

WORDHOOD AND ALIGNMENT IN DURAZNOS MIXTEC PRONOUNS

Sandra Auderset & Adam J.R. Tallman

The Word, Alignment, Valency:
Workshop in Honour of Fernando Zúñiga
Universität Bern
02.12.2023

1) *ichìntxeé tátá se'-ó mí-ó*

i ³	tʃi ¹ⁿ dɕe ³ e ⁵	ta ⁵ ta ⁵	se ³ ʔ[e ³]	o ⁵	mi ⁵ [i ⁵]	o ⁵
NEG.POT	help	father	child	1PL.INCL	TOP	1PL.INCL

'the father of our children would not help us' [0011; 02:41]

2) *uchìntxeé-na yì'ì*

u ³	tʃi ¹ⁿ dɕe ³ e ⁵	na ³	ʒi ¹ ʔi ¹
NEG.POT	help	3PL.HUM	1SG

'they (pl) will not help me' [0020; 17:00]

* tʃi ¹ⁿ dɕe ³ e ⁵	ʒi ¹ ʔi ¹	na ³	'I help them'
* tʃi ¹ⁿ dɕe ³ e ⁵	na ³	i ¹	'they help me'
* tʃi ¹ⁿ dɕe ³ e ⁵	o ¹		'help us/we help'
* se ³ ʔ[e ³]	no ⁵ ʔo ⁵		'our child'
BUT			
se ³ ʔe ³	na ³		'their child'
tʃi ¹ⁿ dɕe ³ e ⁵	na ³	na ³	'they help them'

3) *kù'-ìn chìntxeé-ì-na káa*

ku ¹ ʔ[ũ ¹]	i ¹	tʃi ¹ⁿ dɕe ³ e ⁵	i ¹	na ³	ka ⁵ a ³
POT.go	1SG	POT.help	1SG	3PL.HUM	DEM.DIST

'I will go help them there' [0082; 03:38]

4) *tátò-ò ná chìntxeé nto'ó nùù xíkanu va-ntó*

ta ⁵ t[a ¹]	o ¹	na ⁵	tʃi ¹ⁿ dɕe ³ e ⁵	no ⁵ ʔo ⁵	nu ¹ u ¹	ɕi ⁵ ka ³ nu ³	va ³	no ⁵
father	1PL.INCL	MOD	POT.help	2PL	FACE	INCOMPL.go	ADD	2PL

'our Father be with you (pl) wherever you go' [0088; 32:52]

1) *ichìntxeé tátá se'-ó mí-ó*

i ³	tʃi ¹ⁿ dɕe ³ e ⁵	ta ⁵ ta ⁵	se ³ ʔ[e ³]	o ⁵	mi ⁵ [i ⁵]	o ⁵
NEG.POT	help	father	child	1PL.INCL	TOP	1PL.INCL

'the father of our children would not help us' [0011; 02:41]

2) *uchìntxeé-na yì'ì*

u ³	tʃi ¹ⁿ dɕe ³ e ⁵	na ³	ʒi ¹ ʔi ¹
NEG.POT	help	3PL.HUM	1SG

'they (pl) will not help me' [0020; 17:00]

* tʃi ¹ⁿ dɕe ³ e ⁵	ʒi ¹ ʔi ¹	na ³	'I help them'
* tʃi ¹ⁿ dɕe ³ e ⁵	na ³	i ¹	'they help me'
* tʃi ¹ⁿ dɕe ³ e ⁵	o ¹		'help us/we help'
* se ³ ʔ[e ³]	no ⁵ ʔo ⁵		'our child'
BUT			
se ³ ʔe ³	na ³		'their child'
tʃi ¹ⁿ dɕe ³ e ⁵	na ³	na ³	'they help them'

3) *kù'-ìn chìntxeé-ì-na káa*

ku ¹ ʔ[ũ ¹]	i ¹	tʃi ¹ⁿ dɕe ³ e ⁵	i ¹	na ³	ka ⁵ a ³
POT.go	1SG	POT.help	1SG	3PL.HUM	DEM.DIST

'I will go help them there' [0082; 03:38]

4) *tátò-ò ná chìntxeé nto'ó nùù xíkanu va-ntó*

ta ⁵ t[a ¹]	o ¹	na ⁵	tʃi ¹ⁿ dɕe ³ e ⁵	no ⁵ ʔo ⁵	nu ¹ u ¹	ɕi ⁵ ka ³ nu ³	va ³	no ⁵
father	1PL.INCL	MOD	POT.help	2PL	FACE	INCOMPL.go	ADD	2PL

'our Father be with you (pl) wherever you go' [0088; 32:52]

“Thus a class of pronouns as such can by no means be set up as basic to the language.” (Pike 1944:131)

Pike (1944, 1945)’s view on Mixtec:

- morphology-syntax distinction is not informative in Mixtec – constituency exists but there is no discrete shift in structure from morphology to syntax
- pronouns are not suffixes because
 - they are cognate with free forms
 - the dependent forms are replaced with independent ones in certain grammatical contexts
 - treatment as suffixes would lead to ‘words’ that contain free forms
- pronouns are a type of noun with obligatory short forms in certain positions

“an adequate synchronic account of Mixtec must include affixes, clitics (of various kinds), and free words” (Macaulay 1987: 134)

Macaulay (1987, 1993)’s view on Mixtec:

- there is a morphology-syntax distinction in Mixtec
- clitics are motivated classes (fast speech clitics vs. phrasal affixes)
- dependent pronouns are phrasal affixes
 - can’t be fast speech clitics because pf ‘suppletive’ bound forms
 - restrictions on object position in local participants
 - can be separated from the verb (phrasally attached)
- independent pronouns are words (presumably...)

‘CLITICS’ ARE NOT USEFUL FOR DESCRIPTION OR TYPOLOGY

CURRENT MULTIVARIATE TYPOLOGIES CAN BE STRAIGHTFORWARDLY INTEGRATED WITH ALIGNMENT TYPOLOGIES

PIKE’S VIEW ACCOUNTS BETTER FOR MIXTEC FACTS THAN MACAULAY’S

1. Pronouns in Duraznos Mixtec
2. Previous analyses and debates on wordhood/clitic status; clitics from a typological perspective
3. Methodology
 - multivariate approach to wordhood and ‘clitics’,
 - extension to alignment
 - data and coding
4. Results
 - Clustering: are there clitic classes? do pronouns cluster together?
 - Correlations: do variables align with each other?
 - Random Forest: which variables are important for classification?
5. Conclusion

SOME BACKGROUND ON DURAZNOS MIXTEC

- lexical elements are minimally bimoraic, grammatical elements can be monomoraic
- lexical and grammatical tone:
 - 3 tonemes (low, mid, high)
 - TBU is the mora; each mora is specified for tone
- basic word order: VSO
 - but NPs can be fronted for emphasis/focus
- verbs are obligatorily marked for aspect-mood
 - 3 forms: POT, INCOMPL, COMPL
- nouns have no obligatory marking at all

ta sààn ìsùvá'-ì ìsísínì-va-ì
ta⁵ sã⁵ã¹ i¹-su¹βa⁵?-i¹ i¹-si⁵si⁵ni¹-βa³-i¹
LNK then COMPL-prepare-1SG COMPL-have.breakfast-ADD-1SG
'And then I prepared breakfast.'

