight main parameters, from the most trivial (inflectional class shift) to the most specific (morphophonological subconflative transfer). Parameters 1-3 make up the most powerful subset of diversification processes, and are either ubiquitous or widely diffused in dialects and sub-dialects. Inflectional class shift, for example, although qualified here as ubiquitous, implies nevertheless at least 11 x 2 = 22 theoretical options, each of which may define dialects and subdialects both phylogenetically (as *isoglottic* areas) or ontogenetically (as *typological* subsets of varieties). Parameters 4-8 are endemic trends in the diasystem, so that these processes are less correlated to phylogenetic determinism in the dialect Stammbaum. Parameter 9 is typically a local innovation, resorting more to the *hapax* type than to a sub-dialect characteristic.

Geolinguistic distribution of these parameters still awaits description. The task is made difficult by the discrepancy of lexical options throughout the diasystem. For the sake of consistency we selected here a number of lexemes which tend to be cognates (verbs such as ‘send’, ‘toast’, ‘hide’, ‘die’, wash’, ‘read’, etc.), but many other verbs do not necessarily fulfill this prerequisite, as the compounding strategy deeply embedded in Mazatec (and Oto-Manguean) lexical morphology, allows many local solutions.

Nevertheless, the ALMaz currently endeavours to survey at least one hundred cognate verbs in the Mazatec diasystem. The materials analysed here were gathered through a pilot study aiming at detecting the verbs which are more likely to be cognates in the diasystem. Results are encouraging, as it made possible to revisit former taxonomies (such as Jamieson 1982, Pike 1948), and to work out more parsimonious models for the diasystemic survey of Mazatec inflectional classes.

	Pattern sampling	Example	Diasystem
1. INFLECTIONAL CLASS SHIFT	(SL) I/a ⇔ (HU) I.C/1	NTR.3SG ***b'éya*** ⇔ ***biyaa***	*ubiquitous*
2. COMPLETIVE PREVERB	CPL *y'é-*	NTR.3SG ***b'éxá*** / CPL.3SG ***y'éxá-***	JD, Midlands & Lowlands, SL
3. SUBCONFLATIVE SPLIT	I/a complexified	***b'éxá*** ⇔ ***b'éxá*** / ***bi****xá-*	Northwestern Highlands
4. SUBCONFLATIVE BREAKING	+3 *vs* 1SG *vs* 2SG	MzVF (Soyaltitla) 3SG ***b'éñama*** / 1SG ***ba****texñama*/ 1PL.EX ***bi****xñamajin*	MzVF & endemic
5. STRING COMPLEXIFICATION	TAMV prefixal strings complexified	Ayautla 3 CPL ***tsek'****=etañón* > ***ní****=****tsík'****=etanión*	endemic
6. NEUTRALIZING SUBCONFLATIVE ASYMMETRY	Default preverb as an option	HU NTR 3SG ***sí****xá vs* 2SG ***nì****xái*, but INCPL.3SG ***sii****xá vs* INCPL 2SG ***si****xái*	HU & endemic
7. INCOMPLETIVE OVERMARKING	Complexification of INCPL prefixal or proclitic contrasts	SMA INCPL 3 ***kuak'****=èntjé vs* 1SG ***kuík'****=èntjé vs* 1PL.INCL ***kuák'****=èntjé*	SMA & endemic
8. STEM TEMPLATIC ALLOMORPHY	Stem Templatic allomorphy, stem suff. derivation	MzVF (Soyaltitla) MzVF (Soyaltitla) 3SG *b'é****ñama*** / 1SG *ba****texñama***	MzVF & endemic
9. MORPHOPHONOLOGICAL SUBCONFLATIVE TRANSFER	Subconflative pattern applies to an onset	NTR 3SG ***b'è****xoan vs* 1SG ***b'e****xò vs* 1PL.INCL ***'e****x'ojin*	SJI

Table 25. Matrix of simplex processes for inflectional class diversification in Mazatec

12. Conclusion & prospects

In this chapter we have provided first hand data on a few Mazatec dialects highlighting at least five important issues beyond the specific properties already known about this language in terms of inflectional complexity.

