

Chapter 15

Wordhood in Chorote (Mataguayan)

Javier J. Carol

University of Buenos Aires

This chapter discusses the results of applying constituency tests to Chorote (Iyo-jwa'aja' or "Ribereño" variety, ISO code: crt), a Mataguayan language spoken in Argentina and Paraguay. The outcomes are interesting for a number of reasons.

Firstly, Chorote has the highest number of positions of all the languages surveyed in this volume: 46 positions. There are several reasons for this: the high number of categories expressed (lexical arguments/adjuncts, personal clitics, agreement and TAME markers, adverbs, applicatives), the strict ordering of elements such as adverbs or applicatives, which forces one to assign a distinct position to each of them, and the fact that many morphemes can appear in multiple positions in the template.

Secondly, no clear wordhood candidate emerges from the application of constituency diagnostics. Only one subspan reaches the convergence of three diagnostics, and four or perhaps five more spans the convergence of two diagnostics each. Chorote, thus, provides another example of a language where language-internal evidence suggests no quantal shift from word (morphotactic) to phrasal (syntactic) structure, but rather a smooth cline, morphology-like to syntax-like combination.

Thirdly, Chorote makes apparent some methodological problems in the application of the diagnostics based on ciscategorial selection. A general issue is that it is difficult to define a cross-linguistically valid notion of transcategoriality given the presence of mixed constructions, i.e. those that cannot be neatly categorized as verbal or nonverbal predication. The predicate in Chorote can be syntactically headed not only by verbs but also by Ns, NPs, DPs, pronouns, and negation, which take then most of the usually "verbal" markers. Furthermore, NPs and DPs can take some of the "verbal" TAME markers even when they function as arguments. All these facts pose questions regarding how ciscategoriality should be defined as a comparative concept, since it is not clear whether it should be defined with respect to verbs or to predicates in general.



1 Introduction

This chapter describes and discusses the application of a set of constituency tests to the verb complex of Chorote, a Mataguan language spoken in the Argentine and Paraguayan Chaco. The data come from extensive original fieldwork (over 120 days in yearly trips from 2005 to 2011) and text corpora, namely the collection of 33 short stories compiled in Drayson et al. (2000), a New Testament translation into Gomez & Drayson (1997), cited by the corresponding verses, and a few other scattered materials.

Chorote is interesting for typological issues in constituency for a number of reasons. First, it has more positions than any other language surveyed this volume: 46 positions. The number of categories expressed (lexical arguments or adjuncts, pronominal clitics, agreement and TAME markers, adverbs, and a number of applicatives) is relatively high. Furthermore, many of them display a strict order with respect to each other, which forces one to assign a distinct position to each of them. But this alone does not suffice to explain the number of positions in the template. Many morphemes can occur in many different positions in the template. Some TAME markers, for instance, may occur bound not only to the verb (or non-verbal predicate) but also to other word-classes or phrases. Moreover, there is a critical distributional distinction between DPs headed by demonstratives, and NPs (or DPs headed by possessives). The former appear to the right of the predicate that selects them – verb, noun or adposition/applicative, and the latter to the left. This results in a duplication of many positions dedicated to verbal arguments, and also those of the applicatives/adpositions (these are a single lexical class that attach to the verb (“applicatives”) or to the nominal (“adpositions”)).

Secondly, no clear wordhood candidate emerges from the application of the constituency tests used in this chapter. The diagnostics discussed here are both morphosyntactic (free occurrence, non-interruption, non-permutability, subspan repetition, deviation from biuniqueness) and phonological (accent, *yod*-insertion, palatalization). All but one are fractured into two or more subtests. All in all, the highest convergence level is two diagnostic subtests per layer, which is reached by five or perhaps six subspans. There is significant convergence in the left edge of the domain containing the predicate: 14 subtests converge on the position of the verbal prefixes and a further six subtests on the verbal stem or predicate. However, there is much less convergence on the right edge. The highest convergence is again three subtests, reached by three or maybe four positions. But if we do not count subtests of the same diagnostic, convergence is reduced to two diagnostics in two of those positions. Chorote thus provides another example of a language where language-internal evidence suggests no quantal shift from word (morpho-

tactic) to phrasal (syntactic) structure, but rather a relatively smooth cline from morphology-like to syntax-like combination (Tallman 2021).

Thirdly, Chorote reveals some methodological problems in the application of the diagnostics based on ciscategorical selection. A general issue is that it is difficult to define a cross-linguistically valid notion of transcategoriality given the presence of mixed constructions, i.e. those that cannot be neatly categorized as verbal or nonverbal predication.¹ The ciscategoriality diagnostic for verbs can be formulated as follows: if a morpheme can only co-occur with verbal bases but not with other word classes, then it belongs to the verb word, i.e. it is ciscategorial to it. But the predicate in Chorote can be syntactically headed, not only by verbs, but also by nouns (N), noun phrases (NP), determiner phrases (DP), pronouns, and negation, which then take most of the usually “verbal” markers. Furthermore, NPs and DPs can take some of the “verbal” TAME markers even when the former function as arguments. All these facts pose questions regarding how ciscategoriality should be defined as a typological variable (or perhaps as a comparative concept), since it is not clear whether it should be defined with respect to verbs or with respect to predicates in general.

The rest of this chapter is organized as follows: §2 gives information on the language and its speakers; §3 presents the predicate planar structure, §4 presents morphosyntactic wordhood diagnostics except for those related to ciscategoriality, §5 presents phonological diagnostics, §6 presents diagnostics based on ciscategorical selection and discusses some implications of the Chorote facts, and §7 concludes.

2 The language and its speakers

Chorote is a Mataguayan language spoken by about 2000 people in Argentina (Salta Province) and about 500 in Paraguay (Boquerón county), where it is known as Manjúi. The language is fairly vital, and children learn it in both countries, although much less in communities surrounding the cities (Tartagal in Argentina, Mariscal Estigarribia in Paraguay).² The family is often called “Matakoan” in

¹I would like to thank Adam Tallman for pointing this issue out to me.

²The 2010 census in Argentina gives the number of 2270 people who recognize themselves as Chorote (Instituto Nacional de Estadística y Censo 2012). The *Encuesta Complementaria de Pueblos Indígenas* [Complementary Poll on Indigenous Peoples] conducted in 2004–2005 documented 1700 Chorote speakers under a total of 1768 participants (Instituto Nacional de Estadística y Censo 2005). The 2012 census in Paraguay gives the number of 582 Manjúi (Dirección General de Estadísticas 2012). Based on my own personal field experience, I would say it is reasonable to assume that most or almost all those who recognize themselves as Manjúi speak the language and that probably a majority of them are monolingual.

English-speaking literature, according to the convention of naming the family by adding *-an* to the name of the better-known language. But since *Matako* is considered pejorative, the Spanish-speaking literature prefers *mataguayo* (Mataguan). This chapter focuses on the riverside or *Iyojwa'aja'* variety (ISO code: crt), spoken in Argentina in communities by the Pilcomayo river and surrounding the city of Tartagal.

The language has a complex phonology. It has six phonological vowels; /a e i o u a^{*}/. The phoneme /a^{*}/ (or rather /ɑ/) is realized as /e/ after a palatal(ized) phone. In the same environment, /e/ is realized as /i/ and merges with phonological /a/ elsewhere. When no palatal(ized) phone precedes it, stressed /i u/ vowels lower to mid vowels, but do not merge with phonological /e, o/, which in turn are open in such environment. The surface contrast is thus roughly [e o] versus [ɛ ɔ]. In the notation used here, lowered phonological /i u/ are transcribed <ɛ ɔ>; the practical spelling used in the communities, including in educational and religious texts does not distinguish them from phonological /e o/, using <e o> in either case.

Plain consonants are /p t kⁱ k hw hl l s h m n w j ʔ/ (<p t ky k jw jl l s j n w y 'y>). There is a series of laryngealized consonants /p' t' kⁱ' k' ts' ʔm ʔn ʔl ʔw ʔj/ whose phonological status is debatable, at least for non-stop sounds (Carol 2014a, Gutiérrez & Nercesian 2021). I transcribe the laryngealized consonants as <C'> for stops and as <'C> for sonorants. A widespread process of progressive palatalization creates palatalized allophones for almost any consonant, including laryngealized ones. They can mostly be regarded as a sequence of C plus a palatal glide, which I transcribe as <y>. Palatalized phones are accordingly transcribed as <Cy> (although recall that velars <ky, k'y> have phonemic status; they can have in turn palatalized counterparts, which are transcribed <sy, ts'y>). To avoid confusion, from now on I will use the orthographic conventions even in phonological representations, e.g. /y/ represents a palatal glide, except for the fact that I will use /h/ and /hC/ to correspond to orthographic *j* and *jC* respectively.

A sequence of a consonant and a glottal stop produces a laryngealized consonant, e.g. *y+ 'ut* → *'yut* '(s)he puts in'; a sonorant plus a voiceless laryngeal (transcribed as <j>) gives <jC> e.g. *in+jetik* → *ijnetik* 'someone's head'. In these cases, the corresponding glosses are separated by a colon, e.g. 3:put for *'yut*. A voiceless laryngeal is lost after an obstruent. Before a pause, any sonorant is laryngealized: for vowels, a final <'> is added, and for consonants the stop is added before the consonant, e.g. *jlam* '(s)he/it' → *jla'm /_##*; sometimes an "echo vowel" is also inserted, e.g. *jla'am*.

Syllable structure is CV(C). To avoid onsetless syllables, /y/ is inserted between a suffix/enclitic that begins with a vowel and a base that ends in a vowel, and a glottal stop <'> is inserted elsewhere. This will be described in more detail in

§5.3. Onset position can only be filled by one plain or laryngealized consonant (or by clusters of C and glottal stop if these are not considered single phonemes), including their palatalized allophones. Neither palatalized or phonemically laryngealized consonants nor /w, h/ occur in coda position. However, laryngealized consonants can occur in final position because of the phrase-level process mentioned above that laryngealizes any sonorant before pause. The language displays stress marking, realized as an increase of intensity in the stressed syllable. The stress falls in the first or second syllable of the stem and is thus not fully predictable. In the practical spelling used here, an acute accent is used to mark any stress that does not fall on the first syllable of the stem, and any orthographic word has no more than one stress. Raising and neutralization processes often occur in unstressed vowels.

The practical spelling adopted here follows the one used in the speech communities and introduced by missionaries, but differs in the following respects: 1) the orthographic word can only have one accent, so I write *i-ni* 'wenis 'they see each other' or *ta-ka leja'n* 'they wash (antipassive)' (stressed syllables in boldface) instead of *ini'wenis*, *takaleja'n* as in the missionaries' spelling; 2) I transcribe the palatalized consonants as *Cy* instead of *Ci* before vowels, e.g. *kya* instead of *kia* 'that (moving away or disappeared)', 3) I use graphical accents when they don't fall on the first syllable of the stem (missionaries' spelling uses no accents), and 4) I distinguish *ɛ*, *ø* from *e*, *o*.

The language has an active-inactive alignment, and phrases can be both head-initial and head-final, depending on the "weight" of the complement - complements headed by a demonstrative are "heavy" and follow the phrase head, while other complements are "light" and precede it. This will be dealt with in detail in §3.3.

In glosses, no distinction is made between clitics and affixes, and morphemes are always separated by "-", i.e. "=" is not used. The main reason is that the distinction between affix and clitic is mainly based on the notion of "word", which is what this chapter seeks to scrutinize. As will be seen, it is not easy to determine what a word is in Chorote. Furthermore, the distinction often relies on assumptions related to cis-/transcategoriality (affixes have categorial preferences, clitics do not) which are complicated to apply in Chorote compared to other languages.

3 The predicate planar structure

The predicate planar structure of Iyojwa'aja' Chorote has 46 positions (see Table 1).