SOME BACKGROUND ON DURAZNOS MIXTEC

- lexical elements are minimally bimoraic, grammatical elements can be monomoraic
- lexical and grammatical tone:
 - 3 tonemes (low, mid, high)
 - TBU is the mora; each mora is specified for tone
- basic word order: VSO
 - but NPs can be fronted for emphasis/focus
- verbs are obligatorily marked for aspect-mood
 - 3 forms: POT, INCOMPL, COMPL
- nouns have no obligatory marking at all

ta sààn ìsùvá'-ì ìsísínì-va-ì
ta⁵ sã⁵ã¹ i¹-su¹βa⁵?-i¹ i¹-si⁵si⁵ni¹-βa³-i¹
LNK then COMPL-prepare-1SG COMPL-have.breakfast-ADD-1SG
'And then I prepared breakfast.'

1ST AND 2ND PERSON PRONOUNS

GLOSS	independent	dependent	
1SG	yǐ'ì [ʒi¹ʔi¹]	ì [i¹]	~ ì ~ [i⁵¹]
1PL.INCL	(míí-ó) [mi⁵i⁵o⁵]	ò [o¹]	~ ó ~ ò ~ [o⁵] ~ [o⁵¹]
1PL.EXCL	ntì'ì [ⁿdi¹ʔi¹]	ntì [ⁿdi¹]	~ nti ~ ntí ~ [ⁿdi³] ~ [ⁿdi⁵]

GLOSS	independent	dependent	
2SG.NHON	yó'ó [ʒo⁵ʔo⁵]	un [ũ³]	~ ùn ~ ún ~ n ~ [ũ¹] ~ [ũ⁵] ~ [̃]
2SG.HON	(míí-ní) [mi⁵i⁵ni⁵]	ní [ni⁵]	
2PL	ntó'ó [ⁿdo⁵ʔo⁵]	ntò [ⁿdo¹]	~ ntó ~ [ⁿdo⁵]

3RD PERSON PRONOUNS

	dependent		classifier	related noun
3SG.M	rà [ra ¹]	~ ra ~ rá ~ [ra ³] ~ [ra ⁵]	txà ~ rà [tɕa ¹] ~ [ra ¹]	txàa 'man' [tɕa ¹ a ³]
3SG.F	ñà [ɲa ¹]	~ ñá ~ àn ~ án ~ [ɲa ⁵] ~ [ã ¹] ~ [ã ⁵]	ñá [ɲa ⁵]	ñà'a 'woman' [ɲa ³ ʔa ¹]
3PL.HUM	nà [na ¹]	~ na ~ ná ~ [na ³] ~ [na ⁵]	na [na ³]	--
3GEN	ñà [ɲa ¹]	~ ñá ~ àn ~ án ~ [ɲa ⁵] ~ [ã ¹] ~ [ã ⁵]	ñà [ɲa ¹]	ñà'a 'thing' [ɲa ¹ ʔa ³]
3ZOO	rì [ri ¹]	~ rí ~ [ri ⁵]	txì [tɕi ¹]	kìtxì 'animal' [ki ¹ tɕi ¹]
3WOOD	dùn [ðũ ¹]	~ dún ~ [ðũ ⁵]	tú ~ tù [tu ¹] ~ [tu ⁵]	ìtun 'tree' [i ¹ tũ ³]

FUNCTIONS OF PRONOUNS

1. + VERB: argument
2. + NOUN: possessor
3. + relator noun/preposition: oblique
4. fronted element ('focus')

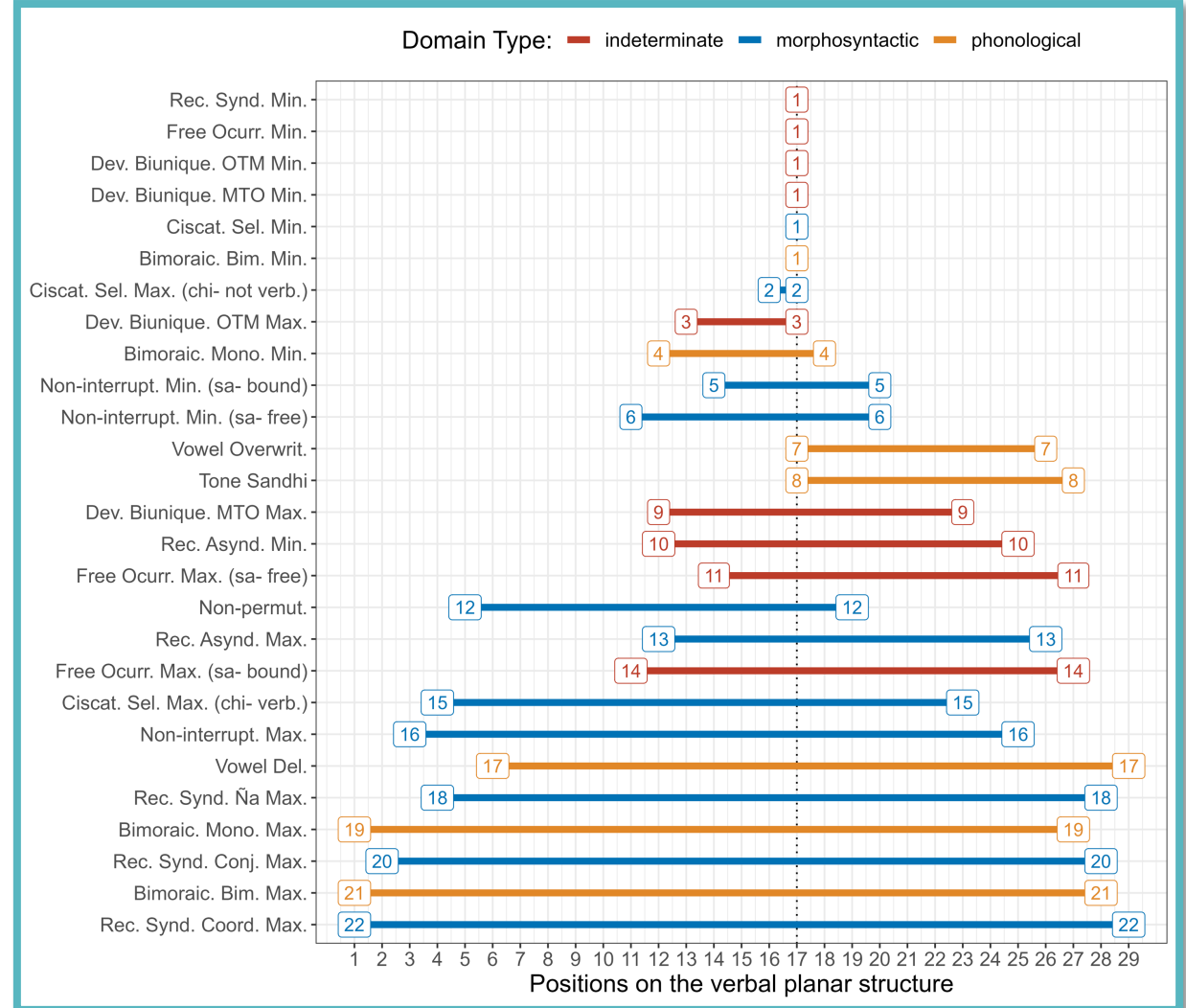
chi nì txiin-rí-ri míí sùkùn lo'o-rí
because MOD COMPL.grab-3ZOO-3ZOO TOP neck small.SG-3ZOO
because it [dog 1] grabbed him [dog 2] by his little neck [0007; 00:32]