The first point could be called the *empirical gap*: Mazatec inflectional class systems are still awaiting comprehensive diasystemic description and modeling. In other words, we need more fieldwork. Since Pike's (1948) and Jamieson's (1982, 1988) seminal work, most scholars have relied on secondary data. They have done well, and theoretical achievements have

indeed been very significant on this empirical basis. In the realm of phonology, the G&K's model of voice quality and autosegmental realignment, which turns out to be heuristic for linguistic typology as a component of universal grammar, has been contrived without any fieldwork, on the basis of P&P's former model. Nevertheless, especially in morphology and morphosyntax, we still miss the inner typological diversity of Mazatec as a (diasystemic) whole. A rapid glimpse at some paradigms from two subdialects close to HU, and at a dialect yet undescribed, like SJ, Soyaltitla or SL, shows that Mazatec is still a continent to discover, of which we know but a few places.

The second point could be called the *heuristic metasynthesis*: synergy between models gives encouraging results. In Table 5, we made Mazatec inflectional class models interact in order to get an analytical grid to observe variation. The *ALMaz* model sketched in this matrix draws a synthesis between many competing models, such as Bull's (1984) morphophonological (cyclical rules) model (for morphophonological rules), Pike's (1948) Light Verb+Polyvalent Root Model and Jamieson's (1982) inflectional class stem model. Such a synthesis resorts to a dimension of complexity, which could be called *metatheoretical complexity*.

The third point should be called the *unavoidable simplexity* principle. Simplex patterns and 'tricks of the trade' in local inflectional class self-organization should be the main targets, when trying to disentangle complex inflectional patterns. Both complexity and simplexity are interdependent terms of a dialectic process in qualifying facts. Many components of the Mazatec verbal inflectional class system, which had been described or considered as complex, may turn out to be actually simplex, as the prosodic constraints induced by the Lexical Tone Taxonomy, which belongs to Pike's root tone class model. For Pike, Mazatec roots are polyvalent, and give the lexical tone of the stem –they literally 'give the tune' to all the concatenated elements, which simply adjust according to local constraints, such as the NTR & CPL 1SG preverbal tone lowering (for more details, see Léonard & Fulcrand 2016). Preverbs may only have a high (H) or a mid tone (h) at lexical level, as in Huautla (inherited from their lexical origin). Whatever the morphosyntactic concatenation may be with any PV or agreement marker endings, they still cling to their inherited lexical tone pattern, so that both root tone class and PV tone are robust lexical properties, only modified locally, i.e. by 1SG or INCPL marking. Most changes happen elsewhere, especially in the

preverb complex, resorting to two simple phenomena: either an OCP adjustment (merely allotonic), or plateauing. The same applies to rules of stem selection: some dialects activate conflation in the INCPL (like SMA), others do not (HU). Some have complex CPL or INCPL proclitic strings, other do not.

The fourth point should be labelled interaction in 3-D. Interaction holds firmly as a basic concept in complexity research. But not only as analogy or transfer of structural properties, as we have seen so far: embedding and dependency are interactions too. Any intricate aspect is interactive, as are such processes like string complexification, subconflation breaking and incompletive (or any other marked TAMV category) mingling, as we have seen so often in the data surveyed here. For instance, not only do preverb string complexification, conflation breaking and incompletive overmarking interact but so do the various TAMV markers between each other. In this way, microsyntax, morphosyntactic seriality and lexical formation constraints interact as much with these parameters as they do between themselves, so that inflectional classes are constantly renegotiated by speakers within the pool of existing solutions the diasystem offers to them –and this process starts as early as very young speakers acquire the language.