Table 1: The predicate planar structure

Position	Type	Elements
(1)	Zone	Conjunction or conjunctive locution; interrogative <i>ma(?)</i>
(2)	Zone	Topic: DP {A, S}, Adverb, Adverbial clause
(3)	Zone	Focalized adverb or AdvP
(4)	Slot	Complementizers <i>ti, ka</i>
(5)	Zone	N', N, light DP, Pronoun/Adverb
(6)	Zone	N', N, light DP, Pronoun/Adverb
(7)	Slot	Mirative <i>-p'an</i> ; reportative <i>-jen</i>
(8)	Slot	Prospective <i>ja</i>
(9)	Slot	Incompletive <i>-ta(j)</i>
(10)	Slot	Negation <i>je</i>
(11)	Slot	Indirect evidential <i>-t'i</i> ; mirative <i>-p'an</i>
(12)	Slot	Interrogative <i>-mi</i>
(13)	Slot	Demonstrative <i>kyak</i> 'that (way)'
(14)	Slot	Cross referencing active/nonactive markers
(15)	Slot	Reflexive/reciprocal <i>ni(n)</i> ; antipassive <i>ka</i>
(16)	Slot	Predicate base
(17)	Slot	Participle <i>-k</i>
(18)	Slot	Causative; antipassive; verbalizer
(19)	Slot	Concord 1
(20)	Slot	Concord 2
(21)	Slot	Perdurative <i>-jli</i>
(22)	Slot	Momentary <i>-a</i>
(23)	Slot	Irrealis <i>-a</i>
(24)	Slot	Reportative <i>-jen</i>
(25)	Slot	Indirect evidential <i>-t'i</i>
(26)	Slot	Mirative <i>-p'an</i>
(27)	Slot	Incompletive <i>-ta(j)</i>
(28)	Slot	Interrogative <i>-mi</i>
(29)	Slot	Concord (3PL, A/S) <i>-is</i>
(30)	Slot	Applicative: orientation <i>-a(j)</i>
(31)	Slot	Applicative: instrumental <i>-e(j)</i>
(32)	Slot	Applicative: distal <i>-ey</i>
(33)	Slot	Applicative: locatives <i>-jiy, -jam</i>
(34)	Slot	Applicative: punctual locative <i>-'e</i> , distributive/comitative <i>-k'i</i> , and possibly others

(35)	Slot	Oblique marker (only realized if position 36 is filled)
(36)	Slot	Applicative
(37)	Slot	Pluractional/downwards <i>-jen</i>
(38)	Slot	Pluractional/iterative <i>-’ni(j)</i>
(39)	Slot	Remote past <i>-pe(j)</i> ; perfect <i>-(’V...)je(j)</i>
(40)	Zone	Temporal, aspectual and discourse particles
(41)	Slot	Locatives
(42)	Zone	N, N’, light DP
(43)	Zone	N, N’, light DP
(44)	Slot	Adpositions-applicatives
(45)	Zone	Heavy DP (A, S, O, Obl, Possessor of)
(46)	Zone	Heavy DP (A, S, O, Obl); S

The analysis whereby Chorote has 46 positions is conservative for two reasons. First, at least some of the particles in position 40 might be rigidly ordered with respect to one another, which might require that we split this position into many. These particles include *pet* ‘please’ (among other meanings), *’ne* ‘then’, *-na’a* ‘later’, among many others. Some appear always in the same order in texts, though no exhaustive elicitation work has been conducted to test whether this order can be altered.

Secondly, texts produced by elderly speakers occasionally show a few morphemes in positions not recorded in Table 1; this will be shown in §3.2. My main consultant, >60 years old when the fieldwork was conducted (2007–2011), accepted these forms in elicitation, though he did not produce them spontaneously, nor did other consultants represented in my corpus. Therefore, I decided to build a table with a “standard” version of the language and exclude these cases.

On the other hand, the number could be reduced if we assume some internal constituent. In effect, one could assume an initial host position (e.g. a complementizer head) and a clitic cluster bound to it. Under such an analysis, the number of positions that hosts fronted elements (which is especially high in the Chorote varieties discussed in §3.2), as well as the many positions for some bound morphemes, could each be reduced to one. The positions could also be reduced if some morphemes (e.g. adverbs) proved to modify others, so that they would constitute a zone, rather than distinct positions.

The rest of this section is organized as follows. §3.1 describes the orthographic word in Chorote, §3.2 discusses the fact that some elements can occur in more than one position, and §3.3 explains the distribution of NPs and DPs in Chorote.

3.1 The orthographic word

This section is concerned with the orthographic “word” in Chorote, which largely corresponds to the practice of writing spaces by missionary linguists. Such a notion offers an idea of what the traditional notion of word is like in the language. In broad terms, it corresponds to a stress domain. Thus, the orthographic word containing the verb core usually has the unstressed personal prefixes of position 14 as its left edge. But when the unstressed prospective marker *ja* of position 8 surfaces, it usually becomes the left edge and, more rarely, the complementizer *ka* of position 4 (most usually written in a separate word).

However, when the stressed reflexive or antipassive markers of position 15 surfaces, the correlation between the orthographic word and the stress domain is broken. Drayson and the authors of religious texts include it in the orthographic word altogether with the verb core, so that the word has two accents - in position 15 and in position 16, e.g. *i-ni-‘we’en* (14-15-16) ‘(S)he sees himself/herself’. Nevertheless, some native speakers split this into two orthographic words, with only one stress each: *i-ni ‘we’en* (14-15 16).

As for the right edge, the remote past marker of position 39 can be the right edge or appear in a different orthographic word, which is consistent with its facultative stress. Something similar occurs with the markers of position 40, some of which are stressed, while others are not.

But when the oblique second person markers of position 35 surface, the correlation with the stress domain is again broken. This morpheme bears (secondary) stress, despite which it is written altogether in word with the verb in most texts. (Although I have not tested it, it is possible that some speakers would write it as a different word).

In the two cases mentioned above in which, in the missionaries’ writing, orthographic word does not correspond to a stress domain, some sort of ‘morphological word’ criterion seems to be involved. Namely, the stressed morphemes of position 15 form a morphological domain with the verb core of position 16 according to all the tests that will be discussed below. Similarly, the marker of position 35 forms a domain, at least, with the material immediately to its left, if not with the verb core too.

3.2 “Promiscuous” elements

Some TAME markers may appear bound not only to the predicate but also to other hosts that are often defined as phonological domains, and some of them even occur as free particles: see the incomplete *-ta(j)/-tye(j)* bound to the verb

in (1a) and to the prospective pre-verbal particle in (1b); the indirect evidential *-t'i(y)* bound to the verb in (2a) and to negation (also pre-verbal) in (2c); the interrogative *-mi* bound to the verb in (2b) and to negation in (2c), and as a sentence-initial particle in (2d);³ and the mirative *-p'an* bound to the verb in (3a), to negation in (3b) and to an initial DP/NP in (3c).⁴

(1) Incompletive *-ta(j)/-tye(j)*

a. post-verbal

Y- ik- tye

v: 14 16 27

3- leave -INCOMP

'(S)he was leaving/about to leave (but didn't).'

b. After prospective

ʃo- ta y- ik

v: 8 9 14 16

PROSP- INCOMP 1.IRR- leave

'I would leave', 'I intended to leave'.

(2) Indirect evidential *-t'i(y)* and interrogative *-mi/ma*

a. post-verbal evidential

'Yijén- t'iy- i

v: 14:16 25 31

3:know- EVID- AP.INS

'(S)he must probably know.'

b. post-verbal interrogative

ʃi- woy- e- t'i- tye- mi [ka Ø- 'wen- a syunye

v: 14 16 20 25 27 28 46 - - - -

2- LV- 2PL- EVID- INCOMP- INTER COMP- 2.IRR- see- 2PL DEM

³I tentatively assign this particle the same position as the conjunctions, because they do not introduce a topic or a focus, and there is evidence they precede adverbial clauses. Since many of the neighbouring positions do not seem to occur in interrogatives, the precise relative position of *ma* is difficult to determine.

⁴Cases where morphemes are basically unsegmentable but occur over a whole span are glossed with an x:y notation, where x is the left-most position occupied by the morph and y is the right-most element. An example is *'yijén-* 'THIRD PERSON+know' in (2a), which covers positions 14 through 16. Where a position involves more than one morpheme, as is often the case in zones, the position is enclosed in brackets and only the first one receives a number. An example is *kya si'yús* 'the fish' in (2a).

i'nyó'...]

-

man

'Did you intend to see that man?' (Mt 11:8)

- c. Evidential and interrogative after negation

ǀJe- t'i- mi 'náján- ay- i..?

v: 10 11 12 14:16 20 31

NEG- EVID- INTER 2:know- 2PL- AP.INS

'Don't you (pl.) know.'

- d. Sentence-initial interrogative

ǀMa y- am- 'nijne?

v: 1 14 16 38:39

Inter 3- go.away- PLACT:PRF

'Are they already gone?'

- (3) Mirative *-p'an*

- a. post-verbal

ǀA- sɛnyan- p'an- taj [kya si'yús]!

v: 14 16 26 27 45 -

1- roast- MIR- INCOMP DEM fish

'I have roasted the fish in vain!' (E.g. because you are not hungry)

- b. After negation

ǀJe p'an Ø- ɛs- i- ji [ka i'nyát i'nyó Ø- wujw-

v: 10 11 14 16 31/32 32/33 46 - - - -

NEG MIR 3- be.good- AP.INS AP.LOC COMP water man 3- be.big-

a- ja'm]

- -

IRR- AP.LOC

'It is not good if a man gets full of water.' (Drayson et al. 2000: 116)

- c. After initial NP/DP

[As- taj] p'an je Ø- wujw- a- k'i, [ti a- 'wen-

v: 5 - 7 10 14 16 22 34 46 - -

2PL.POSS- sight MIR NEG 3- big- MOM- AP.DISTR COMP 1- see

k'i ja- pɔ i'nyó'...]

- - - -

-AP.DISTR DEM- PL man

'It turns out that you (pl.) do not remember when I saw those men.'

(Lit. 'your sight is not large enough to') (Mt 16:9)

Many, if not all of these morphemes, have the appearance of second position clitics, with differences regarding how far they can move to the left, the mirative *-p'an* and the reportative *-jen* reaching the leftmost positions. Therefore, the high number of positions might be somewhat deceiving. If the approach mentioned above turned out to be correct, each morpheme of the clitic cluster would have its own position in the planar fractal method, but the different positions of each morpheme when fronted could be reduced to one: bound to the initial host position. The mirative *-p'an*, for instance, occupies a second position in both (3b) and (3c) but, since in (3b) negation occupies the first position and in (3c) a DP precedes negation, *-p'an* appears to the right and to the left of negation respectively, in two apparently different positions. If the host positions, namely negation in (3b) and the DP in (3c), were the same, so would be that of *-p'an*. We leave more thorough consideration of this possibility for future research.

Moreover, as mentioned above, some texts by elderly speakers occasionally contain morphemes occurring in positions not recorded in Table 1 as in (4).

(4) Elements in positions not recorded in Table 1

- a. Discontinuous first person plural inactive with mirative *p'an* in between

kas p'an ts'elya- k'i'

kas p'an s- 'ila- k'ye

v: Ø 7 14 16 34

1PL MIR 1PL- be.eager- AP.DISTR

'We wanted to keep (eating)!' (Drayson et al. 2000: 40)

- b. Discontinuous first person plural inactive with negation *je* in between

kas- é s- oyme'n

v: Ø 10 14 16 38

1PL- NEG 1PL- be.drunk;JEN

'We are not drunk.' (Gerzenstein 1983)

- c. Indirect evidential *-t'i* attached to reflexive/reciprocal

i- ni- t'i 'wen- k- in- a- ja'ajme [ja Santiago]

v: 14 15 Ø 16 17 18 22 33:39 45 -

3- REFL- EVID see- PP- VBLZ- MOM- AP.LOC:PRF DEM Santiago

'He appeared in a vision (lit. 'made himself seen') to Santiago.' (1 Cor 15: 7)

The properties of the first person plural inactive *kas-* might have a metrical explanation, at least historically. Capable of forming a closed syllable, unlike the

personal prefixes of the form CV, *kas-* could have constituted a minimal foot (and therefore a minimal word). This has been argued for in the sister Nivacle language by Gutiérrez (2015: 178–179), where a cognate of *kas-* exists, although in the nominal domain. Thus, if a morpheme needs to be not smaller than a foot in order to appear left-dislocated and/or to be a host, *kas-* or more precisely its first segment, is the only verbal person marker that meets this condition. (In the “standard” variety the first person plural inactive occupies position 14 only, and there is no reason to assume two segments.)

3.3 Distribution of DPs/NPs

Another factor that multiplies positions in Chorote is the fact that complements may appear to the right or to the left of their heads. Complements headed by a demonstrative are heavy and surface to the right, while other complements are light and surface to the left. Examples (5a) and (5b) illustrate this in the nominal domain. Here the possessor appears to the right in (5a) and to the left in (5b).

(5) Heavy and light complements in the nominal domain

a. Heavy complement

jl- as na Juan

3POSS- son DEM Juan

‘Juan’s son.’

b. Light complement

Juan jl- as

Juan 3POSS- son

‘Juan’s son’

The phenomenon in the verbal domain is shown in (6), where (*puwa*) *alenas* ‘the dogs’ is the subject of a transitive verb (A). But in the first sentence it is a heavy DP and appears post-verbally, in position 45, whereas in the second sentence it is light (=no demonstrative) and appears pre-verbally, in position 5. (Light NPs/DPs are licensed to the right when followed by an irrealis nominal marker, however; see (14), (15) and (17) below).

(6) Heavy and light complements in the verbal domain

<i>i-</i>	<i>'wi'in</i>	<i>[pu- wa alena- s]_(A)</i>	<i>[Alena- s]_(A)</i>	<i>i-</i>	<i>jyan-</i>	<i>a-</i>	<i>'ni</i>
v: 14	16	45	- - -	5	-	14 16	22 38
3-	find	DEM- PL	dog- PL	dog-	PL	3- chase-	MOM- PLACT

[ja- pa 'najáte]_(O)

45 - -

F- DEM rabbit

'...the dogs find [some rabbit]. The dogs then chase the rabbit.' (Drayson et al. 2000: 48)

The same conditioning on the position applies when the NPs/DPs are complements of adpositions/applicatives, but in this case it affects the position of the heads as well. In effect, Chorote has morphemes that may encliticize to their complements ("adpositions") or to the verb ("applicatives"), depending on whether they take a light complement or not, respectively. It is argued elsewhere (Carol 2011, Carol & Salanova 2012) that they are simply grammatical adpositions that may occur superficially bound to the verb in some cases. They are glossed here indistinctly as "P".

In (7a) there is no light complement, but an optional heavy complement (*na Mosik*); hence, P encliticizes to the verb. (7b) and (7c) exemplify Ps with light complements. These may be NPs, as in (7b), or oblique pronouns, as in (7c), which have different positions in the template. The complex element *oblique pronoun+adposition*, in turn, attaches to the verb. Therefore, the same applicative/adposition (the distal in this case) may appear in three different positions: 32, 36 and 44.

