**Appendix 1**

**Prefixes, classifiers and monosyllabic words**

**Note: Please do not cite without consulting author at jonamith@gmail.com. This is a draft and may contain errors including some that can easily be corrected.**

ã ẽ ĩ õ ũ

Note: Analysis of what constituted a clitic is ongoing and some forms presented in the second part of this appendix perhaps might be considered phrasal, not word, enclitics. Monomoraic monomorphemic words are for now divided into several classes, the definition and boundary of which will undoubtedly change.

Appendix 1.1 The pronominal enclitics are generally alternative monomoraic forms of independent dimoraic words. The enclitics are most commonly added to a verb (referencing an argument) and to nouns (referencing a possessor). Certain adverbial markers (e.g., the intensifier ni42) often occur between the verb stem and the enclitic. In this study enclitics are preceded by an equal sign, '=' (e.g., Ku3xi3=un3, ‘¡Eat!’)

**1. Derivational and inflectional prefixes**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ku3- | inchoative verbalizer | to wind up;  to become | *Ku3-* and *ndu3*- are alternative inchoative markers that occur mostly before adjectives and nouns, converting them into verbs (e.g., de *kwã 4ã*2 ‘yellow’ forms *ku3-kwã4ã2* ‘ponerse amarillo’ | see beow |
| ndu3- | inchoative verbalizer | to wind up;  to become |
| nda3-  ndi3- | iterative | to repeatedly [verbal action];  to once again [verbal action]; | Although nda3- and ndi3- seem basically to be iteratives at times the sense of the iterative + verb is quite distinct from that of the verb itself. In other cases although a verb seems to be derived from *nda3* + verbal root the “verbal root” has not been documented to exist without the nda3- or ndi3- "prefix". no existe solo. The context of alternation between *nda3*- and *ndi3*- has not been determined. |  |
| sa4-  ja4- | causative | to make, induce, provoke Sub to verb | The causative marker increases verbal valency adding an agent to the verbal event. The causative marker appears to derive from the verb *a3sa3* (in many Mixtec languages, *ka3sa3*). The conditions for alternation between *sa4-* and *ja4*- have not been determined. |  |
| ni1-  ni4- (negative alternant after ba143) | completive | —— | The tone of the completive is 4 after a negative marker. Negation is not manifested by a 14 ascendant tone on the first mora. |  |
| ka3 | ? | ? | There are various roots that occur with an initial *ka3* syllable (e.g., *ku3-nde3ta3* 'to be standing, to stand up' *ka3nde3ta3* 'to stand up vertically (long objects, plural)'. The nature of the *ku3-* and *ka3-* alternation needs to be researched. For now *ka3* has not been written as a prefix but rather part of a (derived) stem. |  |

Verbs that alternate ku3-root (potential) > root (habitual)

|  |  |  |
| --- | --- | --- |
| Potential | Habitual |  |
| ku3ndu'4u4 | i4ndu'3u4 | to be seated (sg) |
| ku3-so3ko2 | i4so3ko2 | to be hungry |
| ku3-nda3tu3 | nda4tu3 (progressive)  cf. ju4-nda3tu3 (habitual) | to be waiting |
| ku3-ndi3ta3, ku3-nde3ta3  kwe3ta3, kwi3ta3 | ndi4ta3, nde3ta3 (stative) | to be standing (pl) |
| ku3-ni'4i4 | ni'4i4 | to have in ones hand |
| ku3-ñu3u3 | ñu4u4 (progressive)  ju4-ñu3u3 | arrastrar |
| ku3-ñu'3u3 | ñu'4u4 (habitual NOTE: perhaps better interpreted as a stative) | to be contained in, inside of |