The fifth point leads us back to a prerequisite: models producing inventories and simulations for the way inflectional taxonomies work in grammar and how they are embedded in the lexicon. We need of a comprehensive multidialectal Mazatec dictionary, with a reliable inflectional class grid to sort verbs, and a solid polyvalent root inventory, taking alignment and valence into account. And while we are in want of an inflectional class model which might simulate to some extent how these systems are learned by native speakers, in this paper we have made a first attempt.

The *ALMaz* project intended to address all these questions and however scarce the resources may still be, we hope this paper will foster more attempts in this direction. However, right from the beginning, the motto should be: focus strongly on grammar and modeling already from the start, and try to disentangle intricacy into simplexity.

Abbreviations

Localities: HU: Huautla de Jiménez; JD: Jalapa de Díaz; MzVF: Mazatlán Villa de Flores; SJ: San José Independencia; SMA: Santa Maria de la Asunción; SPIx: San Pedro Ixcatlán.

ADJ = adjective; agreement markers = agreement subject; CPL = completive aspect; ef = ending fusion; EXCL = exclusive; FORM = formative; IMPER = imperative; G&K = Golston & Kehrein (1998); /H/ = high tone; /h/ = mid-high tone; IC = inflectional class; INCL = inclusive; INCPL = incompletive aspect; /L/ = low tone; /M/ = mid tone; MP = morphophonological; MoV = motion verb; N = noun; NTR = neutral aspect; PV = preverb; PFM = Paradigm Function Morphology; P&P = Pike & Pike (1947); PL = plural; Subcfl = subconflation; SG = singular; T = tone; TAMV = tense, aspect, mood & voice; V = underspecified vowel; V'V, VhV = rearticulated vowels; |…| morphemic representation (underlying concatenative pattern); /…/ phonological representation (lexical level); […] phonetic representation (postlexical level).

Acknowledgements

This work is part of the program "Investissements d'Avenir" overseen by the French National Research Agency, ANR-10-LABX-0083 (Labex EFL), in particular PPC11 – Complexity, Geolinguistics & (Morpho)Phonological Patterns (strand 1) and Cross-Mediated elicitation, EM2 (strand 7). We also thank the *Institut Universitaire de France* for funding fieldwork in the period 2010-14, within the framework of the MAmP project (*cf.* http://jll.smallcodes.com/home.page). This research benefited extensively from the PFM formalization designed initially by Alain Kihm (see Léonard & Kihm 2010, 2012, 2014), and from unvaluable help by Julien Fulcrand in transcribing the data, recorded in 2012 and 2013 either by Jean Léo Léonard (JLL) or by Jaime R. Calderón (UNAM) and JLL during, or at the margin of Mazatec Grammar and Literacy Workshops.

References

Anderson, Stephen R. 2015. Dimensions of morphological complexity. *Understanding and measuring morphological complexity,* M. Baerman, D. Brown & G. G. Corbett (eds), 11-28. Oxford: Oxford University Press.

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* 6: 93-117.

Carrera, Rodrigo. 2014. Fonología del mazateco de San Lorenzo Cuaunecuititla, MA dissertation, Benemérita Universidad Autónoma de Puebla, available on academia.edu.

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

Golston, Chris & Kehrein, Wolfgang. 2004. A prosodic Theory of laryngeal contrasts. *Phonology* 21: 1-33

Gudschinsky, Sarah. 1958a. *Proto-Popotecan. A Comparative Study of Popolocan and Mixtecan*. *International Journal of American Linguistics* 25(2) supplement.

Gudschinsky, Sarah C. 1958b. Mazatec dialect history: a study in miniature. *Language* 34: 469-481.

Jakobson, Roman. 1957. *Shifters, verbal categories, and the Russian verb*. Cambridge: Russian Language Project, Department of Slavic Languages and Literatures, Harvard University.

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

Jamieson, Carole Ann. 1982. Conflated subsystems marking person and aspect in Chiquihuitlán Mazatec verb. *International Journal of American Linguistics* 48(2): 139-167.

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

Jamieson, Carole Ann. 1996. *Diccionario mazateco de Chuiquihuitlán, Oaxaca.* Mexico City: Instituto Lingüístico de Verano.