(7) Heavy and light complements with Ps

a. P bound to V; optional heavy complement

Ø- Tajl-	ej-	e	<i>wek</i>	([<i>na Mosik</i>])
v: 14 16	31	32	40	(45 -)

3- come.out- AP.INS- AP.DIST finally DEM Mosik

'(S)he finally obtains it (from Mosik).' Lit. 'comes out with it from Mosik.'

b. P bound to a light nominal complement

Ø- Tajl-	e	<i>tewk-</i>	i
v: 14 16	31	43	44

3- come.out- AP.INS river- AP.DIST

'(S)he finally obtains it from the river.'

c. P bound to a pronoun

Ø- Tajl-	a	'a-	i	<i>wek</i>
v: 14 16	31	35	36	40

3- come.out AP.INS- 2- AP.DIST- finally

'(S)he finally obtains it from you.'

4 Morphosyntactic diagnostics

This section discusses the morphosyntactic diagnostics, except for those related to ciscategorial selection, which will be discussed in more detail in §6.

4.1 Free occurrence (16-16; 4-40)

This abstract type identifies the minimal free form, i.e. a complete utterance that is a single free form. The test can be fractured into *minimal* and *maximal*. The *minimal* subtype identifies the smallest possible minimal free form that contains the predicate head. This corresponds just to position 16, i.e. the verb root or non-verbal predicate, which stands alone as an utterance in imperatives, e.g. *kasit* ‘stand up’.

The *maximal* subtype identifies the span defined by the largest minimal free form that contains the predicate head, i.e. the largest possible span containing a predicate head (typically a verb) plus the more distant bound elements to the right and to the left, such that no other free form intervenes. This defines the subspan 4-40. I have no examples of this subspan in main clauses in spontaneous speech, but (8a) shows an example in an embedded clause. The left edge of this span is occupied in main clauses by the complementizers *ka* and, more rarely, *ti*.⁵ *Ka* selects for the irrealis mode on the predicate and behaves as a phonological proclitic. It heads some complement, conditional, temporal and in general future-oriented embedded clauses, as well as some main clauses including optative, hortative, and prohibitive. An example of *ka* on the left edge in a main clause is given in (8b).

(8) Free occurrence *maximal*

a. Subspan 4-40 in embedded clause

<i>ka</i>	Ø-	'nes-	ta-	<i>na'a</i>
v: 4	14	16	27	40
COMP 3.IRR- arrive- INCOMP- later				
‘When (s)he/it arrives.’				

b. Left edge in main clause

<i>Ka</i>	y-	ik
v: 4	14	16
COMP 1.IRR- go.away		
‘I’m leaving, may I go’.		

⁵The complementizer *ti* introduces temporal completives and others selecting *realis*. Examples of *ti* in main clauses are not as clear as those of *ka*; one of them is the second *ti* of (48). The interrogative *ma* does appear in main clauses, but its inclusion in this position is only tentative.

Positions 1 through 2 include phrases and other elements that can occur as free forms and are thus excluded from the span. Notice that there are free forms between 4 and 16 but, since they are not obligatory, they are irrelevant for this diagnostic.

The right edge is more problematic. In 8a it is represented by the adverb *na'a* 'later' in position 40. Example (9) below is another example of *na'a* as right edge but in a main clause; *pet* 'please' also belongs in position 40.

- (9) Free occurrence (maximal): right edge (Drayson et al. 2000: 70)

	<i>Jo-</i>	<i>kyu-</i>	<i>nye'e pet</i>	<i>ts'iji</i>	<i>[i-</i>	<i>'wit-]</i>	<i>e</i>
v: 16	40	40	40	41	43	-	44
	go-	a.while-	later	please	there	1POSS-	place- AP
	'Go to my place (i.e. my house) there.'						

This adverb does not occur as a free form in my material. But at least some of the other adverbs of position 40 are free forms, e.g. *t'e*,⁶ which can function as an answer to questions with the meaning 'who knows'. For other elements in the same position I have no conclusive evidence regarding their bounded character; some can appear fronted in position 2 (e.g. *ta'a* 'immediately', 'already') and have been spontaneously uttered alone at least in metalinguistic uses, unlike applicatives and unstressed TAME markers, but perhaps the latter holds as well for the remote past *pe(j)* of position 39; notice that both positions 39 and 40 host optionally stressed elements. As for material to the right of position 40, there is no doubt it must be excluded from the span. The locative adverbs of 41 *ts'iji* 'there' or *niji* 'here' are clearly free forms. Further to the right the only bound elements are the adpositions in position 44, but they occur bound to nouns (in position 43), which in turn are free forms; the adpositions can also occur bound to verbs functioning as applicatives, but they are then analyzed as occurring in a different position, see §3.3.

4.2 Non-interruption (14-39/38; 7-41; 14-22)

The diagnostic of non-interruption identifies the span of positions that includes the predicate head and cannot be interrupted by some interrupting element (Tallman 2021). The diagnostic is fractured according to how the interrupting element is defined. The *single free interruptor* subtest defines the interruptor as any free

⁶This has the same underlying form as the evidential *-t'i* of position 23, namely /t'ey(h)/. The difference in the surface forms correspond to regular differences between stressed and unstressed vowels.

form and is the most straightforward version of the diagnostic. This gives the span 14-39 or maybe 14-38, see (10).

(10) Non-interruption - *single free interruptor*

a. Span 14-39

Y- am **pej**
 v: 14 16 39
 3- go.away **REM.PST**
 ‘(S)he/it left.’

b. Span 14-38

Y- am- **’ni**
 v: 14 16 38
 3- go.away- **PLACT**
 ‘They left’, ‘(s)he/it left repeatedly’

Regarding the left boundaries, position 13 hosts a demonstrative pronoun *kyak* (and less frequently other pronouns) that indicates distancing from the speaker (either through the speaker’s or through the subject’s movement), but which functions also as a locative or a manner adverb ‘this/that way’. This is the first interruptor found to the left of the predicate head. The rest of the demonstratives in this paradigm occur most usually before negation in positions 5 and 6, but *kyak* is documented between the negation of position 10 (and presumably its enclitics of 11 and 12) and the verbal prefixes of position 14 when the verb is *-wo*, a light verb meaning ‘do’, ‘become, be’ (among many other meanings) and its derivatives. (The expressions with *-wo* ‘do, become, be’ might be somewhat lexicalized, but since the verb can bear TAME morphemes and inflect for person, I still consider it a verbal head in position 16 and not an auxiliary verb)

(11) Single interruptor *kyak* on the left

a. Je **kyak** i- yo- ø

v: 10 13 14 16 32
 NEG DEM 3- do- AP.DIST

‘It is not like that, it is not the same’.

b. Jlam’ne je **kyak** ji- won- ay- i [na- pø as- lejwa- s]

v: 1 10 13 14 16 20 32 45 - - - -
 but NEG DEM 2- do- 2PL- AP DEM- PL 2PL.POSS- fellow- PL

[ka Ø- en- a'yi ni syuni ti- jnajyi]

46 - - - - -

COMP 2.IRR- put- 2PL:AP DEM DEM 3- be.straight:AP.LOC

‘But you neglect justice with your fellows’ (Lit ‘you don’t do the same to your fellows when you make justice’).’ (Lc 11:42)

To the right of position 16 the closest free forms are some of the adverbs of 40, or perhaps the remote past of 39 if considered a free form - it can be uttered alone at least in metalinguistic uses, unlike applicatives and monosyllabic adpositions, and also fronted in other varieties of Chorote. Neither can these or any other free forms intervene between positions 13 and 40 (or 39), so the span defined by this diagnostic subtype is 14-39 (or -38).

The interruptor can also be defined as a construct that contains more than one free form, i.e. a *multiple free interruptor*. The interruptors of positions 13 and 40 (or 39) seen above are single forms, so they do not count in this version of the diagnostic; the same holds for negation (position 10), which is also a free form. The most typical multiple free interruptors are NPs (Tallman 2021). Recall that even light NPs may be multiple interruptors, since they can consist of a possessive construction with two Ns, like the one shown in (5b). In fact, these light NPs are the interruptors that function as the boundaries for this diagnostic. On the left side, light NPs appear in positions 5 and 6; although usually only one of these positions is filled, (12) illustrates the need to postulate two distinct positions for light pre-verbal NPs.

- (12) Multiple interruptor to the left. Two NPs as interruptors (Rom 4: 13)

[Si- nya] [ji- 'lij] i- wijnam [pa Abraham]
v: 5 - 6 - 14 16:33 45 -

1PL.POSS- father 3POSS- speech 3- give:AP.LOC DEM Abraham

‘God (lit. ‘our father’) gave his word to Abraham.’

To the right of the predicate head the first complex interruptor is the NP of position 42 exemplified in (13) and (14). (14) shows a construction with a complex light NP to the right of the verb, which is licensed by the irrealis nominal marker *-a* that follows the NP. The construction will be explained in more detail later in this section. (The NP means literally ‘son of boy’, but *jlas* is here a diminutive: ‘young/little boy’.) In the following position 43 there is also an NP (complement of the adposition of position 44) that can be complex. The example in (13) illustrates the need of postulating two different zones for positions 42 and 43:

- (13) Positions 42 and 43 (Drayson et al. 2000: 114)

Ø- *Laj* [i'nyát -a] [s- ate jl- as-] i'
 v: 14 16 42 - 43 - - - 44
 3- not_exist water -IRR 1PL.POSS -pitcher 3POSS- son -AP.LOC
 'There was no water in our (little) pitchers.'

- (14) Multiple interruptor to the right

a- wo [jwemik jl- as- a'].
 v: 14 16 42 - - -
 1- LV boy 3POSS- son- IRR
 'I was a young boy.' (Drayson et al. 2000: 122)

In sum, the multiple interruptors closest to the predicate head are in positions 6 on the left and 42 on the right. Thus, the span defined by this diagnostic is in principle 7-41. An example including both 7 and 7-41 is lacking; (15) shows the span 7-38, and 41 was shown in (9).

- (15) Multiple free interruptor: left edge in position 7

Kyak- p'an i- yo- ø- pi [pa i'nyó] [ti i- yo- ø
 v: 5- 7 14 16 32 38 45 - 46 - - -
 DEM- MIR 3- do- AP.DIST- REM.PST DEM man COMP 3.IRR- LV- AP.DIST
ka i- wo aye'wu- ye']
 - - - - -
 COMP 3- do shaman -IRR
 'That is what a man used to do when he wanted to become a shaman.'
 (Drayson et al. 2000: 134)

Note however that position 7 is occupied by the mirative and (less frequently) by the reportative. If these morphemes turned out to be second position clitics as discussed in §3.2, such that the pronoun *kyak* here occupied a clause-initial position rather than position 5, then the left boundary should be the first positively fixed element that follows the interruptor NP. This would give us the prospective particle *ja* of position 8.

Finally, the interruptor can also be defined as a *non-fixed* element. This subtype defines the subspan 14-22, exemplified in (16). In this example, the relevant part is in an embedded clause introduced by the complementizer *ti*. Therefore, the positions are given for the embedded clause only. In the main clause, enclosed in brackets, *sek yi'i* 'there is [the fact]' can be translated as 'then' and selects most usually for a verb with the momentary morpheme in position 22.

- (16) Non-fixed interruptor: subspan 14-22

[Sek y- i -'i] ti y- am- a'.
 v: [- - -] 4 14 16 22
 [DEM 3- be -AP.PUNCT] COMP 3- go.away- MOM
 'Then (lit. there it is) (s)he left.'

In this subtype, the left boundary is still the demonstrative pronoun *kyak* that occurs in position 13 among others, i.e. the same as in the *single free interruptor* version, but the right boundary is now the irrealis marker *-a* of 23. This morpheme occurs bound to non-verbal predicates and certain 'adjective-like' verbs (Class V in Table 2 below) indicating irrealis mood, as the embedded clause in (3b).; another example is shown (17a). In the remaining classes of verbs, irrealis is realized by means of a special set of personal prefixes in position 14. The suffix *-a* appears furthermore following nouns or light NPs in a predicative construction, like the one shown in (14). Another example of this can be seen in (15), whose relevant part is repeated and adapted in (17b).⁷

- (17)
- Irrealis*
- suffix/enclitic

- a. On the predicate head

Ka Ø- wujw- a'
 v: 4 14 16 23
 COMP 3- be.big- IRR
 'If it were big.'

- b. On the noun in constructions with light verb

i- yo [aye'wu- ye']
 v: 14 16 43 -
 3- LV shaman- IRR
 '(He) becomes a shaman.'

Thus, the span of interruptability by a non-fixed element is 14-22. An important issue, however, is that the momentary morpheme *-a* of position 22 has the same form as the irrealis suffix/enclitic, and they never co-occur. Attempts to

⁷The multiple meanings of this construction, consisting of a light verb *-wo/-yo* 'do, become, be' and an N(P) followed by the irrealis, depend on the N(P) involved: 'become', 'look for', 'build', 'use' or even 'be', among others (Carol 2015: 909–910). In these constructions the existence of the referent of the (N)P is not asserted, e.g. with negated existential verbs, existential verbs under conditional clauses, constructions with the meaning 'looking for' etc., or the referent comes into existence by the event itself, e.g. with verbs meaning 'build', 'become' and others; the meaning 'be' of (14) is the only exception.

elicit both together consistently failed, even in contexts where both are selected for.⁸ Clearly, they are two different abstract morphemes, since they can cooccur when irrealis is marked by a special set of personal prefixes, and that is the reason why two different positions have been assigned; besides, only the irrealis is non-fixed. But the choice in the linear order between them, in the absence of empirical evidence, has a purely theoretical motivation: aspectual morphemes are usually expected to be more internal than modal ones.