**2. Classifiers (with associated independent and enclitic forms)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Full** | **Classifier** | **Enclitic** | **Reference** | **Discussion** |
| ta1a3  *hombre* | ta1 | =ra[1,2,3] | masculine sg | Note that the enclitic is distinct from the classifier. |
| ña'3a4  *woman* | ña4 | = ã4;  =ẽn4 (after /a/)  =ña4 | feminine sg | Note the alternation although both classifier and enclitic contain a nasal vowel. The enclitic *=ña4* is the rarer of the forms and is used mainly by elders. |
| ———— | na1 | = na[1,2,3] | human plural |  |
| ña'1a3  *thing, object* | ña1 o ya1 | =a[1,2,3] | inanimate object | At times animals and humans may be referenced by the inanimate object enclitic, that is, the animates are treated as perhaps less animate. Instead of using the enclitics for humans and animals, the enclitic or classifier for objects is utilized =*a[1,2,3]*, *=e[1,2,3]*, *ña1*o *ya1*. |
| i3tun4  *árbol* | tũ4 | =rũ4 | wooden object | Note that the enclitic is distinct from the classifier. Used only for wood and objects made of wood with a few exceptions. Airplanes, cars, trucks, tops, tables and other things traditionally made of wood now no longer true for many. |
| ki3ti4  *animal* | ti4 | =ri4 | animal | Note that the enclitic is distinct from the classifier. Used only for animals although as noted above, animals may be referenced by the enclitic for inanimate object. |
| *ta1kwi42*  *agua*  (*nduTtaT* or a cognate word is used in other Mixtecan languages) | ta4 | =ra4 | water, liquids | Note that the enclitic is distinct from the classifier. Note that *ta4* (*=ra4*) is not used for all "liquids". For example, for atole and honey *ña1*o *ya1* (=a[1,2,3]) is used. |