ntì'ì ta ntákani-va ini-ntì jí'-an ví
1PL.EXCL LNK INCOMPL.worry-ADD inside-1PL.EXCL with-3GEN DM
a nosotros nos sorprendía pues [0088; 00:12]

ta jaan kontòo ikán kùsì-rà jí'in-rì yikì-va
LNK then INCOMPL.stay DEM.DIST POT.sleep-3SG.M with-3ZOO mountain-ADD
'and there he stays to sleep among them [sheep] in the mountain' [0057; 03:47]

IS MIXTEC ISOLATING OR POLYSYNTHETIC?

Pos.	Type	Elements	Forms
(1)	slot	connectors, question marker	<i>an, ta, távâ, chii, etc.</i>
(2)	slot	question words	<i>nishi, ntxáa etc.</i>
(3)	slot	focus (S/A/P, OBL, etc.)	
(4)	slot	realis negation	<i>kòó</i>
(5)	zone	adverbials	<i>xàà, sa'a, và'a, vitxi etc.</i>
(6)	slot	additive	<i>va</i>
(7)	slot	intensifier	<i>kuà'à, tóntó</i>
(8)	slot	intensifier/again	<i>ntxivà'a, yáá; tiki~ti</i>
(9)	slot	intensifier/again	<i>tiki~ti; ntxivà'a, yáá</i>
(10)	slot	additive	<i>va</i>
(11)	slot	modals	<i>nì, ná</i>
(12)	slot	perfective, irrealis negation	<i>ì; u/o~i</i>
(13)	slot	causative 'do'	<i>sá</i>
(14)	slot	IRR/PFV class markers	<i>ku; xì</i>
(15)	slot	iterative	<i>nta~nti</i>
(16)	slot	transitivizer 'put'	<i>chi</i>
(17)	slot	verb core	
(18)	slot	additive	<i>va</i>
(19)	slot	reciprocal	<i>ta'an</i>
(20)	slot	temporal adv.	<i>tuun, kîi</i>
(21)	slot	intensifier	<i>tóntó, kuà'à</i>
(22)	slot	intensifier/again	<i>ntxivà'a, yáá; tiki~ti</i>
(23)	slot	intensifier/again	<i>tiki~ti; ntxivà'a, yáá</i>
(24)	slot	additive	<i>va</i>
(25)	slot	inside/being	<i>ini</i>
(26)	slot	S/A	
(27)	slot	P	
(28)	zone	OBL, PP, LOC, adv.	
(29)	slot	discourse markers	<i>ní, ví</i>



CLITICS FROM A TYPOLOGICAL PERSPECTIVE

clitics are phonologically bound but morphosyntactically independent
(Spencer & Luís 2012, Aikhenvald 2002)

What does phonologically bound mean?

- not able to be a complete utterance
- integrated into another prosodic word (not its own prosodic word)
- integrated into a phonological phrase (weak layering)
- integrated into a phonological phrase (+prosodic word)
- undergoing lexical or post-lexical phonological processes? (Tallman *in press*)

What does morphosyntactically independent mean?

- depends on the wordhood test...

DEPENDENT PRONOUNS = PHONOLOGICAL CLITICS?

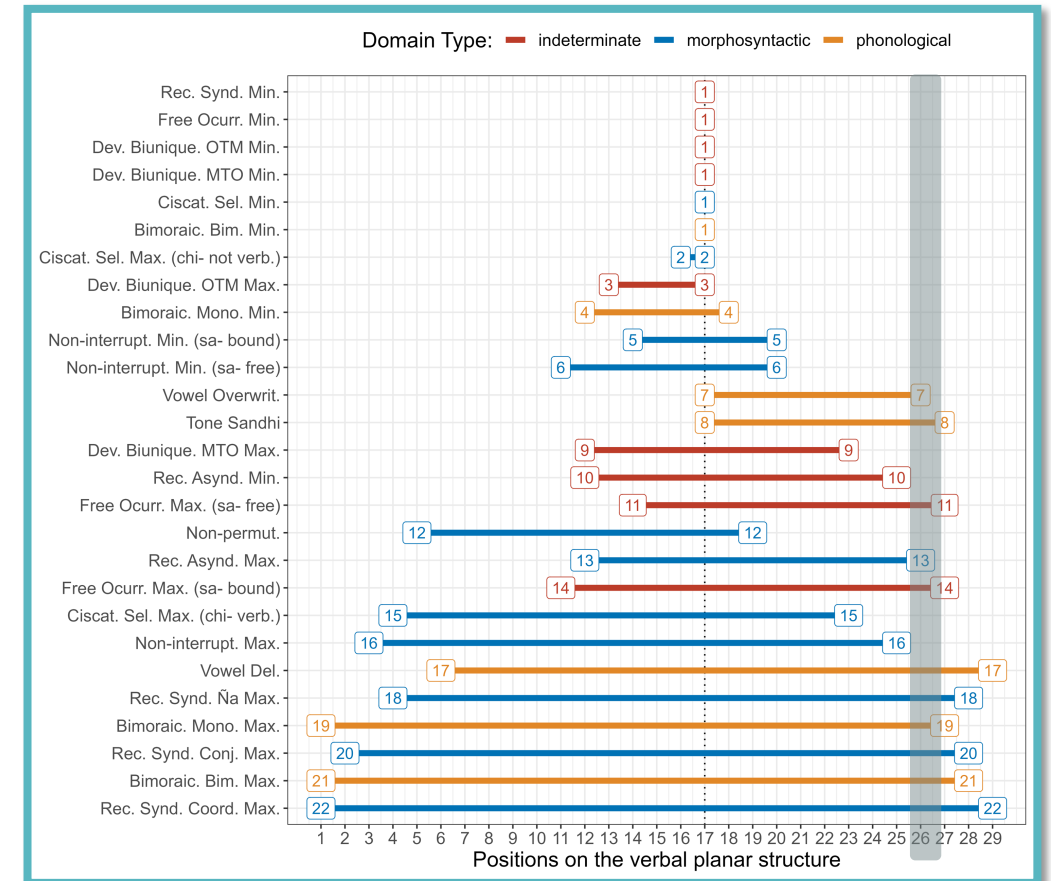
Zwicky 1977, Anderson 2005: 'phonological clitics'

- morphosyntactically independent:

- non-permutability
- non-interruptability – minimal
- ciscategorical selection - minimal

- phonologically dependent:

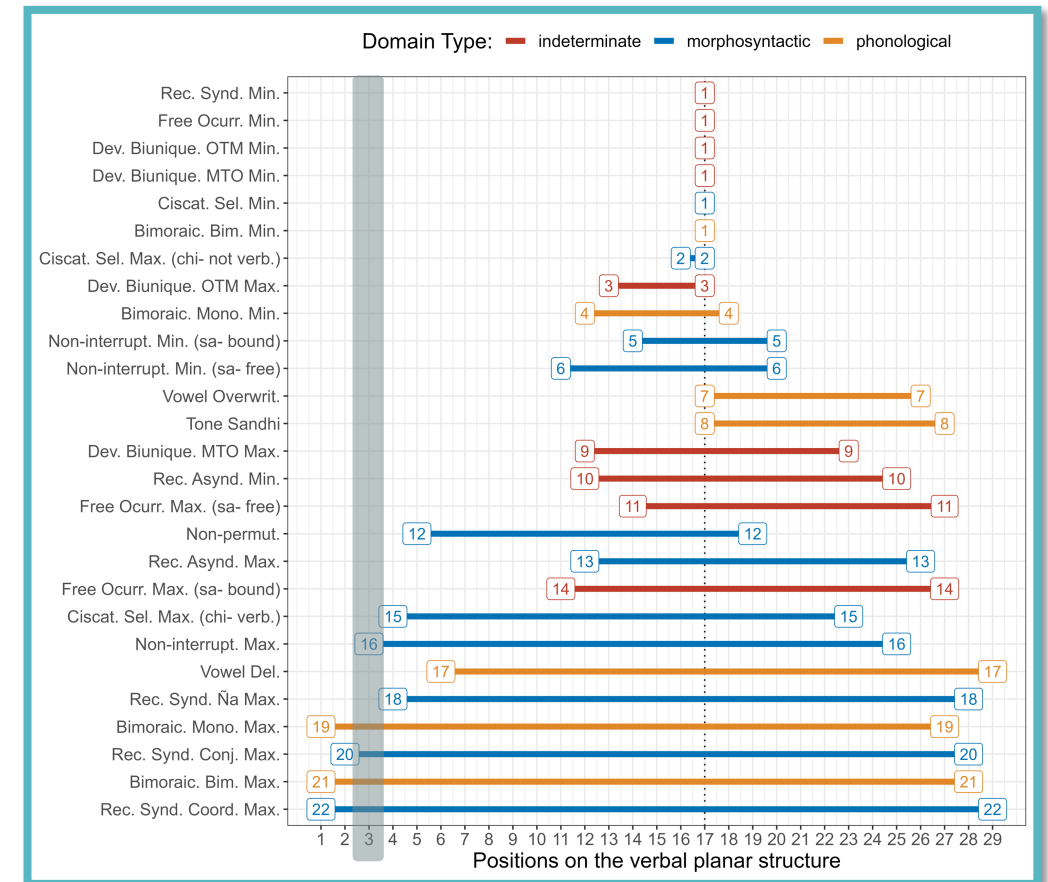
- bimoraicity constraint - maximal
- vowel deletion
- vowel overwriting



INDEPENDENT PRONOUNS = ANTICLITICS?

Zúñiga 2014: 'anticlitics'

- phonological words that are part of grammatical words
- formal independence
 - bimoraicity – minimal
 - vowel overwriting
 - tone sandhi
 - vowel deletion
- distributional dependence
 - recursion in coordination – maximal
 - non-interruptability – maximal



DEPENDENT PRONOUNS = PHRASAL AFFIXES?

Macaulay 1987: fast speech clitics vs. phrasal affixes vs. words

- dependent pronouns are phrasal affixes
 - “elements which are essentially affixal in nature but which can be shown to attach to the phrase rather than to single lexical items” (p.128)
 - can’t be fast speech clitics because of ‘suppletive forms’
 - restrictions on object position in local participants
 - can be separated from the verb
- independent pronouns are words/fast-speech clitics

ARE THERE CLITIC PRONOUNS IN MIXTEC?

Pronouns	Analysis	Phon. Criteria	Morphosynt. Criteria
dependent =	phonological clitics	bimoraicity constraint - maximal vowel deletion vowel overwriting	non-permutability non-interruptability – minimal ciscategorical selection - minimal
dependent =	phrasal affixes	deviations from biuniqueness	interruptability from the verb
independent =	anticlitics	bimoraicity – minimal vowel overwriting tone sandhi vowel deletion	recursion in coordination – maximal non-interruptability – maximal
independent =	words/fast speech clitics	bimoraicity	free occurrence?

ALIGNMENT OF PRONOUNS AND NOUN PHRASES

postverbal (positions 26 and 27):

FORM:

1/2 S=A≠P

3 S=A=P

NP S=A=P

ORDER:

all S=A≠P

preverbal 'focus' (position 3):

FORM:

1/2 S=A=P

NP S=A=P

ORDER:

all S=A=P

RESEARCH QUESTIONS

Do closed class elements in Duraznos Mixtec fall into two or three distinct classes when not cherry-picking wordhood tests?

Do dependent/independent pronouns form a coherent class according to these criteria and their alignment?

Is it helpful for description and/or comparison to classify dependent pronouns as (some type of) clitic?

DATA AND CODING

- language: San Martín Duraznos Mixtec
- data:
 - 13h transcribed & translated corpus of naturalistic speech
 - complemented with elicitation and native speaker judgements
- coding:
 - 108 closed class elements (pronouns, conjunctions, aspect/mood markers, intensifiers, classifiers, temporal expressions etc.)
 - coded for an array of diagnostics seen as crucial for distinguishing affixes/clitics/words (mostly based on Spencer & Luís 2012)

WORDHOOD VARIABLES

Variable	Description	Values
Selection	Which categories can X modify?	N and/or V and/or A
Permutability	Is X permutable with the base? ... with distributionally identical elements?	yes/no/yes with a difference in meaning
Interruptability	Can X be separated from the base?	no/bound/free
Free	Can X stand alone as a complete utterance?	yes/in a related form/no
Scope	Does X have wide scope under coordination?	yes/no
Complexity	Is X composed of morphs?	yes/no
Fossilization	Does X combine with non-recurring bases?	yes/no
Morae	How many morae does X have?	1,2,3
Allomorphy	Does X have segmental allomorphs? ... tonal allomorphs?	

