

# Chapter 11

## Constituency in Yukuna

Magdalena Lemus Serrano

Aix-Marseille Université Laboratoire Parole et Langage

This study provides an overview of constituency in the verbal domain in Yukuna, an Arawakan language of Colombian Amazonia, on the basis of a firsthand corpus of texts. Using the methodology developed in Tallman (2021), we establish a verbal planar structure, to which we applied a total of 30 constituency tests pertaining to different domains (morphosyntactic, phonological, indeterminate), including both cross-linguistic and language-specific tests. The results show that, similarly to many other languages in this book, there is no strict convergence of tests around a single layer across different domains, nor strict convergence within domains. Instead of a strict phonological vs. grammatical word distinction, the results point to a distinction between positions placed before and after the verb core, so that in Yukuna, diagnostics tend to select either a span including person indexes and the verb core to the exclusion of all following formatives, or a span including the verb and its following formatives, to the exclusion of person indexes.

### 1 Introduction

This chapter provides an overview of constituency in the verbal domain in Yukuna, following the methodology in Tallman (2021).

Yukuna (ISO 693-3:ycn, Glottocode: yucu1253) is a North-Amazonian Arawakan language spoken by under one thousand speakers in South Eastern Colombia. This study is based on ongoing work on the Yukuna language by the author, and the Yukuna grammar sketch in Lemus Serrano (2020). All examples come from firsthand data collected during three fieldtrips in various Yukuna speaking communities between 2015 and 2018. The corpus contains roughly five hours of annotated texts.



The examples are transcribed alphabetically with a slightly modified version of the Yukuna writing system used in the Yukuna dictionary (Schauer et al. 2005), based on Spanish. The following alphabetic conventions are used: <j> /h/, <ñ> /ɲ/, <y> /j/, <V'> /V̥/ (creaky vowel), <Ch> /C<sup>h</sup>/ (aspirated plosive), <jC> /Ç/ (voiceless sonorant). High tone is transcribed with an acute accent. The surface manifestation of H tones is very variable, so the same morpheme may be transcribed with or without an accent depending on the context (e.g. past tense *-cha* is found in examples both as *-cha* and as *-chá*). Lastly, each example contains a source indicating the name of its audio file, and the ELAN/Flex line. Examples taken from elicitation come from the first author's field notes.

## 2 Yukuna language and its speakers

Yukuna is an Arawakan language of the Japurá-Colombia branch (Ramirez 2001), spoken in various communities along the Mirití-Paraná River in North-Western Amazonia (Figure 1). The Yukuna language is spoken by the Yukuna and Matapí ethnic groups, who are in intense, long-term contact with the (Tukanoan) speaking groups Tanimuka and Letuama (Fontaine 2001: 57). Despite the overall small number of speakers, the language continues to be transmitted to new generations within the Mirití-Paraná communities, so most ethnic Yukuna and Matapi of all ages speak their language. The relative stability of the language in the Mirití-Paraná contrasts with the sharp decline in vitality of the language when speakers move to nearby towns and cities, where Spanish and Portuguese are the dominant languages (Lemus Serrano 2016: 24).

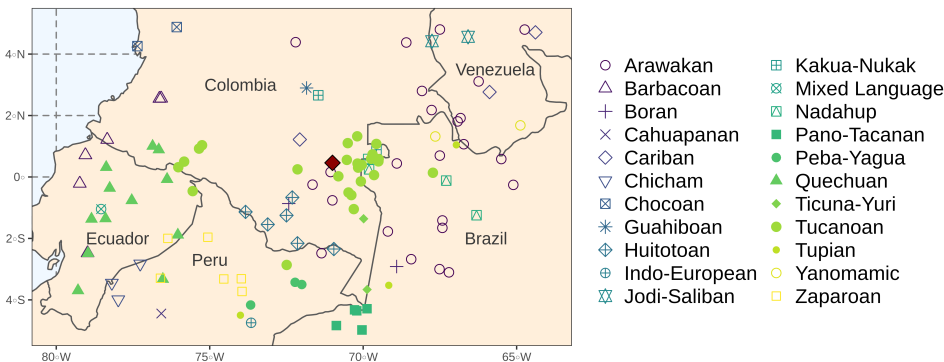


Figure 1: Languages of North-West Amazonia (Yukuna is marked with a red diamond) (Hammarström et al. 2019)

Typologically, Yukuna is a nominative-accusative language. The lexical word classes are nouns, verbs, adverbs, adjectives, and postpositions. Formatives, that is to say, grammatical morphemes, mostly follow a lexical core, and they are concatenative, monoexponential, with little allomorphy. As such, Yukuna can be said to be an agglutinating, suffixing language. Core arguments are not case-marked, and obliques are marked with postpositions. The predominant constituent order is SVO. The Subject is rigidly placed before the verb and it is adjacent to it, whereas objects and obliques are not necessarily adjacent to the verb, they are variably ordered with respect to each other (SVOX/SVXO), and can also be preposed to the verb (XOSV) (see Lemus Serrano 2020: 97–102).

### 3 Planar structure of the verbal complex

This section presents the verbal planar structure of Yukuna. A planar structure represents all the elements that can occur in a clause with a verbal predicate, by flattening all hierarchical distinctions between morphemes, words and constituents. This planar structure does not represent other types of phrases such as noun phrases (NPs) or postpositional phrases (PPs), as they have their own internal structure. Yukuna's planar structure is detailed in Table 1. Positions within the planar structure are numbered (from 1 to 21) and classed per type (slot or zone) depending on whether elements inside mutually exclude one another, or can freely co-occur. For each position, a broad description of the type of elements that it contains as well as the specific forms used. Empty cells correspond to positions filled with elements from open word classes. An in-depth description of all of the markers in this planar structure is given in §5 of the Yukuna grammar sketch in Lemus Serrano (2020).

As shown in this table, the structure contains 21 positions, with the verb core placed in position 10. Grammatical markers of the verb core are placed from positions 9 through 20. Henceforth, I refer to this set of positions as the verb complex. Constituents, obliques, and adverbials are placed either before or after the verb complex, and display quite a lot of variation depending on polarity, speech acts, and information structure. As planar structures aim to avoid a potentially arbitrary classification of verbal clauses into different *constructions*, the planar structure in Table (1) merges together all structures in the language containing a verbal predicate, including all clause types. Doing so raises at least two methodological challenges.