4.3 Non-permutability (14-23)

This diagnostic identifies a span whose elements cannot be variably ordered with respect to each other. In Chorote this defines the span 14-22, already exemplified in (16).

We have already discussed the pronoun *kyak*, which can appear both to the right of negation, in position 13, or to the left, in positions 5 or 6. Thus, *kyak* should be excluded from the span, since its position is interchangeable with that of negation.

As for the right edge of the span, the irrealis morpheme in position 23, exemplified in (17a), only occurs in that position in the verbal template and is therefore not interchangeable with any other element. Thus, I consider it the right edge in this diagnostic. It is true that it can occur to the right of the reportative (position 24), as in (17b), but in that case it belongs in the nominal template and is not relevant for the diagnostic. The reportative, in turn, can occur in a different position besides position 24, so it must be excluded from the span; see (18).

(18) Reportative occurring in position 7

<i>Sek-</i>	<i>jin</i>	<i>y- i-</i>	<i>'i-</i>	<i>pe</i>	<i>[syupi i'nyó' ji-</i>	<i>kyo</i>				
v: 5	7	14 16	34	39	45 - - -	-				
there REP 3- be- AP.PUNCT -REM.PST DEM man 3POSS- hand										
<i>t'isyé^(')n</i>	<i>y- i'lya-</i>	<i>je'</i>].								
-	- -	-								
3POSS:flesh 3- be.dry- AP.LOC										
'There was (hearsay) a man whose hand was dry.' (Mc 3:1)										

⁸These contexts are (a) the sequence *sekyi' i ti* 'there is (the fact) that...', approximately equivalent to 'then', which selects for the momentary morpheme on the verb and (b) the complementizer *ka*, which selects for irrealis. When (a) is in the prospective form, it includes *ka* instead of *ti*: *sek jane'yi ka...*, so that both contexts co-occur. In that case, when the verb marks irrealis with the irrealis set of personal prefixes, both irrealis and momentary co-occur. But when the predicate is non-verbal or a Class V verb, such that irrealis need to be indicated with *-a*, only one *-a* surfaces.

The diagnostic as described above corresponds to *strict* non-permutability. A possible fracture discussed in the introduction of this book and in Tallman (2021) considers scope: *flexible* non-permutability admits inside the span elements whose order can be altered if that entails changes in scope. However, I have not considered this version of the diagnostic for Chorote due to lack of reliable data on scope changes when order is altered. In §3.2 it was shown that the TAME markers of positions 24 through 28 can be fronted to positions 1, 7-9, and 11-12. In some cases, these morphemes occur bound to nominals and have nominal scope, as discussed in Carol (2015), and thus do not occupy a position in the verbal template; but in others, as those of §3.2, it is not clear to me whether the variation in position affects scope in the verbal template.

Despite that, it is worth considering what happens to the right of those TAME markers since, were it confirmed that their variable ordering is sensitive to scope, we would have results for the diagnostic of flexible permutability.

After the TAME markers comes third person plural marker *-is*, which only occurs in position 29, and thereafter the string of Ps functioning as applicatives, which are rigidly ordered in positions 30 through 34 if no complement (NP or pronoun) surfaces. But if an oblique pronominal complement surfaces in position 35, then the P which selects the pronoun as its complement occurs to the right of the latter in position 36, which may alter its relative order regarding other Ps.⁹ In 19a the P called here “orientation” appears to the left of the instrumental, but in 19b the former takes a first person pronominal complement and surfaces thus to the right of the instrumental.¹⁰ This does not relate to semantic scope; hence, position 30 is outside the span of a flexible non-permutability (under the assumption that all TAME markers are potentially able to display scopal variation).

(19) Permutability of Ps

a. Orientation before instrumental in regular order

<i>i-</i>	<i>nyu-</i>	<i>yej-</i>	<i>e</i>
v: 14	16	30	31
3- pass- AP.OR- AP.INS			
‘(S)he helps him/her.’			

⁹Recall that I assume that the Ps always belong to the verbal template, but they can surface as applicatives or adpositions depending on the “weight” of the NP. If an alternative analysis were adopted, according to which the Ps surfacing as adpositions belonged to a distinct template other than the verbal template, there would be no permutability of Ps.

¹⁰The basic allomorphs are *-a(j)* for “orientation” and *-e(j)* for the instrumental. After a vowel, epenthetic /y/ is inserted and included as part of the Ps. This /y/ in turn raises front vowels, as explained in §2, thus the Ps result in *-yej*, *-yij*, respectively. The P “orientation” takes here a suppletive form when the pronominal complement surfaces.

- b. Instrumental to the right when it takes a non-null complement

i- nyu- yij- k'i- 'm
 v: 14 16 31 35 36
 3- pass- AP.INS- 1SG- AP.OR
 '(S)he helps me.'

4.4 Subspan repetition (8-38; 8-39)

For this diagnostic, I consider repetition of a subspan of contiguous positions in a construction that is functionally equivalent to 'and' coordination. In fact, the construction in question probably involves subordination: it is *jla'yi ti/ka*, where *jla'yi* means literally 'his/her/its fellow' and *ti/ka* are the complementizers that select for *realis* and *irrealis* predicate, respectively.¹¹ If we consider elements that *must* be repeated in order to be interpreted, i.e. the *minimal* subtype of this diagnostic, the subspan is 8-38. If we consider the elements that *can* be repeated, i.e. the *maximal* version, the span is 8-39. In (20) we can see evidence for the left edge: failing to repeat the prospective marker *ja* gives an ungrammatical result. The elements to the left of *ja* cannot be repeated.

- (20) Prospective *ja* repeated under coordination

Ja-	'yit-	aj-	a-	'a	[na si-	'lij]	jla'yi
v: 8	14:16	19	31	34	45 -	-	1

PROSP- 1:stab?- 1PL- AP.INS- AP.PUNCT DEM 1PL.POSS- language and

ti ja-	y- amti-	jyen-	a	/*y- amti-	jyen-	a
4 8	14 16	18	19	/*14 16	18	19

COMP PROSP- 1- speak- CAUS- 1PL/ *1- speak- CAUS- 1PL

'Let us write our language and read it.' (From an educational book, Drayson 1999)

In (21a) we can see the repetition of the remote past *pe(j)*, but not in (21b), even though the remote past is interpreted in the second conjunct; both examples come from the same text and presumably the same speaker. This can be taken as evidence that repetition of *pe(j)* is optional.

¹¹Coordination of NPs also involves *jla'yi* (or the feminine *jla'yiki*), but with determiners instead of complementizers. For 'or' coordinating VPs or clauses, Chorote uses *ni'ne*, with the same complementizers; *ni'ne* could be translated in isolation as 'perhaps, maybe'. The behavior of *ni'ne ti/ka* is apparently similar to that of *jla'yi ti*, but less examples were found.

(21) Remote past *pe(j)* under coordination

a. Repeated

[*Naka ni* *ø paj-* *k'i* *ti* *a-pe'ya-k,* *si-*
 v: [- - - - - - - - - -
 [DEM DEM 3- be.ancient- AP.DISTR COMP 1- hear- 1PL 1PL.POSS-
tyet-e i- 'wi'in,] *a- 'yen-* *a- 'nij-* ***pe*** *jla'yi ti* *a- kyes-*
 - - - -] 14 16 19 38 39 1 4 14 16-
 eye- PL 3- see] 1- look- 1PL- PLACT- REM.PST and COMP 1- touch-
a- 'a- ***pe***
 19- 34- 39
 1PL- AP.PUNCT- REM.PST
 '[That what was at the beginning (lit. 'long ago'), what our eyes saw,]
 what we observed and touched...' (1 John 1:1)

b. Not repeated

[...*a- wo- k-* *i* *s-* *amtiky-e-* *'as-* *e* *naka*
 v: [...- - - - - - - - - -
 [1- LV- 1PL- AP.DIST 1PL.POSS- speech- IRR- 2PL- AP.DIST DEM
syunye] *a- 'wen-* *a-* ***pe*** *jla'yi ti* *a-pe'ya-k.*
 -] 14 16 19 39 1 4 14 16 19
 DEM] 1- see- 1PL- REM.PST and COMP 1- hear- 1PL
 '[We tell you what] we saw and heard.' (1 John 1:3)

Moreover, this is consistent with the general behavior of *pe(j)*, which is “optional” in the sense that it does not need to follow the predicate when the information it provides can be recovered e.g. from previous discourse. The pluractional *'ni(j)* is *not* repeated in the second member of the coordination of (21a) and thus not interpreted (*'ni(j)* in combination with *'yen* ‘watch, look at’ gives literally ‘watch repeatedly’, i.e. ‘observe’.) This is a clear indication of its difference with respect to *pe(j)*.

4.5 Deviation from biuniqueness (14-18; 16-18; 14-29)

This type identifies deviations from the biuniqueness relation between form and meaning, which might be used as indication of word boundaries. The most common case in Chorote is more than one form corresponding to a single meaning. In this regard, the personal prefixes of position 14 are the left edge. There are five different sets for the third person. They are predictable in some cases — transitive verbs always select for Class I *i-/y-*, antipassive verbs for Class III *t-/tV-*, and there

is some correlation between semantic classes and prefixes, but the latter are not fully predictable in intransitive non-antipassive verbs (Carol 2013, 2014b). To the left of the prefixes of 14 no such deviation is ever found.

Table 2: Third person verbal prefixes

	Realis /_C (_k) ~ _V	Irrealis /_C ~ _V	Goes with...
Class I	<i>i-</i> (<i>ya-</i>) ~ <i>y-</i>	<i>in-</i>	transitive, active and inactive
Class II	\emptyset -	<i>in-</i>	active and some inactive
Class III	<i>ti-</i> (<i>ta-</i>)- ~ <i>t-</i>	<i>in-</i> ~ <i>int-</i>	active and a few inactive
Class IV	<i>in-</i> ~ <i>n-</i>	<i>in-</i>	inactive
Class V	\emptyset -	\emptyset -...- <i>a</i>	inactive

The rightmost position where such deviation can be found is 18, which hosts antipassive and causative suffixes. Leaving aside the indirect causative *-jan/-yin*, whose allomorphy is limited and predictable, the antipassive and especially the direct causative suffixes display a strong allomorphy which cannot be predicted on phonological or semantic grounds, see (22)-(23); for simplicity, with causatives only bases ending in vowel are shown. Sometimes the same verbal base is acceptable with two distinct allomorphs, as can be seen with *-po-yi* ‘be full of’ in (23b) and (23c) (again, the epenthetic *y* inserted between vowels is analyzed as part of the suffix, so in (23c) *-it* becomes *-yit*).

(22) Allomorphs of the antipassive suffix

- a. Regular antipassive with *-jan*, verb *-lej* ‘wash’

ta- ka le- ja’n

v: 14 15 16 18

3- ANTIP wash- ANTIP

‘(S)he does the washing’

- b. Irregular antipassive with *-n*, verb *-jlu* ‘send’¹²

ta- ka jlɔ- n- i

v: 14 15 16 18 31

3- ANTIP send- ANTIP- AP.INS

‘(S)he sends.’

¹²Here the demoted object is reintroduced by the instrumental applicative.

- c. Irregular antipassive with *-ki*, verb *-lan* ‘kill’

*ta- ka- lan- ki*¹³

v: 14 15 16 18

3- ANTIP- kill- ANTIP

‘(S)he kills.’

- d. Irregular antipassive with no suffix, verb *-sinyan* ‘grill’

ta- ka sinyan

v: 14 15 16

3- ANTIP grill

‘(S)he makes a barbecue.’

(23) Allomorphs of the direct causative suffix

- a. Suffix *-jat*, verb *-nu* ‘pass by’

i- nyu- jat

v: 14 16 18

3- pass_by- CAUS

‘(S)he makes [someone] pass by.’

- b. Suffix *-nit*, verb *-po-yi* ‘be full of’

i- pyo- nit- i

v: 14 16 18 31

3- be.full- CAUS- AP.INS

‘(S)he fills with it.’

- c. Suffix *-it*, verb *-po-yi* ‘be full of’

i- pyo- yit- i

v: 14 16 18 31

3- be.full- CAUS- AP.INS

‘(S)he fills with.’

- d. Suffix *-jVnit*, verb *-pu* ‘exist’

i- pyu- junit

v: 14 16 18

3- exist- CAUS

‘(S)he creates, brings into existence, imports.’

¹³ An alternative analysis, perhaps historically more accurate, consists in splitting *-ki* into *-k*, the participle of position 17, and *-i(y)*, a verbalizer in position 18 that creates denominal verbs. In any case, the right edge of the domain under discussion would still be position 18.

- e. Suffix *-t*, verb *-’uy* ‘enter, get in(to)’

’*yu-* *t*

v: 14:16 18

3:be.full- CAUS

‘(S)he puts (it) in (e.g. a pocket).’

No deviations of this kind occur beyond position 18. In position 19 there appear the concord suffixes of 1PL. Although this morpheme has at least three allomorphs, their distribution is phonologically conditioned: *-Vk* after *-j*, *-k* after *V* and *-a(j)* elsewhere, e.g. *alej-ek* ‘we wash’, *awo-k* ‘we fish’ and *a’wen-a(j)* ‘we see’, respectively.