**3. Monomoraic independent words**

|  |  |  |  |
| --- | --- | --- | --- |
| **Form** | **Category** | **Gloss** | **Brief summary** |
| ã3 | interjection | ah! | this interjection is equivalent to Spanish 'ah' with the difference that the Mixtec form is nasalized. It may or may not be a borrowing from the Spanish. |
| ã4  na4 | hortative modal | I hope that [verb in potential] occurs  I wish that [verb in completive] had occurred  let's verbo! [verb 1st inclusive] | * used with a verb marked with the 1st person inclusive the function word acts as a modal cohortative * used with verbs inflected for persons other than 1st person inclusive ã4/na4 marks a counseling or desire in regard to an action to either might have been realized (with the verb in completive) or that the speaker wishes would occur in the future (with the verb in the potential) |
| ã4  na4 | subordinator | if | * the function words *ã4* and *ndi4* surround a subordinate conditional clause. A search through the corpus (2013-11-04) revealed 30,027 lines with *ã4*/na4 followed by *ndi4*; 10,230 had the two-word sequence *ã4*/na4 *ndi4*. * when the verb is in the potential aspect *ã4*/na4 indicates a possibility that if realized will produce a future consequence (since it is not a counterfactional but an irrealis it does not use the *nu4u4* marker) |
| ã4  a4  na4 | interrogative | ¿ ã4 phrase ? | a clause introduced by either of the three function words in the first column is a question |
| ba143 | negative | no | The most common use of *ba143* is to negate habitual aspect verbs. Optionally *ba143* may be preceding by *ni3*. This occurs 187 times in the corpus (2013-11-03) or 6.33% of the occurrences of *ba143*.  *ba143* also negates adjectives predicated with the zero copula (e.g., *Ba143 che4e13 ta1a3 kan4* 'That man is not fat').  Note the following statistics  *Ba143* occurs 2,956 times in the corpus (2013-11-03) and of these   * 2300 times is followed by a word whose first tone is 4 (indicative of the habitual form and many adjectives) * 317 times it is followed by *ba'1a3* ('good') * 47 times it is followed by *kwã'1ã4* ('going') * 23 times it is followed by *ba1xi3* ('veniendo') |
| ba3 | evidencial modal | surely | This is one of the most common words in YM, occurring 12,144 times in the corpus (2013-11-03). It has the following meaninigs:   * in phrase-final position it signifies that there has been a change of opinion or it marks a statement expressing the opion that what occurred was to have been different * *ba3* neagive clause *bi3*| evidential statement that the speaker deduces the veracity of the proposition that follows though there is no direct evidence (*Ba3 tia'14an2 ku3xi3=un4 bi3* ‘It must be that you haven't eaten’ (said to someone who before it is time to eat said that s/he is hungry'). |
| bã4  ba42 or ba43  be4 | copula inflected for person | to be | The copula in habitual aspect is *ku4u4*, which is invariably followed by an enclitic. When the enclitic is vowel initial the copula+enclitic sequences sometimes reduce. Note the following forms:   * *ku4u4=an4* (fem.sg) => *bã4* * *ku4u4=a3* (phrase final) or *ku4u4=a2* (phrase medial) => *ba43* or *ba*42 * *ku4u4=e4* (1pl incl) => (sometimes) *be4*.   In all these cases as the change to the surface form is quite irregular the transcription convention is to write these irregular reduced forms. |
| bi2 | adverb | ¡ya! | adverb generally used after an imperative to request that the addressee hurry up in carrying out the requested action |
| bi4 | modal | ¡hasta! | modal by which the speaker indicates surprise, by which he or she indicates that it is hard to believe that what the clause expresses is true. The most common collocations are *nda4a2 bi4* (Sp. trans. as *¡hasta dónde!*) and *ka1 bi4* . |
| chi1  chi2 | modal : quotative | dicen | as a quotative chi1/chi2 indicates that la proposition expressed by the preceding clause is something that the speaker has heard but that s/he cannot directly attest to its validity |
| cho4;  cho41 | modal | could it be that? | *ã4* cláusula *cho4* | *cho4* is a clause-final modal that with the interrogative *ã4* at the beginning marks the question, asking whether something that the speaker thought probable effectively will occur. In a corpus search (2013-11-03) of 99 occurrences of *cho4* (plus 9 of cho41) 66 were at the end of a clause that began with *ã4*. |
| ja1  xa1 | adverb | ya | A very common adverb (7,213 occurrences in a corpus search 2013-11-03), *ja1/xa1* almost always occurs at the beginning of a verbal clause. |