Hruschka, Daniel, Christiansen, Morten H., Blythe, Richard A., Croft, William, Heggarty, Paul, Mufwene, Salikoko S., Pierrehumbert Janet B. & Poplack, Shana. 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* 36.3: 205-211.

Kirk Paul L., Ladefoged, Jenny & Ladefoged Peter. 1993. Quantifying acoustic properties of modal, breathy and creaky vowels in Jalapa Mazatec. *American Indian Linguistics and Ethnography in Honor of Laurence C. Thompson*, A. Mattina & T. Montler (eds), 435-40. Missoula: University of Montana.

Léonard, Jean Léo. 2010. Enquêtes exploratoires pour l'*ALMaz* (*Atlas Lingüístico Mazateco*). Elicitation croisée, entre typologie et codification d'une langue Oto-Mangue. *Géolinguistique* 11: 59-109.

Léonard, Jean Léo. 2013. Le verbe mazatec : typologie et diasystème. *Amerindia* 37.1: 211-251.

Léonard, Jean Léo. 2014. *An xo'boo*, ou mazatec de Puebla – un fil d'Ariane écolinguistique dans le labyrinthe dialectal mazatec. *Les minorités invisibles: diversité et complexité (ethno)sociolinguistiques,* D. L. Ksenija (ed.), 290-304. Paris: Michel Houdiard Editeur.

Léonard, Jean Léo, dell'Aquila, Vittorio & Gaillard-Corvaglia, Antonella. 2012. The *ALMaz* (*Atlas Lingüístico Mazateco*): from geolinguistic data processing to typological traits. *STUF* 65(1): 78-94.

Léonard, Jean Léo, dell'Aquila, Vittorio & Gaillard-Corvaglia, Antonella. 2014. Emergence forte *versus* émergence faible en écologie linguistique : le cas du mazatec. Proceedings of the Symposium 2014 of the Institut Universitaire de France *Nature et Culture*. Saint-Etienne: Presses Universitaires de Saint-Etienne. <http://axe7.labex-efl.org/node/250>

Léonard, Jean Léo & Fulcrand, Julien. 2016. Tonal inflection and dialectal variation in Mazatec. *Tone and Inflection. New Facts and New Perspectives,* E. Palancar, Enrique & J.L. Léonard (eds), 163-198. Berlin: Mouton de Gruyter.

Léonard, Jean Léo & Kihm, Alain. 2010. Verb inflection in Chiquihuitlán Mazatec: a fragment and a PFM approach. *Proceedings of the HPSG10 Conference*, S. Müller (ed.), 288-306. Paris: CSLI Publications.

Léonard, Jean Léo & Kihm, Alain. 2012. Classes flexionnelles du mazatec et diasystème. Empirisme critique et formalization. *Bulletin de la Société de Linguistique de Paris* 56: 379-446.

Léonard, Jean Léo & Kihm, Alain. 2014. Mazatec verb inflection: A revisiting of Pike (1948) and a comparison of six dialects. *Patterns in Mesoamerican Morphology,* J.L. Léonard & A. Kihm (eds)*,* 26-76. Paris: Michel Houdiard Editeur.

McCarthy, John. 1986. OCP Effects: Gemination and Antigemination. *Linguistic Inquiry* 17: 207-263.

Miestamo, Matti, Sinnemaki, Kaius & Karlsson, Fred (eds). 2008. *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins.

Mufwene, Salikoko. 2001. *The Ecology of Language Evolution*, Cambridge Approaches to Language Contact. Cambridge: Cambridge University Press.

Mufwene, Salikoko. 2013. Complexity perspectives on language, communication, and society. *Complexity Perspectives on Language, Communication and Society,* A. Massip-Bonet & A. Bastardas-Boada (eds), 197-218. Belin/Heidelberg: Springer.

Myer, Lois & Alvarado, Benjamin Maldonado (eds). 2010. *New World of Indigenous Resistance. Noam Chomsky and Voices from North, South and Central America*. San Francisco: City Light Books.