The exponence of irrealis could also be seen as a case of deviation from biuniqueness, since it is realized both through a suffix *-a* and a distinct set of personal prefixes. However, it is of a different kind, because the different forms appear in different positions - 14 and 23. Besides, the occurrence of one or the other exponent is fully predictable on categorial grounds (see (17) and text above). In sum, the diagnostic analyzed up to now is inapplicable to irrealis exponence, which is better treated as a distinct morphological category rather than as a question of allomorphy.

The previous account describes a language-specific fracture of the diagnostic, called *inflectional class* in Table 4 (§7) because its left boundary includes prefixes that define the inflectional class of the verb. An alternative fracture is possible, where only “derivational” morphology is included; this is referred to as *fossilization* in Table 4. In this case, the right edge is still position 18, but the inflectional personal prefixes of position 14 would fall outside the span. The antipassive morpheme of position 15 too, but for a different reason: even if considered derivational, its allomorphy seems to be predictable. The basic allomorph is *ka*, as shown in (22); if the base begins with a glottal stop followed by a stressed vowel, an epenthetic *n* is inserted, which fuses with the glottal in *’n*, cf. *taka’neyasan* ‘teach (intransitive)’ < *ta-kan-’éya*san*, cf. *’yiyas* ‘(s)he teaches [someone]’ < *y-’éya*san*, while with bases beginning with a vowel its form is *k-* e.g. *ta-k-ámtijnye’n* (14 15 16:18) ‘(s)he talks’, cf. the transitive *y-amti-’ni* (14 16 38) ‘(s)he talks [about someone]’. Thus, the left edge is the predicate head of position 16, and the span is defined as 16-18.

There is a third fracture which considers *extended exponence*. This phenomenon is seen in the personal prefixes of position 14 on the left, and on the verbal third person plural marker *-is* in position 29 on the right, both of which are exponents of third person; *-is* cross-references third person subject of transitives and also intransitives with an oblique introduced by an adposition or applicative, see (24).

Notice there are other cases of extended exponence between these edges: the antipassive morpheme *ka* in position 15 may determine changes in the root, as in 22c, where the root becomes deaccented, as well as the verbal plural markers of positions 19 and 20, which are specific to first and second person, respectively, and thus show extended exponence of these features.

- (24) Extended exponence - third person: subspan 14-29

<i>y-</i>	<i>am-</i>	<i>is-</i>	<i>i</i>	(<i>ja-</i>	<i>pa</i>	<i>jlɔsyɛ</i>)
v: 14	16	29	31	(45	45	45)
3-	go.away-	3PL-	AP.INS	f-	DEM	girl

‘They take (the girl) away.’

5 Phonological diagnostics

5.1 Accent (15-16; 16-34; 4-40)

Three subtypes are considered under this rubric. *Accent minimal-minimal* is defined as the minimal span containing the positions where the accent can appear in utterances with only one accent. In such cases, the accent falls almost always on the verbal root or non-verbal predicate of position 16, but in fossilized, irregular antipassives, it falls on the antipassive morpheme in 15, as in (22c). Thus the span is 15-16.

The *minimal-maximal* subtype considers the maximal span where no position other than the predicate head, that is, position 16, can bear stress. Since position 15 can bear the stress alongside with position 16 in regular and some irregular antipassives, as well as in reflexives, (cf. in *i-ní ’wé’en* (14-15 16) ‘(s)he sees himself/herself’) then it should be excluded from the span and the left edge should be placed in position 16. As for the right edge, the oblique second person pronoun *-a* of position 35 bears secondary stress and, optionally, another main stress, as in (25).¹⁴ Thus the *minimal-maximal* span is 16-34.

- (25) *Accent minimal-maximal*: stressed pronoun of position 35 as right boundary

<i>si-</i>	<i>tyánt’ya</i>	<i>á-</i>	<i>j</i>
v: 14	16	35	36
1SG-	know	2-	AP.INS

‘I know you (sg.)’

¹⁴This oblique pronoun is usually written together with the previous element in Chorote texts. However, according to the convention adopted in this chapter – no more than one stress per orthographic word – it is written separately (see §3.1). The second person plural oblique marker *-(’)as* could arguably be segmented as *-(’)a-s*, where *-s* would be the plural also found in nouns after a vowel. For the initial glottal stop, see below in this section.

The *maximal-maximal* subtype includes the longest possible span with only one accent. Since positions 15 and 35 are not necessarily present in a word, they do not define a boundary in this subtype. The *accent maximal-maximal* subspan is 4-40, already exemplified in (8a), repeated here as (26); see also (8b) and (9) for examples of the edges in spans occurring in main clauses.

(26) Accent *maximal-maximal*: subspan 4-40

ka	Ø-	'nes-	ta-	na'a
v: 4	14	16	27	40
COMP 3.IRR- arrive- INCOMP- later				
'When (s)he/it arrive.'				

Notice that other particles in position 40 are stressed, e.g. *tá'a* 'immediately; already', *pet* 'please' and others, as well as the locative adverbs of position 39 *ts'ijí*, *'niji* and others, as well as all the elements occurring to their right.

None of the fractures proposed for this test gives position 14 as a left limit, which is the position with the highest number of convergences for the left edge. This is because position 15 is stressed. Otherwise, the left edge for the minimal-maximal subtype would be position 14. In fact, position 14 is a left edge in the minimal-maximal subtype, but of a domain that excludes the verb core of position 16 and includes only the verbal prefix of position 14 and the reflexive or the antipassive of position 15. In the following section we will see another potential diagnostic which points to position 14 as a left edge.

5.2 Another potential diagnostic related to accent (14-)

There is another potential phonological diagnostic which has not been included in Table 3. For other languages of the family, namely Wichí and Nivaclé, an iambic rhythmic type has been proposed (Nercesian 2014, Gutiérrez 2015, Gutiérrez 2016). This can be clearly seen in the following alternation in 'Weenhayek Wichí, where long vowels (written as geminates) regularly correspond to stressed vowels in Chorote: *qasiit* 'stand (up)' (imperative) vs. *ta-qaasit* '(s)he/it stands' (Claesson 2016). In Nikulin & Carol (Forthcoming) it is argued that this may have been the default stress pattern in Proto-Mataguayan, activated when no underlying accent is present in a three-mora window at the left margin of the "word" (i.e. of the stress domain). If applied to Chorote, and if the iambic structure should be aligned with the left edge of the verb word, this would suggest a span in which position 14 constitutes the left edge, since it is the leftmost possible initial syllable of a iamb containing the verbal root; see (27), where stress is indicated with an acute accent for the sake of clarity:

(27) Iambic structure of the verb

ta- kásit

v: 14 16

3- stand

‘(S)he stands (up).’

In Chorote, a default left-aligned iambic type would also explain why most non-possessed nouns bear the stress in the second syllable of the stem, e.g. *ajwén-ta* ‘chicken’, but possessed nouns in the first one when the stem begins with a consonant and the possessive prefix has the form (C)V-, e.g. *i-pyúsi* ‘my beard’: in either case, the iambic structure is preserved. There are still some cases of alternation in adpositions that take possessives to indicate the complement: *kyajwé* ‘under, in the lower part of’ versus *ji-kwáje* ‘under it, in its lower part’. This would also explain the rare cases where the stress does not fall in the first vowel of a verbal stem, e.g. *’najwél* ‘(s)he is shy/ashamed’, explained as /n+’ahwél/ (3-be.shy), where the prefix is C-, rather than (C)V, and thus does not add a syllable.¹⁵

Nevertheless, there is no clear evidence that this iambic pattern is still synchronically productive in Chorote. In verbal stems, the position of the stress is fixed: contrast the *’Weenhayek* example above with Chorote *kásit* ‘stand up’ (imperative), *ta-kásit* ‘(s)he/it stands (up)’. In any case, it is an indication of a metrical criterion for defining the left edge of the phonological word which might have been productive for a long time in the (pre)history of the language.

5.3 Insertion of /y/ between vowels (16-32, 16-34, 16-44)

When two vowels are in hiatus across a base-suffix (or -enclitic) border, epenthetic /y/ is inserted. With only one documented exception,¹⁶ this applies consistently to any suffix/enclitic element of the predicate up to the distal applicative -e(y) of position 32, see 28a; recall that morphemes that can surface both as applicatives or adpositions, like the distal, are glossed here as ‘P’ in either case. The /y/ is considered to be part of the suffix/enclitic in the template, and is not

¹⁵Under this hypothesis, cases like *y-imi’n* [yími’n] (3-love) ‘(s)he loves’, where the iambic structure is not preserved, are explained by assuming an underlying long vowel in the first syllable of the stem in the proto-language. This is not mere speculation, but what actually happens in present-day *’Weenhayek Wichí*, cf. *ya-huumin* (3-love) ‘(s)he loves’. In sum, the accent in the proto-language (whatever its nature was) would have fallen in the second *mora*.

¹⁶The exception occurs between the light verb -wo ‘do, become, be’ and the distal -ey, where no /y/ is inserted; instead, the allomorph -y of the distal is selected, resulting in the form -wo-y. An example of this is shown in (32).

segmented as a different morph elsewhere in this chapter unless explicitly indicated. In positions 33 and 34 the Ps begin with a consonant, e.g. -'e or -k'i, so there is no context for insertion; however, as these elements belong to the same class as the distal, it seems reasonable to consider some version in which the right edge extends to position 34. To the right of 34, the only case of /y/ insertion occurs between the P of position 44 (acting there as postposition) and its host, the (N)P of position 43, see 28b.. We take thus that the right edge is position 32 in the *minimal-minimal* version, 34 in the *maximal-minimal* version, and 44 in the *maximal-maximal* version of the /y/ insertion diagnostic. In the latter case I assume that the NP is part of the verbal domain; see §3.3 on the distribution of applicatives/adpositions and NPs. Since /y/ insertion does not occur between a prefix (or proclitic) and a base, the left boundary has to be position 16.

(28) Insertion of epenthetic /y/

- a. *Minimal-minimal* - between the applicative of position 32 and its host: span 14-32

/a- ho- ey/ → o- jo- y- i
v: 14 16 32 → 14 16 32 32
1- go- AP.DIST → 1- go- EPEN- AP.DIST
'I went there.'

- b. *Maximal-maximal* - between the postposition of 44 and its host: span 14-44

o- jo [']Iwit'osi- y- i /*o- jo- y- iwit'osi- y-
v: 14 16 43 44 44
1- go Tartagal- EPEN- AP.DIST /1- go- EPEN- Tartagal- EPEN-
i
AP.DIST
'I went to Tartagal.'

Now let us address some analytical issues that deserve consideration. To the right of position 44 all elements begin with a consonant, so what remains to be considered is only what happens between positions 34 and 44.

Before an underlying initial vowel, Chorote regularly inserts a glottal stop whenever no /y/ is inserted. This can be seen before the initial vowel of *Iwit'osi* '(the city of) Tartagal' in (28b)¹⁷. We do not expect /y/ insertion there, since

¹⁷Actually, in normal/fast speech the first vowel of the N is assimilated to the preceding vowel across the glottal stop, so *Iwit'osi* 'Tartagal' becomes ['owit'osi]. Vowel assimilation across laryngeals is a regular process in Chorote.

Iwit'osi is not a suffix or enclitic. We also find that a glottal stop, rather than /y/, is inserted at the beginning of a P in position 44 when it is a polysyllabic adposition, see (29) (the relevant inserted glottal stop is added between brackets, since it is not written in Chorote spelling). Contrast this with the insertion of /y/ before position 44 when the P is monosyllabic, as in (28b).

(29) Glottal stop insertion before polysyllabic P in position 44

a. Following an N(P)

y- i a'lénta [']**apé'e**
 v: 14 16 43 44
 3- be horse **AP.over**
 '(S)he is on the horse.'

b. Not following an N(P)

i- jyo [']**apé'e** [jlaják tikíjnaki']
 v: 14 16 44 45 -
 3- go **AP.over** DEM mountain
 '(S)he climbed that mountain.'

I take this, together with the fact that polysyllabic adpositions can bear stress (unlike monosyllabic ones), as an indication that they make up a stress projecting domain, so they are not phonologically bound elements.¹⁸ Thus, /y/ insertion occurs *inside* this stress domain, i.e. between bound elements, and glottal stop insertion *between* these domains.

There are still two places between positions 34 and 41 where *prima facie* a morpheme-initial vowel may occur: in position 35, with the second person oblique pronoun *-a* (singular), *-as* (plural) seen in (25) and repeated here as (30a), and position 39, with the perfect *-Vje(j)/-V...je(j)*, seen in (30b). In both cases a glottal stop occurs.

(30) Inserted or underlying glottal stop?

a. Before oblique second person marker

si- tyant'ya [']**a-** (j)
 v: 14 16 35 36
 1- know 2- **AP.INS**
 'I know you (sg.).'

¹⁸ As one of the editors points out, a question for future work arises here - whether adjacent stress domains form a larger prosodic domain. Perhaps some stressed syllables are stronger than others, forming a larger prosodic domain. If so, adpositions might be candidates for 'weaker' syllables.

b. Before the perfect marker

y- <i>am-</i>	a-	' <i>aja</i>
v: 14 16	22	39
3- go.away-	MOM-	PRF
'(S)he/it left again.'		