The first challenge concerns the issue of synchronically ambiguous clefts. In Yukuna, there are several structures that could be analyzed either as non-verbal

Table 1: Yukuna verbal planar structure

Pos.	Type	Elements	Forms
(1)	slot	connectors	
(2)	slot	adverbial interrogatives	<i>náje, méño'jó</i>
(3)	zone	adverbials	
(4)	slot	negation	<i>unká</i>
(5)	slot	constituent focus/interrogative	
(6)	slot	negation	<i>unká</i>
(7)	slot	indefinite proforms	<i>ná, méké, etc.</i>
(8)	slot	S/A NP	
(9)	slot	person indexes (S/A)	<i>nu-, pi-, ri-, ru-, etc.</i>
(10)	slot	<b>verb core</b>	
(11)	slot	valency	<i>-ta, -ñaa, -ka</i>
(12)	slot	negation	<i>-la</i>
(13)	slot	tense	<i>-cha, -je, -khe, -jika</i>
(14)	slot	nominalization, mood	<i>-ka, -kare</i>
(15)	slot	gender/number	<i>-ri, -yo, -ño</i>
(16)	slot	imperative	<i>-chi, -re, -niña</i>
(17)	slot	middle	<i>=o</i>
(18)	slot	perfective	<i>=mi</i>
(19)	zone	habitual, frustrative	<i>=no, jlá</i>
(20)	slot	discourse markers	<i>=ko, =ja</i>
(21)	zone	P, PP, OBL, ADV, remote past	

predicates (pseudo-clefts), or as verbal predicates used in focalization. Pseudo-clefts have the same syntactic structure as equative non-verbal predicates formed by juxtaposing the predicate and the argument without any copula. The focused constituent is placed in clause initial position, followed by a clausal nominalization in the position of the argument, as in (1).

- (1) *Ná kéelé wáa' -ri?*

INDF DEM call -M

'Who is that one calling?'

ycn0068,123

Some uses of these pseudo-clefts are synchronically ambiguous, as their surface structure is almost identical to that of main verbal clauses. To avoid arbitrarily excluding a type of verbal predicate, these ambiguous cases were also

integrated into the planar structure. Their inclusion required adding additional positions, as the focused constituent can be separated from the verb complex by various elements (notably pre-verbal negation marker *unká*), while the S/A NP in non-focused clauses is strictly adjacent to the verb complex. This leads to two different positions in the template where the S/A NP can be placed: position 5 when focused (2a), and position 8 when not focused (2b).

- (2) a. *Kawayá iphí -cha -ri.*  
           5       10     -13 -15  
           deer    arrive -PST -M  
           ‘The deer arrived.’ycn0041,156
- b. *Unká iná i’jna -lá matha’ -jé.*  
           6     8     10   -12 10     -  
           NEG INDF go   -NEG cut     -PURP  
           ‘One does not go cutting.’ycn0119,29

The second methodological challenge concerns word order variations. As stated previously, Yukuna displays variable ordering of some constituents (objects and obliques) as well as some formatives (notably, negation). However, despite the attested variability, the ordering of elements is constrained by several restrictions, especially for elements in positions before the verb complex. In order to fully capture these ordering restrictions, I opted for increasing the number of slots before the verb complex, placing the same elements in multiple positions, and simply placing the freely ordered elements after the verb complex in a single zone (21). For instance, the negation marker *unká* is placed variably with respect to focused constituents (position 5), but it is obligatorily placed immediately before indefinites (position 7), so it is placed in two different positions in the planar structure (positions 4 and 6). The variable positioning of the negative marker *unká* with respect to focused constituents is illustrated in examples (3a) and (3b).

- (3) a. *Unká na=jló nu= yuí -la -je rikhá.*  
           4     5       9= 10   -12 -13 21  
           NEG 3PL=to 1SG= leave -NEG -FUT 3SG  
           ‘I will not leave it to them.’ycn0092,117
- b. *Rikhá unká amá -la nukhá.*  
           5     6     10   -12 21  
           3SG NEG see -NEG 1SG  
           ‘He did not see me.’ycn0117,93

While increasing the number of positions captures all ordering possibilities, multiplying the positions within the planar structure also has the disadvantage of leading to uncertainty in the numbering of elements in examples. For instance, in a given occurrence of standard negation without focused constituents, indefinite pronouns or an overt A/S NP like in (4), it is not possible to know for certain that the negative marker *unká* is in position 4 or 6. In these cases, I assume that negation is placed in position 6 unless there is an overt focused constituent placed before it, as in (3). These arbitrary decisions were necessary for coherency in numbering throughout the chapter, but note that they have no consequence for the results of the diagnostics applied.

- (4) *Unká ru= ajá -lá -cha.*  
 6 9= 10 -12 -13  
 NEG 3SG.F= fly -NEG -PST  
 ‘She did not fly.’ ycn0041,29

## 4 Diagnostics and layers

We applied a total of 30 constituency tests to the planar structure in (1). Each test is given an ID number, and is assigned to one of three domains used in this volume (morphosyntax “MS”, indeterminate “IND” and phonology “PH”), a name (name of the diagnostic), a fracture (different interpretations of the same diagnostic), and a span size (minimal vs. maximal). Figure 2 provides the results of the tests. Each line provides the name of the test, the span of positions identified (from 1 to 21), and the size of the layer identified (numbers in squares at the edges of each line). Tests are sorted by relative size of the identified span, from largest (top) to smallest (bottom). For instance, the last test in (1), the minimal application of one kind of subspan repetition test (lexical nominalizations) selects a span of positions from 10 to 11, meaning that it has a size of one.

The following sections discuss diagnostics per domain, and provide examples for each identified layer. Morphosyntactic diagnostics are presented in §5, indeterminate diagnostics, which concerns both phonological and morphosyntactic factors, are presented in §6, and finally, phonological diagnostics in §7.