| ja143  xa143 | adverb | then, and then, afterwards | This adverb is generally found preverbally, particularly before the habitual verbs *ka4chi2* and *ka'4an2* ('to say' and 'to speak'), before the completive ni1-ki3xa3a4 (and its alternate forms, 'began') and before verbs of movement in the progressive, *kwa'1an1* 'to be going' and *ba1xi3* 'to be coming '. It is also found in the collocation *i3kan4 ja143* 'and then'. |
| ja3 | modal | really? | this modal is associated with verbs and adjectives to inidicate that the speaker has changed his/her mind and requests confirmation of the addressee that something that had been in doubt is true. |
| ka1 | adverb | still | Ka1 is usually postverbal and most commonly occurs (296/2126 in the corpus 2013-11-04) after *i4yo2* or its negative *ko14o3* (existential copula). It also frequently occurs after *nda4a2* (575 times) |
| kã4 | adverb (apocopated form) | allá | This is a common apocopated form of the adverbial deictic *i3k*ã*4* (‘over there’). It is most commonly used after the classifiers/relativizers (*ta1*, *na1*, *ya1*, *ti4*, *tun4*; about half of the time) and sometimes after *tã3*. It is most frequently followed by *ndi4* (about 25% of occurrences). |
| kwa14 | negative | no | This form is used rarely, and almost always only before the few verbs whose potential aspect form begins with a 14 tone contour on the first mora. After *kwa14* the tone pattern of the initial mora on the verb may change. Thus *yu'14bi2* > *kwa14 yu'3bi2* (45/73 occurrences of *kwa14*) and xi14ko3 > kwa14 xi4ko3 (20/73 occurrences). *cha'14bi3* > *kwa14* *cha'4bi3*. The remaining occurrences of *kwa14* are before *xi14ta3* ('cantar') > *kwa14* *xi4ta3* ('jalar'), *kwa14* *ti14bi3* ('descomponerse' no change to *ti4bi3*), and *cha'14bi3* > *kwa14* *cha'4bi3*. |
| ma42 | interjección | ¡ ! | *ma42* is often phrase final and indicates speaker astonishment or surprise; in narratives it is a discourse marker that calls listener attention to the interesting aspect of what had been said. Although it occurs only 28 times in the corpus and never after *chi1* ('dicen'), Rey Castillo mentioned that it does commonly occur after *chi1*. |
| na4 | modal | see ã4 |  |
| na4 | subordinator | see ã4 |  |
| na4 | interrogative | see ã4 |  |
| nda1; ta1nda1 | adverb; preposition | including; up to | *Nda1* has an alternative form, *ta1nda1*. Together the two occur (2013-11-04) 2,558 times in the corpus, most commonly followed by *nda4a2* (194 times) and *mi4i4* (138 times), both interrogatives meaning 'where?'. The terms also occur as part of indefinite pronouns such as *nda1 nda4a2 ka1* 'wherever' or 'whichever [noun] '. *Nd1 nda4a2 ka1* is found 173 times in the corpus.  *Nda1/ta1nda1* also frequently occurs as a "preposition" when followed by a term indicating a spatial or temporal distance. |
| nda143; ta1nda143 | adverb | now just, now only | *Nda143* has an alternative form, *ta1nda143*. Together the two occur (2013-11-04) 1,233 times in the corpus, often followed by a number or quantifier. The words translate loosely as 'now just' or 'now only' ('ya nomás' or 'ya solamente'). |
| ndi4 | subordinator | that (as a complementizer);  because (as an adverbial subordinator) | This is the most common word in Yoloxóchitl Mixtec occurring 44,554 times in the corpus (2013-11-04). Approximately 25% of the time it is followed by a classifier, particularly *ya1* and *ta1*. |
| ni3 | negative element that optionally occurs before verbs in the negative | optional before verbs in negative potential aspect  optional before negative *ba143*used to negate verbs in the habitual aspect  ni3 ... tan3 ni3 ... ; | (Ni3) ku14xi3=ra2 chi1.  'He is not going to eat, they say.'  (Ni3) ba143 xi4xi3=ra2 chi1.  'He doesn't eat, they say.'  The sequence ni3 ... tan3 ni3 ... is a negative noncontrastive correlative conjunction. The *tan3* is obligatory in verbal clauses but optional between nouns and pronouns:  Ni3 yo’4o1 [tan3] ni3 yu1’u1  Neither you nor I (lit., not you, not me)  Ni3 ku3ñu2 ndu3xi4 [tan3] ni3 ku3ñu2 ji1ndi1ki4 ku14xi(3)=2  Neither chicken meat nor beef will I eat.  Ni3 ku14xi(3)=2 tan3 ni3 ko'14o4=2  I am not going to eat nor drink. |
| ni4 | modal | phrase final in modal expressions | *Ni4* is most often used at the end of a cohortative, imperative, or admonitive phrase indicating the urgency with which the speaker speeks. For example, *ã4*  *ku3xi3=o4* can be understood as a strong suggestion, 'Let's eat!', whereas ¿*Ã4 ku3xi3=o4, ni4*? is more of a suggestion: 'Let's eat, no?' or 'It would be good to eat, wouldn't it?'. The imperative *¡Ku3xi3=un4!* can be mitigate, becoming more of a suggestion with the same *ni4*: *Ku3xi3=un4*, *¿ni4?* 'Eat, shouldn't you?' *Ni4* is found 17 times in the corpus [2013-11-07]. |