It is difficult to determine whether the initial glottal stop in (30) is inserted or underlying. In (30a) there is some evidence to consider it underlying, i.e. /'a*/, not /a*/. The evidence for treating the glottal stop here as underlying is as follows; (i) the glottal stop here triggers glottalization of a preceding /s/ into /ts'/, like an underlying glottal stop, and unlike an epenthetic glottal stop: e.g. 'es + 'asé(j) → ests'yase 'it is good for you (pl.)' (with further palatalization due to the previous vowel, see below), and (ii) the glottal stop here labializes after a rounded vowel, like an underlying glottal stop, and unlike an epenthetic glottal stop: i.e. /'/ → /'w/, e.g. ijoyo + 'asé(j) → ijoyo'wasé(j) '(s)he goes to you'. These two processes are not documented in my material with epenthetic glottal stops, but they are with the locative P -'e of position 33, whose initial glottal stop is underlying beyond any doubt, e.g. yiyis + 'e → yiyits'i' '(they) are in...'; 'yu + 'e → 'yu'we' 'it fits in...'.¹⁹

Alternatively, one could consider that the morphemes of (30) belong to a distinct accent projecting domain, like polysyllabic adpositions, and hence glottal stop instead of /y/ is inserted there. There is some historical evidence for this.²⁰ In sum, even if there is no conclusive evidence for glottal stop insertion between positions 34 and 41, there is no evidence at all for /y/ insertion, so the right boundaries for the three versions of this diagnostic hold as determined above, namely as positions 32, 34, and 44.

As for the left edge, since /y/ insertion applies to suffixes only, the left edge can only be the predicate head itself; (31) shows that /y/ insertion does not apply in prefix-base boundary (/y/ is segmented there as distinct morpheme for the sake of clarity). Note that in what follows the first line contains a surface form, and the second line an underlying form to which the processes under discussion apply.

There is no evidence of /y/ being inserted between two positions to the left of the predicate (although it is inserted *inside* positions, e.g. when the position

¹⁹The underlying character of the glottal stop in -'e can be seen in the contrast with the distal and other Ps with initial subjacent vowel, which expectedly take /y/, while -'e does not: yi+ -'e → yi'i' '(s)he/it is at (a precise place)' vs. yi+ -ey → yiyi '(s)he/it arrives (at a distant place)'.

²⁰The Wichí cognates for this second person marker show a long vowel (Claesson 2016), i.e. a bimoraic morpheme, like polysyllabic adpositions.

contains an N(P), such as in positions 1 through 3). Therefore, in any version of this diagnostic the left edge is position 16.

(31) No /y/ insertion without suffixation/encliticization

a. Before the personal prefixes

ø- *ɛmi'n* / *a- *y-* *imi'n*
/a- *imin*/

v: 14 16

1- love / *1- EPEN- love

'I love it/him/her.'

b. Before the reflexive-reciprocal base

a- *nín-* *ɛmi'n* / *a- *ni* *y-* *imi'n*
/a- *ni* *imin*/

v: 14 15 16

1- REFL love / *1- REFL- EPEN- love

'I love myself.'

This diagnostic has two serious limitations: (i) it cannot be used to define the left edge of a word, since it applies to suffixes/enclitics only, and (ii) it cannot be applied to any of the positions where only items beginning with a consonant exist.

However, an alternative formulation might be interesting. In effect, notice that in (31b), instead of the glottal stop, an epenthetic *n* is inserted between the final vowel of the reflexive/reciprocal of position 15 and the initial one of the verb. Therefore, if the diagnostic were formulated as “non glottal stop insertion”, rather than “/y/ insertion”, position 15 should be added to the span as its left edge.

5.4 Palatalization (14-16; 14-40; 14-18/25; 14-40; 14-16; 14-46)

Underlying /i, y/ palatalize all consonants, while epenthetic [i] and underlying /u/ palatalize only coronals. The former is referred to as “first palatalization” and the latter as “second palatalization”. Thus, for instance, prefixes of the form *i-* (possessive and irrealis active first person, *realis* active third person) palatalize any consonant because *i* is underlying there (first palatalization), but prefixes of the form *Ci* (nominal and verbal) palatalize coronals only, because the *i* in such cases is not underlying but derived (second palatalization) (Carol 2014b).

Palatalization usually means $C \rightarrow Cy$, but also $w/ \rightarrow /y/$ before rounded vowels, $/ky/ \rightarrow /sy/$, and $/k'y/ \rightarrow /ts'y/$ (notice that $/k^{(')} /$ and $/k^{(')}y/$ are distinct phonemes; surface $k^{(')}i$ reflects subjacent $/ky^{(')}i/$). Neither the context nor the process itself

are always transparent for two reasons: (i) /y/ is regularly dropped in coda after triggering palatalization, and (ii) Cy causes raising of a following *e* into *i*, among other vowel changes, and thus /Cye/ appears superficially as *Ci*; see (32).

- (32) Palatalization of /w/, deletion of /y/ in coda and /e/ → /i/ after palatal
- | | | | |
|-----------|------------|---------|-----------|
| <i>i-</i> | <i>yo-</i> | ∅ | <i>pi</i> |
| /i- | wo- | y | peh/ |
| v: 14 | 16 | 32 | 39 |
| 3- | LV- | AP.DIST | REM.PST |
- ‘(S)he said/wanted/did it.’

A number of domains can be identified based on palatalization phenomena in Iyojwa’aja’ Chorote. On the one hand, diagnostics can be subdivided into *minimal* and *maximal*: a set of *minimal* subtypes that define contiguous sub-spans of positions that trigger and/or undergo palatalization whenever the relevant context exists, and a set of *maximal* ones that define the largest possible span where *all* the occupied positions trigger or undergo palatalization (in other words, a span outside of which no palatalization is known to occur inside the verbal template). On the other hand, the diagnostics can be classified according to the target and environment of palatalization. We can consider A) the “first” palatalization as a whole, B) the “first” palatalization excluding that of /k^(ʰ)y/, which is somewhat exceptional, and C) palatalization of coronals only, regardless of whether they are affected by the first or the second palatalization rule. In all, six different domains arise; see Table 3 for a summary of these tests.

Table 3: Palatalization diagnostics

Subtype	Specific fracture	Left edge	Right edge
Minimal-A	With k ^(ʰ) y	14	16
Maximal-A	With k ^(ʰ) y	14	40
Minimal-B	Without k ^(ʰ) y	14	18/25
Maximal-B	Without k ^(ʰ) y	14	40
Minimal-C	Coronals only	14	16
Maximal-C	Coronals only	14	46

The need for postulating the subtype B is based on the absence of palatalization of *-k* in positions 17 and 19, which is unexpected, because no other diagnostic places boundaries there. The absence of palatalization in position 17 is especially surprising if we consider that the position corresponds to ‘derivational’ morphology. Moreover, the *j*-initial causatives in position 18 do show palatalization: /i-limi-hat/ → *i-limi-jyet* (3-be.white-CAUS) ‘make white’. These data suggest that it is reasonable to set aside the palatalization of /k^(ʰ)y/ as a special case, thus justifying the B subtype of the diagnostic.

For any version of this diagnostic the left edge is position 14, which hosts prefixes (or proclitics) that trigger palatalization (see e.g. (32)) but do not undergo palatalization of any kind, e.g. *t-amti* ‘(s)he speaks’ is never realized as **ty-amti*, not even when preceded by /i, y/. In the *minimal* subtypes, the contiguous subspan is thus necessarily interrupted in position 14.

It is true that there is no direct evidence that palatalization cannot operate between positions 13 and 14, because the pronouns of position 13 never end in /i/ or /y/; hence, they provide no context for palatalization. But neither other NPs nor any other material ending in /i/ or /y/ trigger palatalization of the prefixes of position 14, so there is no reason to suppose that an eventual pronoun in position 13 ending in /i, y/ would. Furthermore, since position 13 is occupied by a tonic pronoun which is arguably an NP and can occupy other positions as well, I find no reason to suspect that the relationship between positions 13 and 14 should be any different from that between position 14 and any other position to its left.

As for the *maximal* subtypes, position 14 is also the left edge for verbal non-imperative predicates, where position 14 is obligatory. In reflexive-reciprocal and antipassive verbs in imperative mode position 14 is empty but position 15 is occupied by the reflexive/reciprocal *ni* and the antipassive *ka*.²¹ If these morphemes were shown to be palatalized by some element to the left of position 14, the position of that element would be considered to be the left edge. However, there is no way to prove this, since *ka* and *ni* can never be target of palatalization. The regular antipassive marker *ka* does not palatalize because /k/ (unlike /ky/) never palatalizes in the variety under consideration,²² while the reflexive/reciprocal

²¹Of course, it could be argued that the position of the personal prefixes is not actually empty in imperatives, but occupied by an abstract morpheme with zero exponence. Some indirect support for this is found in the Manjúi variety, where an optional *a*-second person prefix for imperative exists. In any case, this implies no changes for the diagnostics.

²²In the Montaraz varieties of Argentina and Paraguay (known as Iyo’awújwa’ and Manjúi respectively), in contrast, /k/ palatalizes, thus the antipassive becomes *kya* after prefixes of the form *i-* (Gerzenstein 1983; Carol 2018). However, even in these varieties *ka* is not palatalized by elements to the left of position 14.

morpheme *ni* should be analyzed as underlying /yne/ or /yni/, and thus palatalization cannot be applied.²³ In non-verbal predicates, where position 15 is also empty, the left edge could be pushed further to the left if the non-verbal predicate of position 16 could undergo palatalization. But this is not the case, hence position 14 remains as the left edge.

What remains to be considered are the right edges. Let us address the A-set first. The *minimal-A* palatalization is limited by the participle /-ky/ of position 17, which does not undergo palatalization, as can be seen in (33), so the span is 14-16. I found no clear examples of a context of palatalization for the participle in the verbal domain, hence (33) comes from the nominal domain; recall also that palatalized phones neutralize with plain ones in coda position, so only examples with onsets are shown.²⁴ (In (33) /Cye/ → Ci; if we had /k/ instead of /ky/ we would expect **amtike*.)

- (33) Lack of palatalization in the participle -ky of position 17.

y- *amti-* **ky-** e'
1SG.POSS- speak- PP- IRR
'my speech (irrealis).'

The *maximal-A* version reaches the TAM particles of position 40, which undergo first palatalization: *pet* 'please'²⁵ → *pit*, *kyu* 'a while' → *syu*, -*na'a* 'later' → -*nye'e*, etc. The subspan is thus 14-40, see examples of the right edge in (34).

- (34) Palatalization *Maximal-A*: right edge in position 40

a. *jwel-* *i* ***syu'***
 /hwel- ey kyu/
v: 16 32 **40**
 tell- AP.DIST a.while
 'Tell him/her.'

²³We can infer this from the following: *ni* is stressed, and stressed *i* only surfaces after a palatal(ized) phone (underlying /i/ is otherwise *e*, a closed mid or very open high vowel). Thus, the previous /n/ must be palatalized, which in turn supposes a previous /y/ which regularly falls in coda.

In the Montaraz varieties the reflexive/reciprocal is *wet* and can be palatalized by a personal prefix of the form /i-/ in *wit* or *yit*, depending on the variety. However, it has not been documented that elements to the left of position 14 are able to palatalize *wet*.

²⁴That there is a context for palatalization in (33) can be seen in the contrast with *t-amti-ts'i-ji'n* 'they speak', with the same root, where -*ts'i* is the palatalized allomorph of the distributive -*k'i* of position 34.

²⁵The translation 'please' (*por favor*) was suggested to me with imperatives. In other cases, however, it is very difficult to find an equivalence. It seems to indicate a benefit for some participant.

- b. *kyak iyo- Ø pit*
 /i- wo- y pet/
 v: 11 14 16 32
 DEM 3- LV- AP.DIST
 ‘This is the way it is.’

The locatives *ts’ijí*, *’nijí* of position 41 do not provide a target for palatalization. The NPs/DPs to the right of position 41 do, but first palatalization is not documented there. First palatalization does occur in the postpositions in position 44, but is always triggered by their complement in position 43 which in turn does not palatalize, so it cannot define a maximal subspan.

Subtype B ignores palatalization of /k^(‘)/, which fails to occur not only in position 17 but also in positions 19 and 35, see (35).²⁶

That *-k’i* is in a palatalization context in 35b. can be seen in the contrast with *’yen-a-jyi’n* (look-2pl-*ĴEN*) ‘watch’, where *-jen#* → *-jyi’n* after the second person plural marker which is underlyingly /ay/; for 35c. it can be seen in the contrast with *ti jna-jyi’* (3-be.straight-P_{LOC}) ‘it goes straight’, where *-ji#* → *-jyi’* after the root which is underlyingly /hnay/. Notice that if palatalization took place in 35c. we would see its traces in the regular *e* → *i* after a palatal and for 35a.