## 5 Morphosyntactic domains

This section presents the results of the morphosyntactic diagnostics applied. In sum, nine different diagnostics were applied, each with multiple fractures, for a total of 14 tests, as provided in Figure 3. Diagnostics concerning both the domains

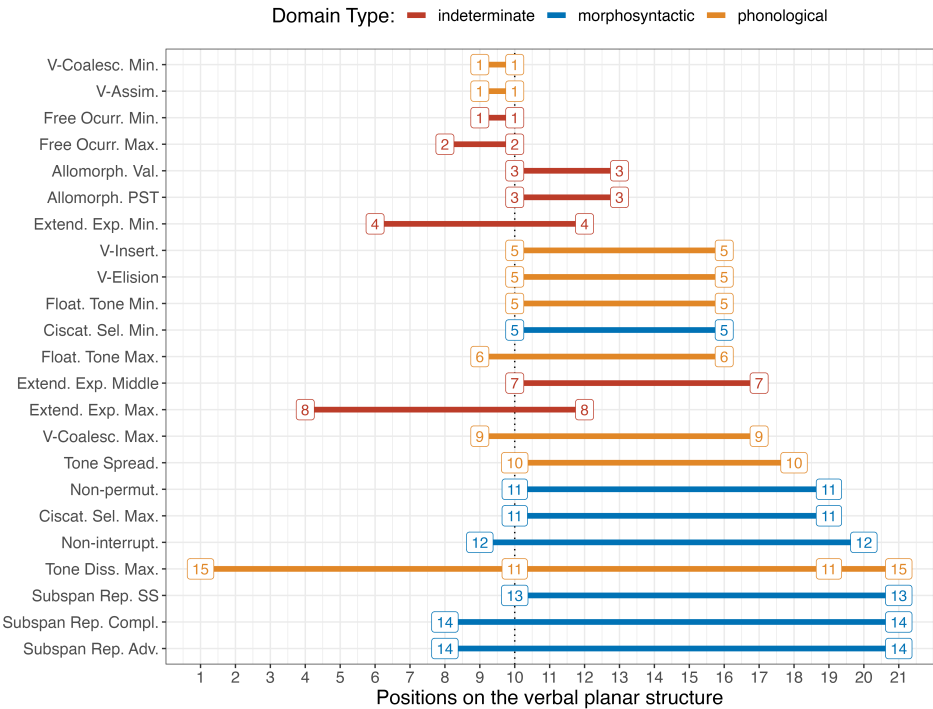


Figure 2: Constituency tests per converging layers

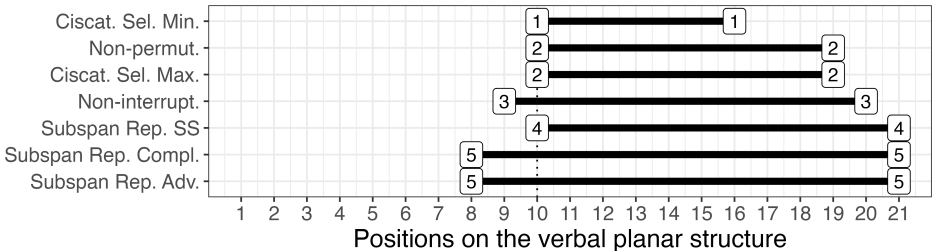


Figure 3: Morphosyntactic constituency tests per converging layers

of phonology and morphosyntax (e.g. free occurrence, opaque allomorphy) are presented in §6 on indeterminate domains.

### 5.1 Non-interruptability

This diagnostic identifies spans within the planar structure that cannot be separated by free and/or promiscuous elements. Different interpretations of this diagnostic require fracturing it into two tests, namely non-interruptability by a free form and non-interruptability by a promiscuous form, holding multiple positions within the planar structure.

#### 5.1.1 Non-interruptability by a free form (9-20)

The first fracture identifies a span from positions 9 to 20, with the core in position 10. The positions placed immediately before and after this span contain the verbal arguments, the S/A NP in 8, and the object plus all obliques in 21. The results of this test are the same regardless of the type of free form, whether simplex or complex. Example (5) illustrates the non-interruptability test, by showing that free forms, such as nominal and pronominal arguments (a noun with a bound person index in the position of the subject 8, and an independent pronoun in the position of the object in 21) are placed before and after the span selected by this test from positions 9 to 20.

- (5) *Kája ri=pirá nó -cha rikhá.*  
1 8 10 -13 21  
then 3SG=pet kill -PST 3SG  
'Then his pet killed him.' ycn0053,33

Examples (6) and (7) illustrate the same test, showing that the same span of the planar structure is identified even when arguments are expressed with complex free forms, such as a complex A/S NP, and a complex object NP respectively (in brackets).

- (6) [*kéelé na=e'wé phe'jǐ*] *kémí -cha*  
8 10 -13  
DEM 3PL=sibling eldest say -PST  
'their eldest sibling said' ycn0189,146
- (7) *kája yáwi tá nó -cha =mi [kéelé nu=yajná michú]*  
1 8 10 -13 =18 21  
then jaguar EMPH kill -PST =PFV DEM 1SG=husband deceased  
'The jaguar already killed my late husband.' ycn0053,88



### 5.1.2 Non-interruptability by a promiscuous form (9-18)

The second fracture of this diagnostic, the test of non-interruptability by a promiscuous form, identifies a slightly smaller span, from positions 9 to 18. This is due to the fact that the habitual marker *=no* can appear in multiple positions within the planar structure (positions 19 and 21), as shown in example (8).

- (8) *ri= puri' -chá =o =no ri=jwa'té=no*  
 9= 10 -13 =17 =19 21  
 3SG= talk -PST =MID =HAB 3SG=with=HAB  
 'he was always talking to him' ycn0041,142

## 5.2 Ciscategorial selection

This diagnostic identifies a span of positions within the planar structure containing formatives exclusive to the verb complex, which only combine with verb roots. The application of this diagnostic required fracturing into a minimal and maximal tests in Yukuna.

### 5.2.1 Ciscategorial selection minimal (10-16)

The minimal fracture of this diagnostic identified a span from positions 10 to 16 including valency markers, verbal negation, tense markers, nominalizers and dependency markers, as well as various imperative mood markers. These categories are either not encoded on other word classes, or encoded with other markers, as is the case for negation, double-marked with particles *unká ... kalé* (9) as opposed to the verbal marking with *unká ... -la* (3).