| nu4 | aspectual | in phrase-final position it marks a counterfactual or a suggestion that something not carried out or that hadn't occurred should have been done | For example, with the modal particle it suggests regret or reproach for something not carried out: *Ã4 ni4-xi3xi3=o4 nu4* 'We should have eaten (e.g., the food was ready at some time in the past but we didn't eat it then and now we wish we had). With a verb in the potential *nu4* indicate an action that did not occur: *Ta1 ku3xi3 ku4u4=2 nu4*, 'He was going to eat (but for some reason didn't )'. The word *nu4* has been found only 10 times in the corpus [2013-11-07]. |
| pi42 | address term | Used almost exclusively among men to establish or indicate a level of familiarity among the participants in a speech act. | This element occurs generally in conversations. Only 9 have been found in the corpus [2013-11-07] |
| rã3 | modal | Used in phrase final position in a conversation | The meaning of this word has still not been determined. |
| rã4 | adverb | Used in phrase final position. | This can be translated into Spanish 'pues' although the precise function and meaning of phrase-final *rã4* still needs to be determined. Apparently it is used after exhortative phrases. Of the 1774 tokens found in the corpus [2013-12-13], 550 were end of line, 285 were followed by a period, 758 were followed by a comma, 179 were followed by a space, and two were followed by a question mark. |
| rũ 42 | modal | Clause final evidencial indicating that the speaker is certain of what he or she has said, apparently because of having seen it. | A-- ¿ *Ã4* ni1-xa'1an1 Juan bi3ko4?  B-- Ni14-xa'1 ã1-ra1, run42.  A-- ¿Did Juan go to the fiesta?  B-- He didn't go, I know (because I saw he didn't).  As of 13 Dec. 2013 *rũ 42* was not found in the transcribed corpus. |
| su3 | modal | The modal *su3* indicates that the speaker considers what he or she has said to be something of awe, to be admired, something that the addressees should consider admirable or surprising. |  |
|  |  |  |  |
| tã3 | conjunction | Generally *tã3* is used in the coordination of two clauses.  It is also used in phrase final position to indicate that what follows is a continuation of the same thought. | For the two coordinating conjunctions "and" and "but" *tã3* is used indiscriminately, context determining the nature of the coordination, if necessary.  *Tio1to1 ku4ni2=yu1 tã3 xu'14ũ4 ku4ni2 Juan.*  I want clothes and/but John wants money.  Cf. to the following conjunctions:    Compare *tã3* to the negative element *ndi4*, which is used and cannot be eliminated in phrases such as the following:  *Ku14ni2=yu(1)=a1 kwa'4a1 ndi4 ya1 ya3a3 ku4ni2=yu1*  I don't want the red one but (rather) the white one.  Note the use of *ji'4 ĩ4* withthe meaning of 'and' between nouns and nominal phrases:  *Tio1to1 ji'4 ĩ4 xu'14ũ4 ku4ni2=yu1.*  'I want clothes and money'  *Ya1 kwa'4a1 ji'4ĩ(4)=a2 ya3a3 ku4ni2=yu1*  I want the red one and the white.  In addition to its use as a conjunction, *tã3* is used in comparative constructions:  *Su4kun1 ka1 ba3 yo'4o1 tã3 yu'1u1*  You are taller than me. |
| tã42  tiã42 | adverbial subordinator | when;  as is | Tã42 xa'1an1=ndu1 ji1ni4 yu3ku4 ni1-ku3un3 sa1bi4.  When we went to the House of Rain (lit., 'head of the hill') it rained.  Tan42 ka4a4 nu14u3 yu3ba4=ra2 ka4a4 nu14u3=ra3.  His face is like that of his father (lit., 'as.is appearance face his.father, appearance his.face') |
| ti4 | adverb | solely  only,  just (in Spanish solamente, nomás, apenas) | The element *ti1* occurs with *nda143* in the sequence *nda143* ... *ti4* to indicate 'now just' ('ya nomás') or 'now only' ('ya solamente'). |
| tĩ3 | adverb (apocopated form) | now | Apocopated form of the adverb *bi3tĩ3*. This short form nevers occurs, as does *bi3tĩ 3*, in phrase-initial position. |
| tõ4 | verb (apocopated form) | ¡look! | *Tõ4* is a shortened form of *ko3to3= õ4*, the verb 'look' in the singular imperative. |
| tu3 | adverb | also (también) | For example:  *A1su3 ɲ3u2 kan4 kwa'1ã1 tu3 yu'1u1ba3*  'To this very village (there) I am also going.' |
| tu3 | modal | ¡it couldn't be! | Epistemological modal of variable placement that indicates surprise by these speaker about a situation, or the difficulty that the speaker has in believing the situation.  *¿A4chu3 ku'1ũ1 tu3 yo'4o1 Sa3lu4bi2 tu3?*  'And you? why would you (why are you) going to San Luis? (perplexed because the person in question never goes to this town)  Apparently, however, the use of *tu3* adds little to a phrase when it is used after an interrogative in isolation:  *¿Nda4a2 tu3?* '  ¿Dónde?' (used when there is an antecedent in the conversation that contextualizes such an inquiry)/ |
| ũ4 | subordinator | then; as a consequence | At the end of a clause it indicates that the following clause is subordinated and occurs as a result of the previous.  *Ã4 ku3xi(3)= ũ4 ũ4, ku3xi3 tu3 yu'1u1 ba3*.  'If you eat, then I will eat as well. |