(35) Lack of palatalization of k^(‘)y

- a. Concord first person plural: position 19
Ø- amti- k- i (- s- i)*
 v: 14 16 19 32
 1- speak- 1PL- AP.DIST
 ‘we talke(d) about (it)’
- b. Oblique marker first person singular: position 35
wen- a- k’i- ’m (-ts’i-’m)*
 /wen- ay- k’V- m/
 v: 16 20 35 36
 give- 2PL- 1SG- AP.LOC
 ‘give (it to) me’

²⁶This suggests considering that at least the palatalization of /k^(‘)y/ might be a ‘lexical’ process, i.e. one that allows exceptions, so that the lack of palatalization in a certain position does not necessarily place it outside the word, if this is a valid wordhood diagnostic. The lack of palatalization of /k^(‘)y/ in position 35, however, is not as surprising as that of positions 17 and 19, since this position is outside the boundaries determined by several diagnostics, and also /ts’/ fails to palatalize in position 35.

c. Oblique marker first person plural: position 35

ti- jna- ts'e- 'm (*-ts'i- m)
 /t- hnay- ts'e- m/
 v: 14 16 35 36
 3- be.right- 1PL- AP.LOC
 'it is our job, it is proper for us'

The *minimal-B* subtype has a right edge in position 18. The causatives -jan, -jat of position 18 palatalize to -jyen, -jyet, as exemplified above. In positions 19 through 25 there is no way to know whether palatalization applies, either because the relevant morphemes begin in a vowel or for other reasons.²⁷ A positive instance of absence of palatalization is the mirative -p'an of position 26, which never palatalizes into *pyan or *pyen; contrast this with (36) with 'wanjli-jen → 'wanjli-jyi'n '(s)he rests', where the same root triggers palatalization of -jen (position 37).

(36) No palatalization in position 26

ø- 'wanjli- p'an- e
 v: 14 16 26 31
 3- remain- MIR- AP.INS
 'also...!', 'even...!' (Spanish *incluso*, *hasta*)

The *maximal-B* subtype defines again the span 14-40, like the *maximal-A*, see examples in (34).

The subtype C considers palatalization of coronals only, which can be triggered by underlying but also by some derived /i/, and also by /u/.

The *minimal-C* version gives just the span 14-16. A positive boundary is the oblique first person plural -ts'e of position 35, which fails to palatalize, as shown in 35c. But between 16 and 35 there is no target for palatalization, or any other way to verify whether palatalization would apply. Most morphemes do not begin in a coronal; others do (positions 21 and 25) but they have the form Ci, and although the concord marker of position 29 -is shows palatalization in -isy before a vowel, the trigger is morpheme-internal, and thus it should not define an edge.

Finally, the *maximal-C* version gives the subspan 14-46. Palatalization affects the initial phonemes of the DPs in positions 45-46 when they begin in a coronal: the demonstratives of the form Ca appear as Ci (where C=coronal), and the

²⁷The morphemes in positions 20 through 23 begin in a vowel or in Ci, and thus contain no target for palatalization. The evidential of 25 is underlyingly /-t'ey/, but since it is unstressed, it becomes -t'i(y), so again a target is lacking. And the reportative of 24 is scarcely documented in my material in this position, and I am not able to confirm or discard palatalization.

demonstratives *jlaja*, *jla'a* as *jlyCa* or *jliCa*, see (37).²⁸ Although the examples available are for palatalization in position 45 only, it seems reasonable to extend the edge to the following position 46, since it hosts elements of the same class, i.e. heavy DPs.

- (37) Palatalization *Maximal-C*: right edge in the DPs of 45-46

y-	i-	'i		[jlyaja Orán]
v: 14	16	34	45	-
3-	be-	AP.PUNCT	DEM:f	Orán

'[It] is in Orán.' (Drayson et al. 2000: 100)

6 Ciscategoriality revised

A domain of ciscategoriality is defined in Tallman (2021) as the span of structure wherein all elements are ciscategorial. If applied to wordhood diagnostics, it means that if, e.g., a morpheme can only attach to verbs but not to other word classes, then it belongs to the verb word. In other words, only ciscategorially selected elements belong to the word. Chorote is interesting in this regard because it allows not only the verb to head the predicate in position 16, but also other word classes — Ns/NPs, pronouns and even negation, which then take most of the usually 'verbal' markers. Furthermore, NPs and DPs can take some of the 'verbal' TAME markers even when they function as arguments. These two facts pose questions regarding how ciscategoriality should be defined as a comparative concept, since it is not clear whether it should be defined with respect to verbs or to predicates in general. Furthermore, Chorote is also interesting regarding typology of transcategoriality (Robert 2003) because it does not display transcategoriality evenly throughout its grammar. Thus, data from Chorote reveal that cis/transcategoriality is a matter of degree.

With respect to wordhood diagnostics, I suggest that two versions of ciscategorial selection diagnostics should be considered: a strict one, specific to verbs, which includes elements that can only be selected by verbs, and a lax one, which considers every element that can only be selected by the predicate head, no matter whether it is verbal or non-verbal, but not by the same categories in non-predicative functions. Importantly, notice that the *lax* subtype does not really define a word class, but rather a set of elements in predicate function.

²⁸The basic allomorphs are *-a(j)* for 'orientation' and *-e(j)* for the instrumental. After a vowel, epenthetic /y/ is inserted and included as part of the Ps. This /y/ in turn raises front vowels, as explained in §2, thus the Ps result in *-yej*, *-yij*, respectively. The P 'orientation' takes here a suppletive form when the pronominal complement surfaces.

Distinguishing between verbal and non-verbal predicates, in turn, forces one to address the question of what constitutes a verb in Chorote. The section is thus organized as follows: §6.1 proposes a definition of verb in Chorote; and §6.2 applies the diagnostics.

6.1 Defining verb in Chorote

In previous work (Carol 2013, 2014b) verbs in Chorote were defined as the words that take the personal prefixes of position 14. This of course makes ciscategoriality diagnostics circular, if those prefixes are used to define the left edge of a span on the basis that only verbs can combine with them, as will be seen below. But this definition is also problematic for a different reason. Class V verbs (see Table 2) take the personal prefixes, but differ from typical verbs in the form of the irrealis, where they take the suffix/enclitic *-a* of position 23, like nominal predicates, and not the irrealis set of personal prefixes; see (38). Furthermore, their third person prefix is always zero, and the plural of any person is expressed through a suffix *-(i)s* identical in form to the most common plural suffix of nouns. I underscore that this *-(i)s*, unlike the *-is* of position 29, is not used just with third person, but with *any* person, and comes immediately after the stem and *before* the TAME morphemes, all of which brings this Class V closer to the nominal domain. Provisionally, I assign this plural *-(i)s* the same position as the predicate head, i.e. 16, just like the nominal plural in nominal predicates.

(38) Exponence of irrealis in different kinds of predicates

a. Typical verb

<i>Ja</i>	n-	<i>ek</i>
v: 8	14	16
PROSP 3.IRR- go.away		
'(S)he/it will leave.'		

b. Nominal predicate

<i>Ja</i>	<i>anéchiyas-</i>	<i>as-</i>	a'
v: 8	16	16	23
PROSP chief-		PL- IRR	
'He will be chief.'			

c. Class V verb

<i>Ja</i>	Ø-	<i>'esy-</i>	e'
v: 8	14	16	23
PROSP 3-		be.good- IRR	
'(S)he/it will be good.'			

d. Plural Class V verb

ʃa *Ø- is-* *ísy-e'*
 v: 8 14 16 16 23
PROSP 3- be.good- **PL- IRR**
 'They are good.'

Being something between nouns and verbs regarding their morphosyntax, and considering the notions they express ('big', 'nice', 'white', etc.) it looks attractive to label Class V verbs as adjectives, as Drayson (2009) does in his dictionary. However, I have preferred to label them "verbs" for two reasons. Firstly, they take the same person indices as verbs in first and second person. Secondly, there is no evidence that these candidate adjectives in attributive function are structurally different from relativized clauses with a (typical) verbal predicate. Even though they appear superficially juxtaposed to nouns, as adjectives do in European languages, the same goes for verbs. The analysis of these is as free relative clauses with a null relative pronoun (Carol 2014b), which is the most usual strategy when the relative pronoun is the subject; see (39b) (otherwise a demonstrative usually surfaces as an explicit relative). Thus, there is no syntactic evidence to assign (39b) a syntax different from that of (39a).

(39) Class V verbs and typical verbs in attributive function

a. Class V verb

Si'yús *Ø- wuj* *in- ka-* *je'*.
 v: 5 5 5 14 16 33
 fish 3- **be.big** 3- have_joy- **AP.LOC**
 'The big fish is tasty.' (Lit. 'contains joy inside')

b. Typical verb

Pi *i'nyó'* *'yijén-* *e* *i- jlyut- i'*.
 v: 2 2 2:2 2 14 16 33
 DEM person 3:**be.wise-** **AP.INS** 3- rub- **AP.LOC**
 'The man who knows (how to make fire) drills (a piece of wood with stick).' (Drayson et al. 2000: 62)

The next question is whether Class V verbs can head a DP/NP in argument function, i.e. assume the typical syntactic function of nouns. Class V verbs can head a DP/NP in argument function, but so can typical verbs, and in the same way. Besides the zero relative shown in (39), demonstratives can also function as relative pronouns, so that a verb preceded by a demonstrative can be 'nominalized' in this way; (40) shows a lexicalized case. In sum, an NP/DP can be both

Dem+N or Dem+V, so here there is no reason to assign Class V verbs a different, more nominal status than that of the other verbs.

- (40) Typical verb heading a DP

ʃana ta- kelisye'n

DEM 3- sing

'radio/tape recorder' Lit. 'the one that sings'

Furthermore, when a DP/NP headed by a typical verb takes part in a construction that requires nominal irrealis (see example (17b) and the text that precedes it), the nominal irrealis morpheme *-a* surfaces, as in any DP/NP headed by a noun, see (41). Therefore, there is again no reason to assign Class V verbs in argumental function a distinct, non-verbal status.

- (41) DP headed by a verb with the nominal irrealis morpheme

Ø- Laj [ya- ka ta- kelisye- a']

v: 14 16 43 - - -

3- not_exist 1SG.POSS- ALI.POSS 3- sing- IRR

'I have no radio/tape recorder.' Lit. 'There is no radio/tape recorder of mine.'

6.2 Diagnostics based on ciscategoriality

The strict and lax versions of ciscategoriality proposed above can combine with the known *minimal-maximal* distinction - a subspan of contiguous positions that satisfy the requirements, or the longest possible subspan which only includes elements that satisfy the requirements, respectively. Or, in other words, a minimal subspan which only includes ciscategorial elements, and a maximal subspan outside of which all elements are transcategorial. In sum, we obtain four diagnostics.

6.2.1 Strict ciscategoriality (14-20; 14-37)

The *strict* (i.e. specific to verbs) *minimal* version has the person prefixes of position 14 and the concord morphemes of position 20 as its edges, see (42). Neither occur in non-verbal predicates, where person/number is indicated through oblique morphemes (and a postposition bound to them), see (43).

- (42) Ciscategoriality - strict minimal: span 14-20

ʃi- 'wen- a'

v: 14- 16- 20

2- see- 2PL

'You (pl.) see it/him/her.'

(43) Person marking in nominal predication

I- lis as- e'm
 v: 16 16 35 36
 1SG.POSS- sons 2PL- AP.LOC
 'You are my sons.'

The left edge is 14 only if we consider that Class V verbs are real verbs. Otherwise, the edge should be placed in position 15 - only transitives can take the reflexive/reciprocal and antipassive morphemes, and Class V are not among them.

The positions to the right of 20 can co-occur with non-verbal predicates. The perdurative of 21 is scarcely documented with non-verbs in my material, but co-occurs with a noun in *jloma-jli* (day-PERD) 'during the day', as well as in 'wena-jli-yi (different_thing-PERD-P_{INST}) 'but' (Spanish 'sino'; Drayson et al. 2000: 'WENAJLIYI), lexicalized as a conjunction. The momentary, in position 22, can also combine with other word classes; see (44) (also 46a-b below); in 44aa-b (as well as in (46), I assume negation occupies the position of the predicate head, i.e. position 16.²⁹ Also the irrealis and the other TAME morphemes that follow combine with non-verbs, as was seen in (17b) and (38b), and will be seen below in (46).

(44) Momentary with non-verbs

- a. [A: E- jetik Ø- a'tye- je'?] B: Je- ye 'ne'
 v: 16 22 40
 [A: 2.POSS- head- 3- hurt- AP.LOC] B: NEG- MOM now
 [A: 'Do you have a headache?'] B: 'Not anymore.'
- b. [Syupa] ti jloma- ye- t'i- jyi...
 v: [-] 4 16 22 25 39
 [DEM] COMP day- MOM- EVID- PRF
 '[Then] the next day...' Lit. 'when it was day again...'

The *strict maximal* subtype has the same left edge as the *minimal* one. The right edge is position 37, occupied by the polysemous *-jen*, which combines even (though rarely) with Class V verbs, as in (45), where it functions as a plural marker. What makes Class V verbs different is that they pattern with nouns in many respects, as shown in (39). Nevertheless, they pattern with the other verbs here and not with nouns, on which *-jen* is not documented.

²⁹That there is a context for palatalization in 33 can be seen in the contrast with *t-amti-ts'i-ji'n* 'they speak', with the same root, where *-ts'i* is the palatalized allomorph of the distributive *-k'i* of position 34.

- (45) Ciscategoriality - strict maximal: subspan 14-37

kas- 'wasajne'n
/kas- 'wasan- jen/
 v: 14 16 37
 1PL- be.alive- JEN
 'We are alive.'

Any material to the left of position 14 can combine with other word classes, even the prospective of position 8, which combines with nominal predicates, see (38b). As for material to the right of 37, position 38 is occupied by the pluractional -'ni(j), which usually attaches to verbs, even from Class V, but which can be seen attached to negation in (46).