- (9) *unká o'wé kalé*  
 NEG brother NVNEG  
 'It is not my brother.' ycn0041,126

Before and after these positions we find multicategorial markers such as person indexes (position 9), the middle voice marker *=o* (position 17) and the perfective aspect marker *=mi* (position 18). Person indexes encode possessors on nouns, *=o* encodes reflexivity on nouns (10), and *=mi* encodes former possession on nouns as well (11).

- (10) *na= mená =o*  
 3PL= cropland =MID  
 'their own cropland' ycn0058,41

- (11) *pají =mí*  
house =PFV  
‘an abandoned house’ ycn0079,18

### 5.2.2 Ciscategorial selection maximal (10-19)

The ciscategoriality test allows a second interpretation in Yukuna, as there are verbal-exclusive markers beyond the multicategorial markers in positions 17-18, namely, the frustrative mood marker *=jlá* in the position 19. In order to capture this, the ciscategoriality diagnostic was fractured into two tests, a minimal test and a maximal test. The latter identifies a larger span ranging from positions 10 to 19, outside of which all elements are transcategorial.

Lastly, while Yukuna has non-verbal predicates, no additional fracturing was necessary to capture the behavior of these predicates with respect to ciscategorial selection. Yukuna displays an alternation between a zero copula clause type where the non-verbal predicate displays no ciscategorial verbal markers at all, and a verbal copula clause type, where the verbal copula combines with all ciscategorial verbal markers.

### 5.3 Non-permutability (9-18)

This diagnostic seeks to identify which spans from the planar structure contain rigidly ordered elements. The application of this diagnostic excludes zones, positions within the planar structure that contain variably ordered elements, as well as promiscuous elements, placed in various positions within the planar structure (see §5.1). The identified span ranges from positions 9 to 18. This span includes only one lexical element, the verb core, and its grammatical markers up until the perfective marker. At the right edge of this span we find a zone of variable ordering in position 19 containing the markers *=jlá* ‘frustrative’ and *=no* ‘habitual’, which can be freely ordered with respect to one another without any identified difference in scope. The variable ordering of the elements in position 19 is illustrated in (12-13).

- (12) *ri= nó -cha =jlá =no kamejéri*  
9= 10 -13 =19 =19 21  
3SG= kill -PST =FRUST =HAB animal  
‘he kept trying to kill animals’ elicited

- (13) *ri= nó -cha =nó =jlá kamejéri*  
 9= 10 -13 =19 =19 21  
 3SG= kill -PST =HAB =FRUST animal  
 ‘he kept trying to kill animals’ elicited

Before the left edge of this span, we find a promiscuous element, the subject NP, which is found in positions 5 for focused constituents, and in 8. The different positions of subjects were illustrated previously in examples (2) and (3).

#### 5.4 Deviations from biuniqueness: extended exponence

This section focuses on a specific type of instance of deviations from biuniqueness, namely, cases of extended exponence involving discontinuous markers. There are two such cases in Yukuna, used as two different diagnostics, and three different tests. Other instances of deviations from biuniqueness (allomorphy) are discussed in §6.2.

##### 5.4.1 Discontinuous stems with =o (10-17)

The first diagnostic concerns extended exponence of verbal stems with the middle marker =o, and identifies a span of positions starting from the verb core 10 up to 17. As a derivational device, the use of =o displays many idiosyncrasies. There are, for instance, multiple cases where the stem is not at all attested without the middle marker, and where the semantics of these stems are not compositional. A case in point concerns the verb *jecho*’=o ‘run’ (see example (17) below), where the stem is discontinuous, and includes the verb core in position 10 and the marker =o in position 17.

##### 5.4.2 Discontinuous negation minimal (6-12) and maximal (4-12)

The second diagnostic concerns extended exponence of verbal negation, which is obligatorily double-marked with the free form *unká* in positions 4 and 6, and the bound marker *-la* in position 12.<sup>1</sup> As the free form *unká* is promiscuous and appears twice in the planar structure, this diagnostic is fractured into a minimal and a maximal domain. The minimal domain identifies a span from positions 6 to 12, and the maximal domain identifies a larger span, from positions 4 to 12.

<sup>1</sup>Although both markers are required to encode negation, there are a few instances in which only the pre-verbal free form *unká* is used, and the suffix *-la* is omitted: when the far past tense suffix *-khe* is used, and in certain types of subordinate clauses (negative conditional subordinate clauses).

These two spans depending on the position of *unká* are illustrated in examples (3a) and (3b) in §3.

Another type of deviation from biuniqueness is opaque allomorphy. It is discussed in §6.

## 5.5 Subspan repetition

This section presents different types of subspan repetition strategies in Yukuna. These strategies allow the creation of complex sentences, formed by repeating spans from the planar structure. In this sense, an element found only once in the planar structure can appear twice in a complex sentence. Crucially, because the repeated spans form a single sentence, they also display signs of forming a grammatical unit, as they share elements that scope over the entire sentence.

Complex sentences in Yukuna are mostly formed via the use of nominalizers, postpositions, and other subordinating markers. There are many such markers in the language, but they can be grouped into four types, depending on the structural features of the dependent clause. Each of these four types corresponds to a different diagnostic, and some diagnostics are further fractured into various tests, for a total of six tests. For more details on complex sentences in Yukuna, see Lemus Serrano (2020).

### 5.5.1 Complement clauses with lexical nominalizers (10-11); (10-17); (10-21)

The first subspan repetition diagnostic includes complement clauses with lexical nominalizers *-kana* and *-kaje*. These sentences are formed by placing a nominalization (as *motho'-kána* ‘the act of cooking’ in example (14)) in the position of the object argument (in brackets) of a complement-taking predicate (as ‘finish’ in examples (14) and (15)). This verb category includes aspectual predicates (start, finish), modal predicates (want), and perception and cognition predicates (see, hear, know). Complement clauses with lexical nominalizers typically require subject co-referentiality, so that while there is no subject marking on the dependent element, it is understood to be the same as in the main predicate as in (14).