ALIGNMENT VARIABLES FOR PRONOUNS

Variable	Description	Values
postverbal SA	Can X appear in position 26 as an S/A argument?	yes/no
postverbal P	Can X appear in position 27 as a P argument?	yes/no
preverbal	Can X appear in position 3?	yes/no

Pos.	Type	Elements	Forms
(1)	slot	connectors, question marker	<i>an, ta, távâ, chii</i> , etc.
(2)	slot	question words	<i>nishi, ntxáa</i> etc.
(3)	slot	focus (S/A/P, OBL, etc.)	
(4)	slot	realis negation	<i>kòó</i>
(5)	zone	adverbials	<i>xàà, sa'a, và'a, vitxi</i> etc.
(6)	slot	additive	<i>va</i>
(7)	slot	intensifier	<i>kuà'à, tóntó</i>
(8)	slot	intensifier/again	<i>ntxivà'a, yáá; tiki~ti</i>
(9)	slot	intensifier/again	<i>tiki~ti; ntxivà'a, yáá</i>
(10)	slot	additive	<i>va</i>
(11)	slot	modals	<i>nì, ná</i>
(12)	slot	perfective, irrealis negation	<i>ì; u/o~i</i>
(13)	slot	causative 'do'	<i>sá</i>
(14)	slot	IRR/PFV class markers	<i>ku; xì</i>
(15)	slot	iterative	<i>nta~nti</i>
(16)	slot	transitivizer 'put'	<i>chi</i>
(17)	slot	verb core	
(18)	slot	additive	<i>va</i>
(19)	slot	reciprocal	<i>ta'an</i>
(20)	slot	temporal adv.	<i>tuun, kûi</i>
(21)	slot	intensifier	<i>tóntó, kuà'à</i>
(22)	slot	intensifier/again	<i>ntxivà'a, yáá; tiki~ti</i>
(23)	slot	intensifier/again	<i>tiki~ti; ntxivà'a, yáá</i>
(24)	slot	additive	<i>va</i>
(25)	slot	inside/being	<i>ini</i>
(26)	slot	S/A	
(27)	slot	P	
(28)	zone	OBL, PP, LOC, adv.	
(29)	slot	discourse markers	<i>ní, ví</i>