O'Sullivan, David. 2004. Complexity Science and Human Geography. *Transactions of the Institute of British Geography* 29(3): 282-295.

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. *International Journal of American Linguistics* 22: 57-71.

Pike, Eunice. 1967. Huautla de Jiménez Mazatec. *Handbook of Middle American Indians,* 5 *Linguistics*, N. McQuown (ed.), 311-330. Austin: University of Texas Press.

Pike, Kenneth & Pike, Eunice. 1947. Immediate constituents of Mazatec Syllables. *International Journal of American Linguistics* 13:78-91.

Rensch, Calvin R. 1977. Classification of the Oto-Manguean languages and the Position of Tlapanec. *Two Studies in Middle American Comparative Linguistics,* C. R. Rensh & D. Oltrogge (eds), 53-102. Arlington: University of Texas/S.I.L.

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

Silverman, Daniel, Blankenship, Barbara, Kirk, Paul & Ladefoged, Peter. 1987. Phonetic Structures in Jalapa Mazatec. *Anthropological Linguistics* 37-1: 70-88.

Silverman, Daniel 1997. Laryngeal complexity in Oto-Manguean vowels. *Phonology* 14: 235-261.

Suárez, Jorge A. 1983. *The Mesoamerican Languages.* Cambridge: Cambridge University Press.

Schane, Standford A. 1985. Vowel Changes of Mazatec. *IJAL* 51(4): 62-78.

Schram, Judith L. & Pike, Eunice V. 1978. Vowel Fusion in Mazatec of Jalapa de Díaz. *International Journal of American Linguistics* 44(4): 257-261.

Stump, Gregory. 2001. *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

Trubetzkoy, Nikolaj S. 1939. *Grundzüge der Phonologie*. Prague: Travaux du Cercle Linguistique de Prague.

Weinreich, Uriel. 1954. Is a structural dialectology possible? *Word* 4: 388-400.

APPENDIX

Main collaborators for the ALMaz:

- Huautla: Clementina Elodia Cerqueda García. Bilingual school teacher, aged 53 in 2011. Always lived in Huautla. Data recorded by J. L. Léonard, in 2011-13.
- Jalapa de Díaz: Gilberto Martínez Álvarez. Bilingual school teacher, aged 53 in 2013. Trained at Universidad Pedagógica Nacional (UPN de Tuxtepec), author of a school book for literacy in Jalapa Mazatec. Lived a few years in Córdoba Veracruz. Recorded on 28, 29-07-2013 by Léonard.
- San José Independencia: Magneli Manuel Romero. Bilingual school teacher, aged 33 in 2013. Lived eight years in Tierra Blanca (Veracruz), Puebla and now in Tuxtepec, where she lives nowadays. Parents from the former village Nueva Patria, Soyaltpec, Tuxtepec, Oaxaca, drowned by the Miguel Alemán dam. Recorded by Calderón and Léonard the 24-08-2013 at Corral de Piedra, San Miguel Soyaltepec.
- San Lorenzo Cuaunecuiltitla: B.I.C. (Bachillerato Integral Comunitario) student, Abraham Cabrera Gabito, aged 16 in 2013. Always lived in SLC, except one year in San Antonio Eloxochitlán. Grandparents from Matzazongo de Guerrero, Puebla. Recorded on 20-08-013 by Calderón and Léonard.
- Santa Maria de la Asunción (Jiotes): shop keeper, aged 38 in 2013. Lived until 14 at SMS, 20 years in Mexico City and settled back in SMA three years before the recording. Recorded on 14-08-2013 by Calderón,
- Soyaltitla: Guillermina Borilla Torres, B.I.C. (Bachillerato Integral Comunitario) student, aged 15 in 2013. Father from Mazatlán, mother from Huautla. Always lived in Soyaltitla, except three months in Mexico and one week in San Lorenzo, where she had recently entered the local high school. Recorded on 20-08-013 by Calderón and Léonard.