- (14) *ru=      ñapáchi -ya [na= motho' -kána]*  
       9=      10            13 -      10      -  
       3SG.F= finish    PST 3PL= cook    NMLZ  
       ‘She finished the cooking of them.’

ycn0189,20

- (15)
- na= ñapáchi -ya [rikhá]*

9= 10 13 21

3PL= finish -PST 3SG

'They finished it.'

ycn0151,21

Verb roots marked with lexical nominalizers lack all inflectional features of verbs (subject indexation in position 9, and all markers from positions 12 to 16). However, they can still show valency marking, given that their use is lexically determined, as illustrated in the previous section with the marker =o MID. Since valency markers are placed in two discontinuous positions of the planar structure (11 and 17) this diagnostic is fractured into two.

The first fracture identifies a small layer including the verb core in 10 and the valency markers in position 11, as shown in example (16). The second fracture of this diagnostic identifies a larger layer, from the verb core 10, up to the middle marker in position 17. This is illustrated with example (17), where the middle marker =o is an integral part of the discontinuous stem 'study', placed after the event nominalizer -kaje.

- (16)
- majó pi= wakára'a [ri=la'jówa pa' -tá -kana phiyúké]*

3 9= 10 - 10 -11 - 21

here 2SG= order 3SG=ornament return -CAUS -NMLZ entirely

'Order them to bring all of their ornaments over here.'

ycn0018,5

- (17)
- eyá nu= ñapáta [jewíña' -kaje =o]*

1 9= 10 10 - 17

then 1SG= finish transform -NMLZ =MID

'Then I finished studying.'

ycn0018,5

Lastly, a third fracture of this diagnostic is required, given that lexical nominalizations in Yukuna show a hybrid structure, with features of both the nominal and verbal complex. For instance, these nominalizations show possessor marking of the P argument (similarly to English action nouns like in 'the destruction of the city') as in example (14). However, as is clear from example (16), they may also include adverbs and postpositional phrases, which are absent from the nominal complex. The only way to capture this was by applying a third fracture, identifying a span from the verb core (10) up to verbal obliques (21). The presence of oblique arguments in this subspan repetition strategy is illustrated with (18), where the oblique *ri=jló* 'for him' can only be analyzed as part of the verbal complex of the verb stem 'kill', and not of the main verb 'start'. In contrast, the

noun *pú'ju-na* is a possessor, and not a verbal argument, and thus it is not considered as being part of the span from the verbal planar structure identified by the diagnostic.

- (18) *Ri= keño' -chá [pú'ju-na nó -kana ri=jló].*  
 9= 10 -13 - 10 - 21  
 3SG= start -PST rodent-PL kill -NMLZ 3SG=for  
 'Lit. He started the killing of *tintin* rodents for him.' ycn0053,15

### 5.5.2 Complement clauses with *-ka* (8-21)

There are three further subspan repetition strategies used as diagnostics. They differ from the strategy discussed above because they select much larger spans from the verbal planar structure. I use the term Clausal nominalization to refer to this type of structure, which is internally very clause-like, but are externally used as NPs. Clausal nominalizations with *-ka* are used in Different Subject complement clauses, as well as in many types of adverbial subordinate clauses combined with postpositions. This diagnostic identifies a span from positions 8 to position 21. Clausal nominalizations with *-ka* are thus internally almost identical to main verbal clauses, although they lack some of their features (word order freedom, focused constituents, mood marking), but similarly to NPs, they can freely combine with postpositions. The layer identified by this diagnostic is shown in brackets in (19).

- (19) *Kája na= ka'á rikhá leyuná chojé, a'jné [na= ñapáta -ka rikhá]*  
 1 9= 10 21 21 21, 21 9= 10 - 21  
 then 3PL= throw 3SG pot into DIST 3PL= finish -NMLZ 3SG  
*ejená.*  
 21  
 until  
 'Then they throw it into a pot, until they finish it.' ycn0059,29

### 5.5.3 Same subject clause-chaining (10-21)

Another subspan repetition strategy is Same subject clause-chaining with gender and number markers. This diagnostic identifies a span from positions 10 to 21, including almost the entirety of the verbal planar structure, but leaving out all subject markers (person indexes in 9, and subject NPs in 5 and 8), as the Agent participant of the linked clause is understood as being co-referential with the

subject argument of the main clause. Example (20) shows this diagnostic with an instance of same subject clause-chaining (in brackets), where the chained-clause shows a verb core marked for tense and valency, with a postpositional argument. Typically, the choice of the gender and number suffix on the verb depends on the referent. So in (20), the suffix *-yo* F agrees in gender with the subject ‘she’ marked with the person index on the main verb.

- (20) *Pherú ké =ja ru= jecho’ -chá =o, [ja’ -chá -yo =o maloca*  
 3 9= 10 -13 =17 10 -13 - 17 21  
 quick like =EMPH 3SG.F= run -PST =MID fall -PST -F =MID house  
*jupichúmi éjó].*

old toward

‘Then she ran quickly, and (she) ran into the old house.’ ycn0151,105

#### 5.5.4 Adverbial clauses (8-21)

Lastly, the last subspan repetition diagnostic concerns adverbial clauses with adverbial subordinators *-chí* PURP, *-ré* PURP, and *-noja* CONC. The attested patterns suggest that this diagnostic identifies the same span of positions as clausal nominalizations with *-ka* used in complement clauses, from positions 8 to 21, as in example (21). Elicitation would be required to test whether all positions before 8 are excluded from this span. Given that complement clauses with *-ka* and the adverbial clauses presented here are nevertheless functionally and structurally different, I treat them as two different diagnostics (see Lemus Serrano 2020: 115–123).

- (21) *Marí na= a’ -chá nu= jló [pi= wáa’ -chí ri=jló tá me’tení*  
 3 9= 10 -13 21 9= 10 - 21 21  
 PROX 3PL= give -PST 1SG= to 2SG= call -PURP 3SG=to EMPH now  
*ilé=eyá].*  
 21  
 MED=from  
 ‘Here they’ve just given me (his number) for you to call him right now  
 from there.’ ycn0504,9

An important note concerning markers used in subspan repetition strategies in Yukuna is that some of them have grammaticalized as main clause markers as well. For instance, the markers *-chí* and *-ré* used as purposive subordinators are

synchronically distinct from the formally identical *-chí* and *-ré* used to encode first and third person imperatives in main clauses, although they diachronically originate in the same markers. In this latter function, they are included within the planar structure in Table 1 (position 16). However, as subordinators, they are not included within the planar structure, so in example (21), *-chí* is not numbered.