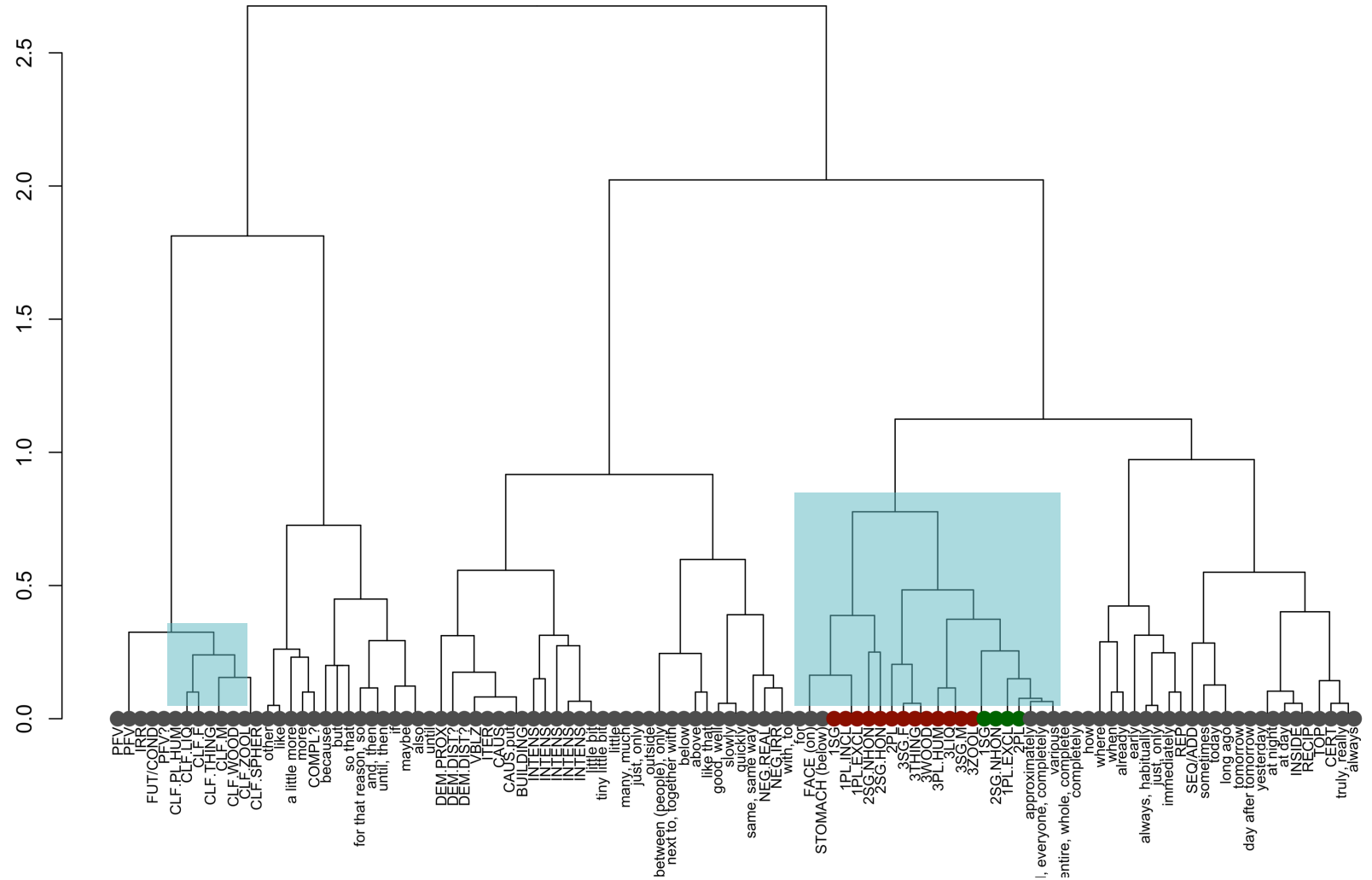
CODING EXAMPLE

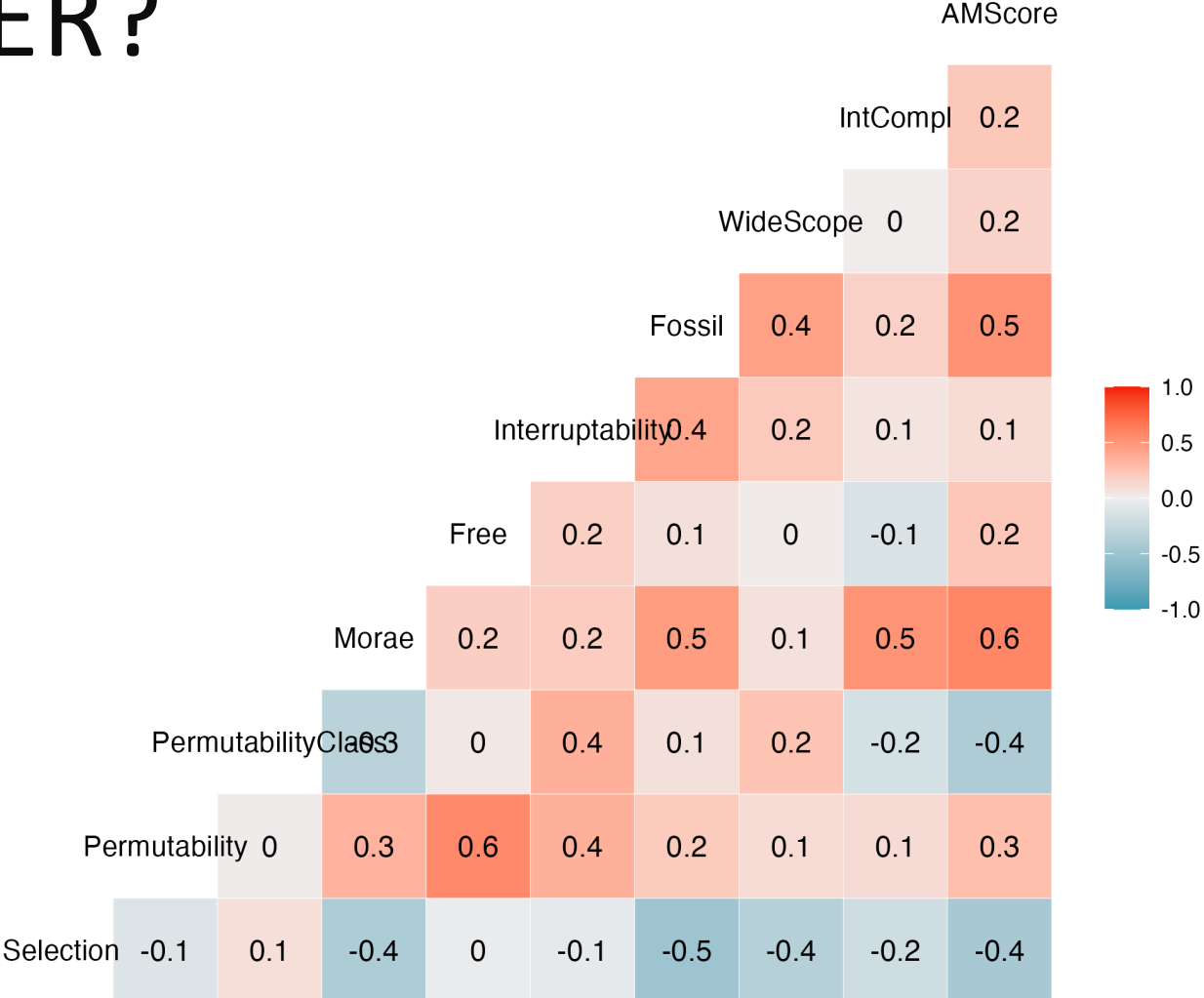
Variable	ì 1SG	yì'ì 1SG
Selection	Nouns, Verbs	Verbs
Permutability Head	no	yes
Permutability Class	yes	no
Interruptability	yes	yes
Free	no	yes
Scope	yes	yes
Complexity	no	no
Fossilization	no	no
Morae	1	2
Allomorphy Seg.	no	no
Allomorphy Tone	yes	no
postverbal SA	yes	no
postverbal P	no	yes
preverbal	no	yes

GLOBAL PARTITIONS?

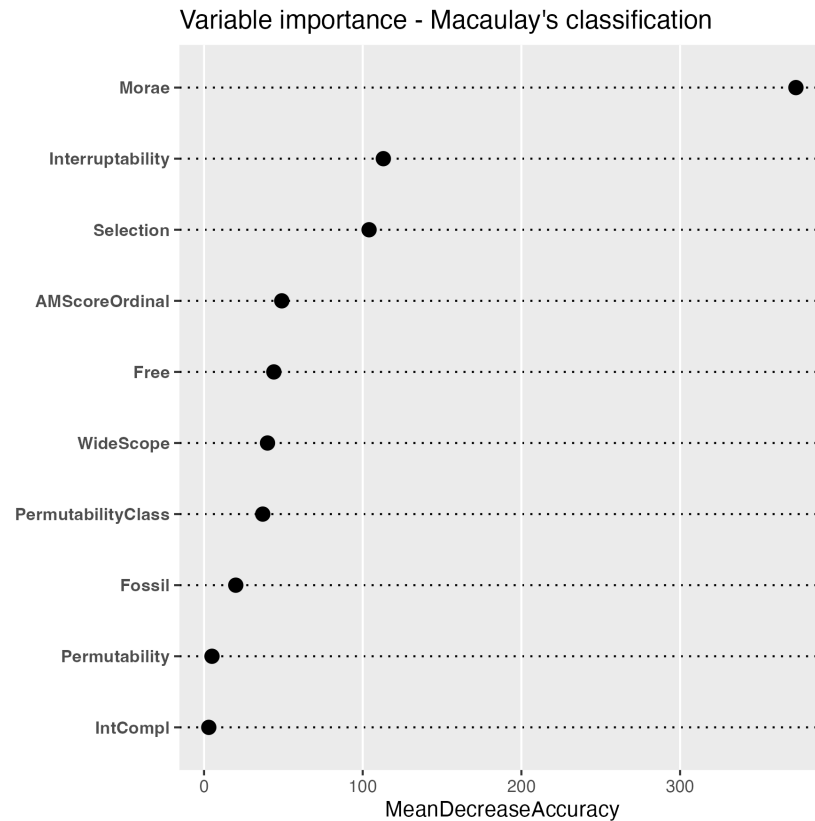
Hierarchical clustering of the data

red = dependent pronouns
green = independent pronouns





WHICH VARIABLES ARE IMPORTANT?

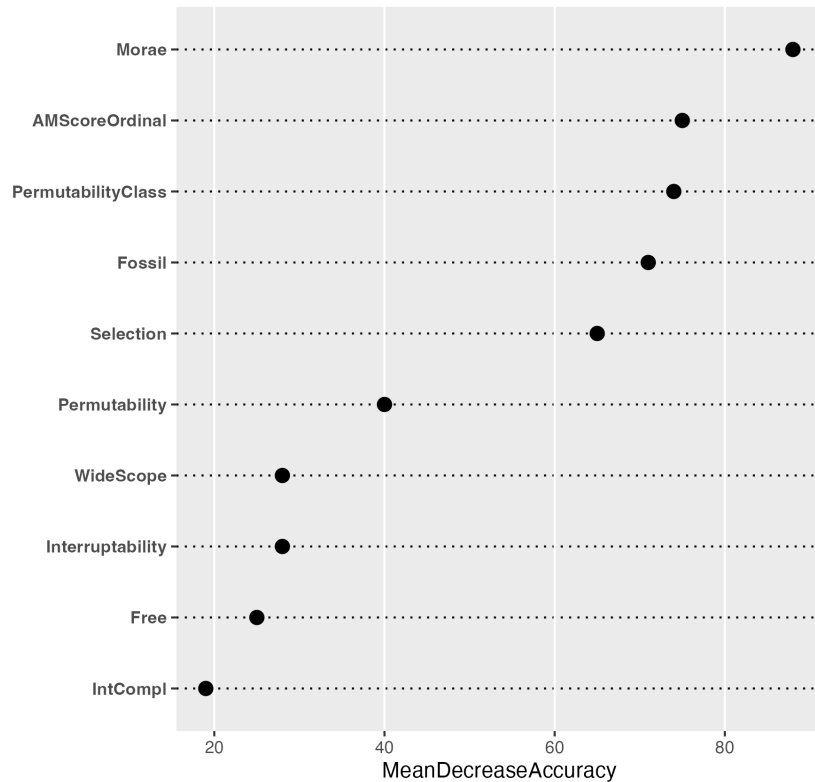


```
No. of variables tried at each split: 7  
  
OOB estimate of error rate: 4.63%  
Confusion matrix:  
      Affix PhAffix Word class.error  
Affix    7      3    0 0.30000000  
PhAffix   1     23    0 0.04166667  
Word      0      1   73 0.01351351
```

baseline	accuracy	difference
68%	85%	17%

WHICH VARIABLES ARE IMPORTANT?