- (46) Pluractional 'ni(j) of position 37 with non-verbs

'Yina je- ye- 'ni wata'a [ka Ø- tojw- a-
 v: 1 16 22 38 40 46 - - -
 I_mean(?) NEG- MOM- PACT so_much COMP 3- be.distant- IRR-
k'i'].
 -
 AP.DISTR
 'I mean, it is not so distant [as the previous place].' (Drayson et al. 2000: 94)

6.2.2 Lax ciscategoriality (14-22; 8-40)

The *lax* subtype is similar to the strict one, but replacing “verb” by “main predicate of the clause”, whether verbal or not. Recall that, as stated above, this cannot define a word class, but an element that displays a predicate function. As a consequence, if an element, e.g. the oblique markers in position 35 that host applicatives, can combine with a non-verbal category, e.g. nouns, only when the noun is in predicate function, this does not mean that the oblique markers belong to the noun class - if they can only combine when the noun heads a predicate, then they belong to predicates, not to nouns themselves.

The *lax-minimal* subtype also has the person prefixes in position 14 as its left edge. By definition, when the predicate is non-verbal, this position, as well as position 15, is simply empty. The right edge is the momentary morpheme in position 22, which only occurs bound to the predicate head, verbal or not, as in (44). The span is thus 14-22; an example of this was provided in (16).³⁰ In turn, the ir-

³⁰The translation ‘please’ (*por favor*) was suggested to me with imperatives. In other cases, however, it is very difficult to find an equivalence. It seems to indicate a benefit for some participant.

realis of position 23 attaches to nouns in argument function when the existence of the entity denoted by the N(P) is not asserted, as in (17b) and (41), and is thus excluded from the span identified by this diagnostic.

The *lax-maximal* version of the diagnostic has the prospective morpheme of position 8 as its left edge; see examples in (38). The right edge is position 40, which contains some adverbs only documented attached to the predicate head. In sum, the subspan is 8-40. An example of the *lax-maximal* subspan is (47).

- (47) Ciscategoriality - lax maximal: subspan 840.

Ja- kas- 'wasan- a- jan- na'a
 v: 8 14 16 23 37 40
PROSP- 1PL- be.alive- IRR- JEN- later
 'We will be alive, we will survive'.

The morphemes *ja* and *-na'a* at the edges of (47) are only documented bound to the predicate head, although other adverbs in position 40 are free, e.g. fronted to position 3; see §4.1. The complementizers of position 4 usually co-occur with predicate heads of any class, but not always - in cases like (48) *sa'am* 'we' can hardly be considered a predicate head; in this example, the first *ti* in position 2 introduces the topic *sa'am* 'we', the second *ti* in the usual position 4 apparently heads the main clause, and the last *ti* in position 46 heads an adverbial clause. Thus, position 4 is excluded from this subspan.

- (48) Complementizer *ti* not introducing a clause

Jlampet ti sa'am ti a- wa- k- i [siuni- wa jloma- s]
 v: 1 2 2 4 14 16 19 32 45 - - -
 but **COMP 1PL COMP 1- be- 1PL- AP.DIST- DEM- PL day- PL**
[ti Ø- 'nes- a- t'i- pi ni Si- nya' jl- amt-
 46 - - - - - - - - - -
COMP 3- arrive- MOM- EVID- REM.PST 1PL.POSS- father 3POSS- word- PL
is].
 -

'But we were already there those days when the Gospel arrived.'
 (Drayson et al. 2000: 106)

7 Conclusions

From the application of constituency diagnostics to Chorote using the methodology advocated in Tallman (2021) there does not emerge an obvious wordhood

candidate. As can be seen in Tables 4 and 5, which summarize the results of applying morphosyntactic and phonological diagnostics, respectively. As stated in the introduction, the maximum of diagnostic subtypes for a complete span is two, for the subspans 4-40, 14-16, 14-22, 14-40, 16-34, and maybe 14-18. If each edge is taken separately, the left edge shows more convergence, with 14 diagnostic subtypes converging on the personal prefixes in position 14, and six on the predicate head in position 16, while the right edge has the highest convergence of five diagnostic subtypes on position 40.

If one takes morphosyntactic and phonological diagnostics separately, looking for separate grammatical and phonological words, the results are not very different. The left edge of a possible grammatical word could reasonably be position 14 (convergence of eight subtypes), but the right edge could be the positions 18, 22, 39 (or -38) or 40, with two subtypes converging in each case; from these, only position 22 shows convergence in position 14 (for both subtypes) in the left edge as well.

As for a possible phonological word, the left edge has two candidates: again position 14, with a convergence of six subtypes, and position 16, with four subtypes. But all six subtypes converging in position 14 correspond to the palatalization diagnostic, while the ones converging in position 16 correspond to accent (one) and *y*-insertion (three) diagnostics. If we take into account the iambic type conjectured for Proto- or Pre-Chorote, there would be a seventh subtype converging in position 14, which would not be related to palatalization. As for the right edge, positions 16 and 40 show the highest number of subtype convergences - three, belonging each to two different diagnostic types. However, position 16 is very problematic as a candidate for the right edge of the phonological word if we consider that it is the position of verbal root and the left edge for several other diagnostics. In turn, position 40, filled by adverbial particles that can encliticize to the predicate head, appears as a more reasonable candidate. Two subtypes that give position 40 as right edge also give position 14 as a left edge, which makes the span 14-40 the only 'candidate' for the phonological word, but with only two diagnostic subtypes converging in it.

Finally, Chorote is interesting regarding the typology of transcategoriality because it shows features of different types. On the one hand, there is a distinct verb word class. Concord morphology can be selected only by certain stems, *verbal* stems. Other words are not "transcategorized" into verbs when they function as predicates. In those cases, the subject is cross-referenced by oblique markers, as seen in (43). In this respect, Chorote is not different from languages with heavy morphology and limited transcategoriality (Robert 2003), except for the fact that it lacks a copula.

Table 4: Morphosyntactic wordhood diagnostics

Abstract type	Subtype	Language specific fracture	Left-edge	Right-edge	Size	Convergence
Freedom	Minimal		16	16	1	1
	Maximal		4	40	37	2
Non- Interruption	Single free interruptor		14	39 (or -38)	26 (or 25)	1
	Multiple free interruptor		7	41	35	1
	Nonfixed interruptor		14	22	9	2
Deviation		Inflectional class	14	18	5	1/2
	Extended	Fossilization	16	18	3	1
	Exponence		14	29	16	1
	Rigid		14	23	9	1
Non- permutability						
Subspan repetition	Minimal	Coordination	8	38	33	1
	Maximal	Coordination	8	39	34	1
Ciscategorical selection	Strict minimal		14	20	7	1
	Strict maximal		14	37	24	1
	Lax maximal		14	22	9	2
	Lax minimal		8	40	30	1

Table 5: Phonological wordhood diagnostics

Abstract type	Subtype	Language specific fracture	Left-edge	Right-edge	Size	Convergence
Accent	Minimal-minimal		15	16	2	1
	Minimal-maximal		16	34	19	2
	Maximal-maximal		4	40	37	2
	Minimal		16	32	17	1
	Maximal-minimal		16	34	19	2
y-insertion	Maximal-maximal		16	44	28	1
	Minimal-A	With k ^(·) y	14	16	3	2
	Maximal-A	With k ^(·) y	14	40	27	2
	Minimal-B	Without k ^(·) y	14	18/25	5/12	2/1
	Maximal-B	Without k ^(·) y	14	40	27	2
Palatalization	Minimal-C	Coronals only	14	16	3	2
	Maximal-C	Coronals only	14	46	32	1

On the other hand, in other respects the language seems to make extensive use of transcategoriality, something which has been correlated with the isolating type (i.e. the type of languages with weak morphology; Robert 2003). For example, an inflected verb can perform a referential function -i.e. head a noun phrase- without any overt transcategorial morphology, as (40) shows. Moreover, many TAME markers can be bound to NPs, clearly taking nominal rather than clausal scope (Carol 2014b, 2015). All this gives Chorote some properties of an ‘omnipredicative language’ (Launey 1994), which in turn underscores that tests based on cis-/transcategoriality deserve further discussion.

Abbreviations

ALI	alienable	LV	light verb
ANTIP	antipassive	MIR	mirative
AP	applicative/adposition	MOM	momentary
CAUS	causative	OR	orientation
COMP	complementizer	PLACT	pluractional
DEM	demonstrative	PP	participle
DIST	distal	PRF	perfect
DISTR	distributive	PROSP	prospective
EPEN	epenthetic	PST	past
EVID	evidential	PUNCT	punctual
INCOMP	incompletive	REFL	reflexive
INTER	interrogative	REM	remote
IRR	irrealis	REP	reportative
JEN	pluractional/downwards	VLZ	verbalizer
LOC	locative		

References

- Carol, Javier J. 2011. Aplicativos/adposiciones en chorote (mataguayo): Algunos aspectos formales. *LIAMES* 11. 51–73.
- Carol, Javier J. 2013. Marcación de argumentos En el verbo En chorote: Intransitividad escindida Y otros sistemas de alineamiento. *International Journal of American Linguistics* 79(4). 491–532.
- Carol, Javier J. 2014a. Esbozo fonológico del chorote (mataguayo). *LIAMES: Línguas Indígenas Americanas* 14. 73–103.

- Carol, Javier J. 2014b. *Lengua chorote (mataguayo): Estudio fonológico y morfosintáctico* (Studies in Native American Linguistics 72). Munich: LINCOM.
- Carol, Javier J. 2015. TAM marking on nominals in chorote. *Linguistics* 53(4). 877–930.
- Carol, Javier J. 2018. *Inamtes fleeizi' Inkijwas ji'lij – Kiláyi Ji'lij*. Diccionario Bilingüe Manjui – Castellano. Curupaity: Secretaría de Políticas Lingüísticas.
- Carol, Javier J. & Andrés Pablo Salanova. 2012. On pseudo-applicatives. *Cahiers Linguistiques d'Ottawa / Ottawa Papers in Linguistics* 37. 22–42.
- Claesson, Kenneth. 2016. ? Vocabulario 'weenhayek-español.
- Dirección General de Estadísticas, Encuestas y Censos. 2012. *IIIer censo nacional de población y viviendas para pueblos indígenas. Resultados finales*. Asunción.
- Drayson, Nicholas. 1999. *Ja'yitaja'a na si'lij jla'yi ti jayamtijiena [Vamos a leer y escribir en Chorote]*. Tartagal: ASOCIANA.
- Drayson, Nicholas. 2009. 'niwak samtis. Diccionario iyojwa'ja'lij – kilay'lij (chorote- castellano). In José Braunstein & Cristina Messineo (eds.), *Hacia una nueva carta étnica del gran chaco VIII*, 91–174. Buenos Aires & Las Lomitas: PICT 32894 (ANPCyT), Centro del Hombre Antiguo Chaqueño.
- Drayson, Nicholas, Sebastian Frías & Julián Gomez. 2000. *Sake' iyo ti iyojwa'jats'e'm. Somos chorotes – nuestras costumbres*. Tartagal: ASOCIANA.
- Gerzenstein, Ana. 1983. *Lengua chorote: Estudio descriptivo-comparativo y vocabulario*. 2nd edn. (Archivo de lenguas precolombinas 4). Buenos Aires: Universidad de Buenos Aires, Facultad de Filosofía y Letras, Instituto de Lingüística.
- Gomez, Julian & Nicholas Drayson. 1997. *Sinia' jlamtis. El nuevo testamento en el idioma iyojwa'ja (chorote)*. Buenos Aires: Sociedad Bíblica Argentina.
- Gutiérrez, Analía & Verónica Nercesian. 2021. La glotal y la glotalización en las lenguas mataguayas. *Forma y Función* 1(34). DOI: 10.15446/fyf.v34n1.79328.
- Gutiérrez, Analía. 2015. *Segmental and prosodic complexity in niva'le: Laryngeals, laterals, and metathesis*. Vancouver: University of British Columbia. (Doctoral dissertation).
- Gutiérrez, Analía. 2016. Patterns of (de)glottalization in niva'le. In Kim Kyeong Min, Pocholo Umbal, Trevor Block, Chan Queenie, Tanie Cheng, Kelli Finney, Mara Katz, Sophie Nickel Thompson & Lisa Shorten (eds.), *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, 176–185. Somerville: Cascadia Proceedings Project.
- Instituto Nacional de Estadística y Censo, Argentina. 2005. *Encuesta complementaria de pueblos indígenas (ECPI)*.
- Instituto Nacional de Estadística y Censo, Argentina. 2012. *Encuesta complementaria de pueblos indígenas (ECPI)*.
- Launey, Michel. 1994. *Une grammaire omniprédicative*. Paris: CNRS.

- Nercesian, Verónica. 2014. *Wichi lhomtes: Estudio de la gramática y la interacción fonología–morfología–sintaxis–semántica* (Studies in Native American Linguistics 74). Munich: LINCOM.
- Nikulin, Andrey & Javier Carol. Forthcoming. *Historical phonology of Mataguayan* (Topics in Phonological Diversity). Berlin: Language Science Press.
- Robert, Stéphane. 2003. Vers une typologie de la transcatégorialité. In Stéphane Robert (ed.), *Perspectives synchroniques sur la grammaticalisation: Polysémie, transcatégorialité et échelles syntaxiques* (Collection Afrique et Langage 5), 255–270. Louvain: Peeters.
- Tallman, Adam J. R. 2021. Constituency and coincidence in Chácobo (Pano). *Studies in Language* 45(2). 321–383. DOI: 10.1075/sl.19025.tal.