## 6 Indeterminate domains

This section presents diagnostics involving both the domains of phonology and morphosyntax, namely, free occurrence §6.1, and opaque allomorphy §6.2.

### 6.1 Free occurrence

This diagnostic identifies a span of positions, including the verb core, which can stand alone as an utterance. In Yukuna, this usually corresponds to lexical roots with their own lexical tone. Only alienable and non-possessible nouns can be used on their own as utterances without combining with any bound markers. Obligatorily possessed nouns and verbs minimally require indexation of possessors and subjects respectively.

This diagnostic is fractured into three tests depending on its interpretation.

#### 6.1.1 Free occurrence minimal (9-10) ; (8-10)

The first two fractures of this diagnostic seek to identify the smallest possible span of positions that could be used as an utterance. In Yukuna, a verbal utterance minimally requires a verb core with subject encoding, achieved either with a person index 9 or an overt subject NP 8. These two options mutually exclude each other, leading to two fractures: a layer including the verb core 10 with a subject person index 9 (22a), and a layer including the verb core 10 and an overt subject NP 8 (22b).

- (22) a. *Ri= iphí -cha.*  
           9=   10   -13  
           3SG= arrive -PST  
           ‘He arrived.’ ycn0041,32
- b. *Ri=i’rí iphí -cha.*  
           8       10   -13  
           3SG=son arrive -PST  
           ‘His son arrived.’ ycn0089,103



### 6.1.2 Free occurrence maximal (10-20)

The third fracture of this diagnostic identifies the largest span of positions including the verbal core that could stand as an utterance. This fracture identifies a span from positions 10 to 20. This span includes only one lexical element (the verb core) and its formatives, with no overt arguments nor obliques, as in (23).

- (23) *Pi= kapií -cha =o =jlá.*  
       9= 10 -13 =17 =19  
       3SG= lose -PST =MID =FRUST  
       ‘You almost got lost.’ ycn0058,66

Note that *=jlá* FRUST has its own lexical tone, and it does not undergo any phonological processes with adjacent forms, but it is still a bound form in the sense that it cannot be used on its own as an utterance, it is a formative of the verb complex.

## 6.2 Deviations from biuniqueness: opaque allomorphy

This section presents cases of opaque allomorphy in Yukuna, as opposed to non-opaque allomorphy, motivated by phonological constraints of the language (see §7.2). In Yukuna, I have identified two such cases, each used as a different diagnostic.

### 6.2.1 Valency allomorphy (10-13)

The first diagnostic, in test 18, concerns opaque allomorphy between the valency markers in position 11 and the tense marker *-cha* in 13. The three valency markers */-ta/* CAUS, */-ɲaa/* APPL and */-ka/* ‘associative’ undergo allomorphy when followed by past tense */-tʃa/*. Their last vowel changes from */a/* to */i/*, so */-ta/* becomes */-ti/* (phonetically produced either as [ti] or [tʃi]) (24-25), */-ɲaa/* becomes */-ɲai/*, and */-ka/* becomes */-ki/*. This instance of allomorphy is opaque. While it only targets morphemes ending in */a/*, it does not apply to all morphemes ending in */a/*, in fact, the negation marker */-la/* NEG does not trigger this allomorphy (26). Note that in examples (25-26) there are two forms of the past tense marker: */ja/* and */tʃa/*. This leads us to the second instance of opaque allomorphy.

- (24) *pa.pó.ta.he nu.khá*  
       /*pi= apô -ta -he nukhá/*  
       9= 10 -11 -13 21  
       3SG= wake.up -CAUS -FUT 1SG  
       ‘you will wake me up’ ycn0053,70

- (25) no.pó.tʃi.ja rik<sup>h</sup>á  
 /nu= apô -ti -ja rik<sup>h</sup>á/  
 9 10 11 13 21  
 1SG= wake.up -CAUS -PST 3SG  
 ‘I woke him up.’ ycn0089,103
- (26) uŋ.ká ri.p<sup>h</sup>á.lá.tʃa  
 /unká ri= ip<sup>h</sup>á -la -tʃa/  
 6 9= 10 -12 -13  
 NEG 3SG= arrive -NEG -PST  
 ‘he did not arrive’ ycn0041,130

### 6.2.2 Past tense allomorphy (10-13)

This diagnostic identifies the same set of spans as the former, from the verb core 10 to tense markers 13. In this instance of allomorphy, the past tense suffix /-tʃa/ changes its form if the preceding syllable (whether the last syllable of the verb core, or a valency marker) is phonologically /ti/ [tʃi], as in (24). Interestingly, the rule does not apply if the preceding syllable is phonologically /tʃi/ [tʃi], showing that this instance of allomorphy is entirely idiosyncratic. Indeed, the verb /itʃá/~/itʃí/ ‘dig’ combines with past tense /tʃa/ leading to the form /itʃí-tʃa/ ‘dug’ and not \*/itʃí-ja/.

## 7 Phonological domains

There are seven phonological processes used as constituency diagnostics in Yukuna, shown in Figure 4. This section presents each of these diagnostics, per phonological domain: diagnostics concerning tonal processes are discussed in §7.1, and segmental diagnostics in §7.2.

### 7.1 Tonal diagnostics

This section presents three tonal processes in Yukuna, namely, floating tone placement §7.1.1, tonal spreading §7.1.2, and tonal polarity §7.1.3.