Variable importance - first cluster part.

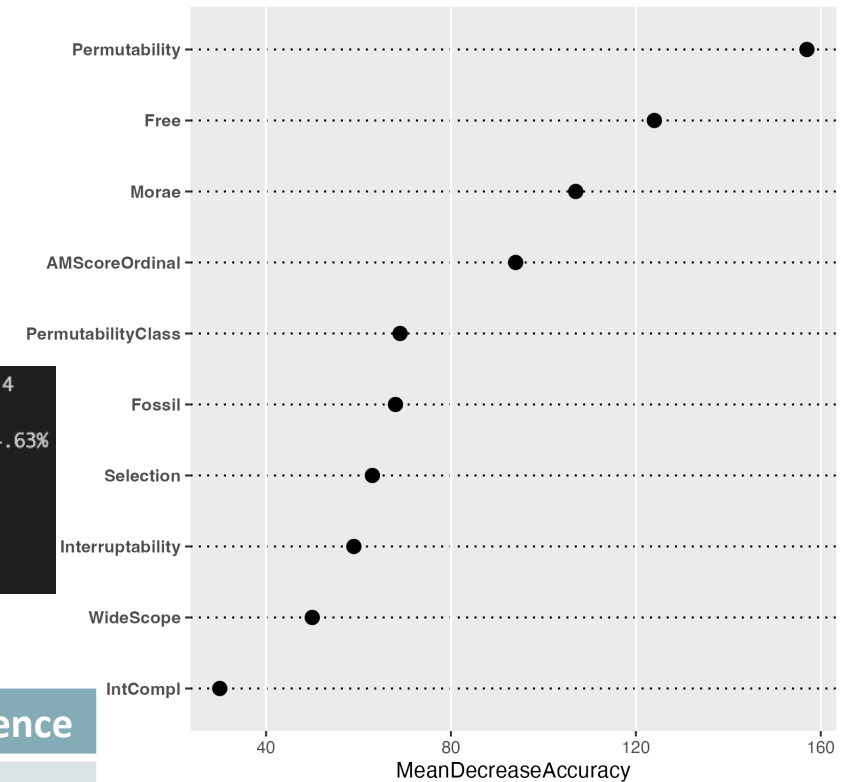


```
No. of variables tried at each split: 2  
  
OOB estimate of error rate: 0.93%  
Confusion matrix:  
  1  2 class.error  
1 79  1    0.0125  
2  0 28    0.0000
```

```
No. of variables tried at each split: 4  
  
OOB estimate of error rate: 4.63%  
Confusion matrix:  
  1  2  3 class.error  
1 28  0  3    0.09677419  
2  0 27  1    0.03571429  
3  0  1 48    0.02040816
```

part.	baseline	accuracy	difference
1st	74%	99%	25%
2nd	50%	98%	48%

Variable importance - 2nd cluster part.



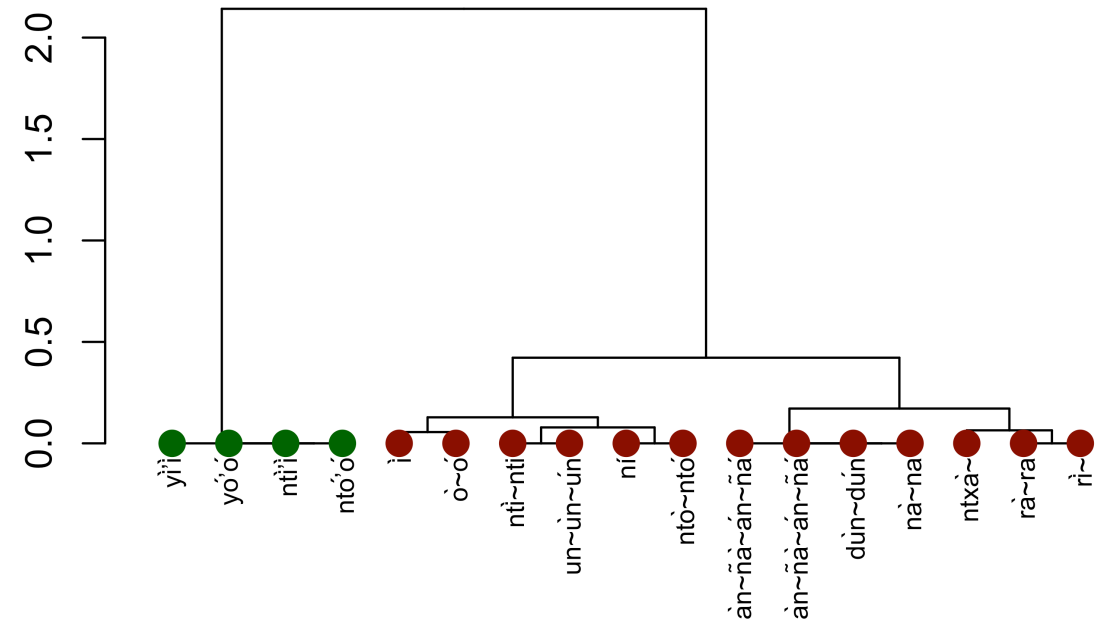
PRONOUN CLASSES?

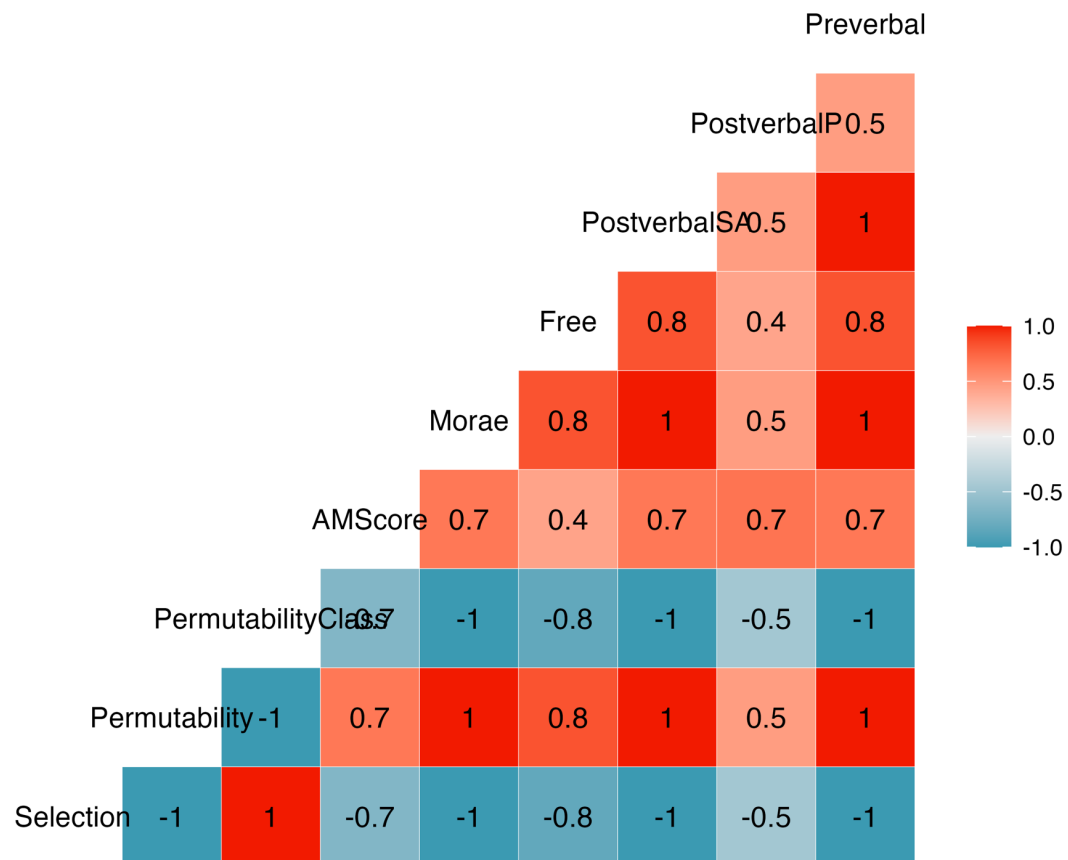
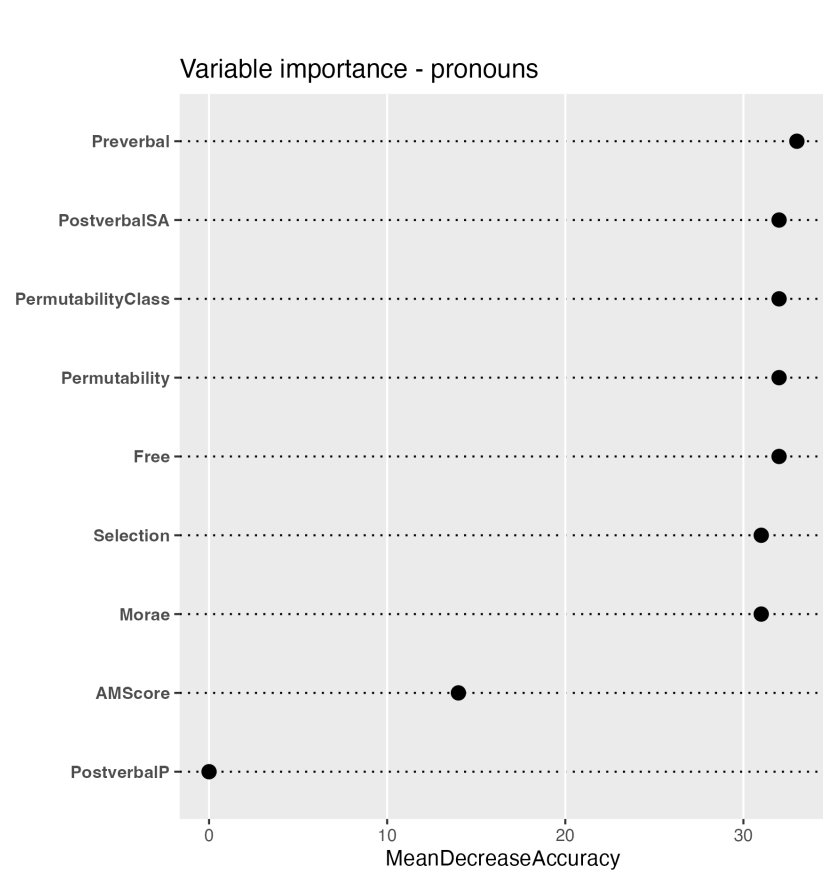
Hierarchical clustering of pronouns

wordhood and alignment variables

green = independent pronouns

red = dependent pronouns





TAKE AWAYS

for Mixtec

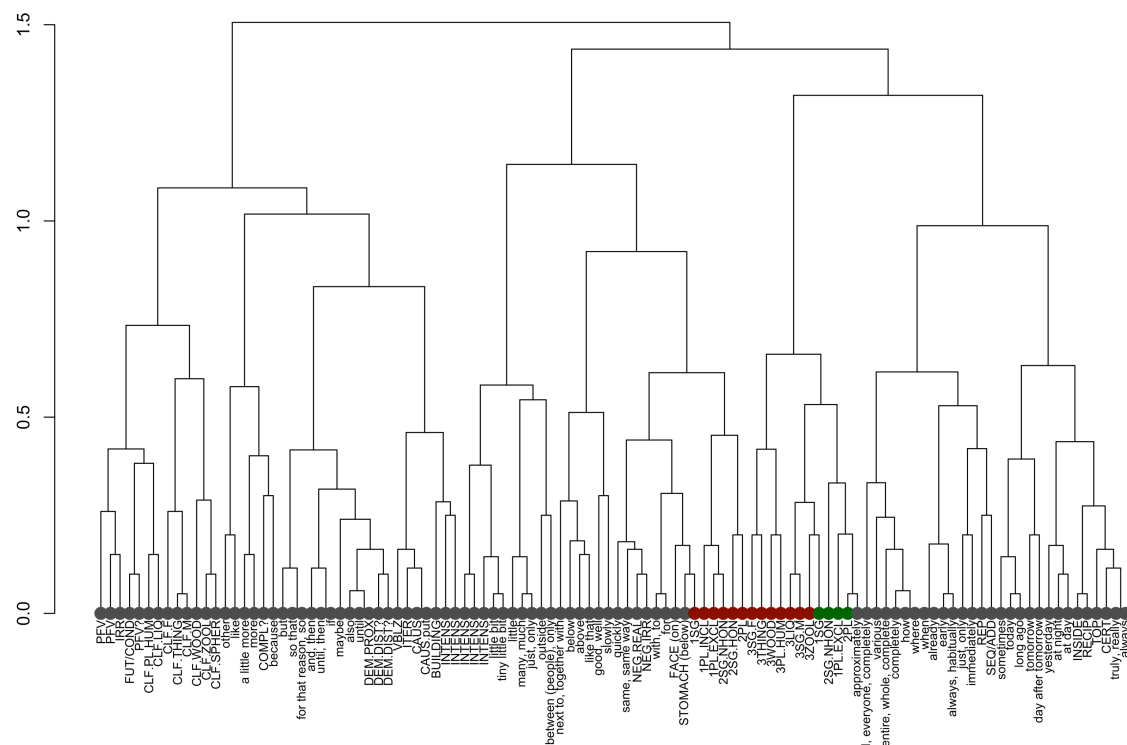
- 'clitics' don't solve any descriptive problems
- there is no language internal motivation for a division between morphology and syntax
- based on two methodologies (word-hood tests on elements or spans)
- pronouns are not a (formal/distributional) class

for Typology

- there should not be a typology of clitics
- alignment typology can be integrated and improved into constituency/wordhood based typology via **positions as 'selectors'**

SIMULATIONS

sampled from data



sampled from uniform distribution