#### 7.1.1 Floating tone placement: minimal (10-16) and maximal (9-16)

Yukuna’s tonal system is a zero vs. H system, with no low tone (underspecified) and two H tones: a spreading H tone (/H/, transcribed as  $\acute{V}$ ) and a non-spreading

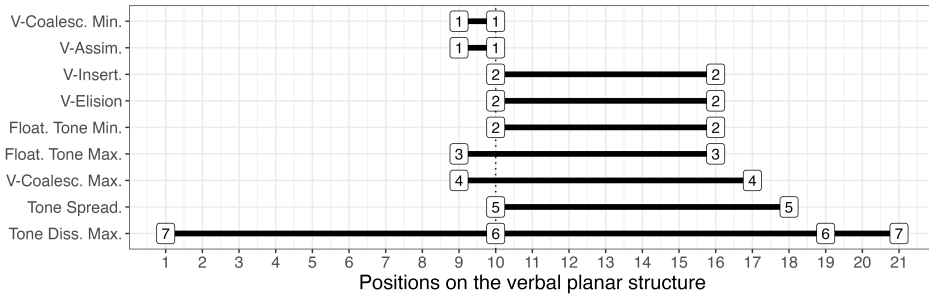


Figure 4: Phonological tests per converging layers

H tone (/HL/, transcribed as  $\hat{V}$ ). Each of these tones can be either bound to a specific TBU, or floating (transcribed as  $V^H$  if spreading, and as  $V^{HL}$  if non-spreading). Inherent tone is a feature of lexical roots, and most bound formatives are toneless, but as detailed next, the surface manifestation of H tones varies depending on the tone type of the root, and on whether the root combines with certain formatives or not (see Lemus Serrano et al. 2021 for a description of Yukuna's tonal system). This system is very similar to that found in 'pitch-accent' languages like Japanese, where verb roots are divided into accented and unaccented roots, and the position of the pitch drop depends on the suffixes that follow (Kawahara 2015).

Floating tones usually attach rightward to the first bound marker placed after the verb core. This diagnostic identifies a span from positions 10 to 16. With most verbs, this process only concerns what follows the verb core regardless of syllable count, as shown in (27), where the root's floating tone (/k<sub>a</sub><sup>HL</sup>/ 'throw') surfaces immediately after the root, on the second syllable of the group.

- (27) [k<sub>a</sub>.ké.ha]  
 /k<sub>a</sub><sup>HL</sup> -keha /  
 10 -  
 throw -PTCP  
 'thrown'

ycn0108,272

However, with a small subset of verbs, the presence or absence of a person index affects the placement of floating tones. This leads to a diagnostic fracture, as this latter application of the floating tone test identifies a larger span of positions, from subject indexes in 9 to imperative markers in 16. For instance, the floating tone of the root /l<sub>a</sub><sup>HL</sup>/ 'do' attaches to the third syllable counting from the left, so the /HL/ tone attaches to the first marker after the core if there is a person

index such as in (28a), but it attaches to the second marker after the root if there is no person index such as in (28b).

- (28) a. [ri.lə .lá.tʃa]  
           /ri= l<sup>a</sup><sup>HL</sup> -la -tʃa/  
           9= 10 -12 -13  
           3SG= do -NEG -PST  
           ‘he did not do’ ycn0041,155
- b. [lə .la.pó]  
           /l<sup>a</sup><sup>HL</sup> -la -po/  
           10 -12 -15  
           do -NEG -PL  
           ‘(they) did not do.’ ycn0117,93

This diagnostic excludes markers beyond position 16. Example (29) illustrates the behavior of the verb root /a<sup>H</sup>/ ‘give’, followed by two bound markers in each case, so the number of syllables is the same. The counting rule applies when the root combines with markers up to position 16, so that in (29a) the floating tone attaches to third syllable from the left, corresponding to the marker /-ri/ M in position 15. In contrast, in (29b), the floating tone attaches to the second syllable from the left, corresponding to the tense marker in position 13, since the marker /=mi/ PFV in position 18 is excluded from the span identified by this diagnostic.

- (29) a. a .he.rí  
           /a<sup>H</sup> -he -ri/  
           10 -13 -15  
           give -FUT -M  
           ‘(he) will give’ elicited
- b. a.tʃá.mí  
           /a<sup>H</sup> -tʃa =mi/  
           10 -13 =18  
           give -PST =PFV  
           ‘have given’ elicited

Note that the perfective marker /=mi/ in (29b) is produced with an H tone. This is due to the fact that despite being excluded from the domain of floating tone attachment, this marker is included in the domain of tonal spreading, which leads us to the final tonal diagnostic.

### 7.1.2 Tonal spreading (10-18)

In Yukuna, once attached to a given syllable, /H/ tones spread one syllable rightward, from the verb core in 10 up to the perfective marker in 18. Beyond this position, markers have very different tonal features. For instance, =*jlá* FRUST in position 19 has its own lexical tone, and in the same position, =*no* HAB displays tonal polarity, so even if preceded by a spreading /H/, it will be produced with low tone. Consider the pair of examples in (30). In (30a) the floating /H/ tone attaches to the first marker after the verb core and spreads rightward to the nominalizer /-ka/. In contrast, in (30b), the /H/ tone does not spread rightward, as the following marker /=no/ is beyond the domain of tonal spreading.

- (30) a. *rì.hĩ.ʃá.ká.no*  
           /ri= *ĩhĩ<sup>H</sup>* -*ʃa* -*ka*    =*no*/  
           9=   10   -13   -           =19  
           3SG= go   -PST -NMLZ =HAB  
           ‘his constant going’ ycn0189,41
- b. *rì.hĩ.ʃá.no*  
           /ri= *ĩhĩ<sup>H</sup>* -*ʃa*   =*no*/  
           9=   10   -13   =19  
           3SG= go   -PST =HAB  
           ‘he constantly goes’ ycn0041,12

### 7.1.3 Tonal polarity minimal (10-19) and maximal (1-21)

In both (30a) and (30b), =*no* HAB is preceded by an H tone, and it is produced with a low tone (untranscribed). Because the marker =*no* is also a promiscuous one, placed in several positions of the planar structure, the process of tonal polarity also applies no matter what element precedes it. See for instance its use in (8), where =*no* appears after a postpositional phrase in position 21. In the absence of any evidence showing that there is a limit beyond which tonal polarity does not apply, this diagnostic is considered as identifying the entire planar structure, from position 1 all throughout position 21.

## 7.2 Segmental diagnostics

There are five segmental processes used as constituency diagnostics in Yukuna, discussed next.

### 7.2.1 Vowel coalescence minimal (9-10)

The first process applies at the juncture between person indexes in position 9 and the verb core in 10. When combined with a vowel initial verb root, the final vowel of the person index and the vowel of the verb root undergo vowel coalescence. The result of this process may be a vowel of a different quality than the two vowels involved, as in (31-32).

- (31) no.há  
       /nu= ahá/  
       9= 10  
       1SG= fly  
       ‘I fly’ elicited
- (32) wɛ.ɲa.hé  
       /wa= ɲa<sup>H</sup> -he/  
       9= 10 -13  
       1PL= go -FUT  
       ‘we will go’ ycn0063,182

### 7.2.2 Vowel coalescence maximal (9-17)

The minimal application of this diagnostic clearly identifies a span from 9 (person indexes) to 10 (verb core), as shown in (31-32). However, there is no evidence that this process does not apply further, beyond the verb core and the formatives that follow it, as it requires a polymorphemic vowel sequence but most markers placed after the verb core have a CV shape. The only exception is the middle marker =o in position 17, and this marker does not undergo vowel coalescence with the preceding vowel, but another phonological process (vowel elision, §7.2.5). As such, in the absence of evidence of non-application of this process before position 17, we consider that one possible interpretation of the vowel coalescence diagnostic identifies a maximal span from positions 9 up to 16, outside of which this process does not apply.

### 7.2.3 Vowel assimilation (9-10)

The second segmental process also applies at the index-core juncture, and concerns person indexes ending in /u/ (nu= 1SG and ru= 3SG.F) and verb cores where the first vowel nucleus is /o/. The vowel /o/ of the verb core triggers total assimilation, turning the /u/ of the index into [o], in a process of vowel harmony, as in (33a).

- (33) a. ro.nó.tʃa  
           /ru= nó -tʃa/  
           9= 10 -13  
           3SG.F= kill -PST  
           ‘she killed’ elicited
- b. ne.nó.tʃa  
           /na= nó -tʃa/  
           9= 10 13  
           3PL= kill -PST  
           ‘they killed’ elicited

This process also affects person indexes ending with /a/ (/wa=/ 1PL and na=/ 3PL), turning the /a/ into an [e], in a process of partial assimilation (33b). Note that the marker following the verb core /-tʃa/ also ends with /a/ and it is not at all affected by the assimilation process. The vowel harmony diagnostic thus only has one application, identifying a small layer (9-10).

#### 7.2.4 Consonant aspiration minimal (9-10) and maximal (9-12)

Yet another process at the index-core juncture is consonant aspiration. This process applies between all person indexes and roots starting with /h/, wherein the /h/ of the root merges with the onset of the person index, and leads to either an aspirated plosive [C<sup>h</sup>] or a voiceless sonorant [C̥], as in (34a-34b) respectively.

- (34) a. p<sup>h</sup>ə.pá  
           /pi= həpá/  
           9= 10  
           2SG= walk  
           ‘you walk’ ycn0053,83
- b. ɲəpá  
           /nu= həpá/  
           9= 10  
           1SG= walk  
           ‘I walk’ ycn0018,7

Note that once the process of consonant aspiration applies, leading to the merger of the two onsets from the person index and the verb core, the process of vowel coalescence applies, leading to the merger of the two vowel nuclei as well. In (34b), the vowels /u/ and /ə/ merge, leading to a creaky [ɔ̰]. As the process

of consonant aspiration requires a polymorphemic /CV.hV/ syllable sequence to take place, once more, the only evidence that it is not applying beyond the index-root juncture is that the markers /-he/ FUT and /-hika/ FAR.HAB in position 13 also start with /h/ and they are not affected by this process. This leads to a diagnostic fracture, where the maximal interpretation of this diagnostic identifies a span from positions 9 to 12.

### 7.2.5 Vowel elision (10-16)

The following process, vowel elision, applies whenever the middle marker =o is used, as this marker systematically replaces the preceding vowel of any element from the verb core in position 10 (example (35a)) up to the imperative markers in position 16 (example (35b)).

- (35) a. nu.ju.ró  
       /nu= jurî =o/  
       9= 10 =17  
       1SG= stay =MID  
       ‘I stay’ ycn0018,8
- b. pi.ɲa.ní.ɲo  
       /pi= ɲa<sup>HL</sup> -niɲa =o/  
       9= 10 -16 =17  
       2SG= escape -PROH =MID  
       ‘Don’t escape!’ ycn0041,68

### 7.2.6 Copy vowel insertion (10-16)

Lastly, the process of copy vowel insertion applies when a root ending with a creaky vowel is not followed by any markers up until position 16. In such cases, if the root is not followed by any markers at all, a copy vowel of the same quality of the creaky vowel is inserted, as in (36a). The same process applies when the root is followed by markers from positions 17-20, such as =no in position 19 (36b).

- (36) a. pi.kaʔ.á  
       /pi= kɑ<sup>HL</sup> /  
       9= 10  
       2SG= throw  
       ‘you throw’ ycn0058,101



- b. naʔ.á.no  
 /na=  $q^H$  =no/  
 9= 10 19  
 3PL= give =HAB  
 ‘they always give!’ ycn0068,181

If the root is followed by a marker before position 17, copy vowel insertion does not take place (37). Note that the inserted copy vowels in (36) carry the H tone, while in (37) it is the marker *-nija* PROH that carries the H tone.

- (37) pi.kə .ní.ja  
 /pi=  $kq^{HL}$  -nija/  
 9= 10 -16  
 2SG= throw -PROH  
 ‘Don’t throw!’ ycn0058,101

As a constituency diagnostic, this process identifies a span from positions 10 to 16. This process is tightly connected to the process of floating tone association, discussed previously §7.1. Indeed, the right edges of the two tests on floating tone association converge on position 16. Beyond this position, it is necessary to insert a copy vowel for roots ending in a creaky vowel, as a support for the floating tone, which cannot attach to the markers to its right (as in (36b)). However, this process is not related to minimality constraints in the language. There are multiple monosyllabic lexical roots that do not undergo such a process as they do not end in a creaky vowel (e.g. the verb /nó/ ‘kill’, the indefinite pronoun /ná/ ‘what/who, something/someone’).

As detailed in this section, most segmental processes in Yukuna are motivated by hiatus avoidance. Indeed, there are no instances of cross-morph hiatus in the language at all. Morpheme internally, however, vowel sequences (analyzed as adjacent vowel nuclei, and not as diphthongs) are attested, and even sometimes spontaneously produced by optional intervocalic elision in speech.

## 8 Summary and discussion

In total, out of the 30 tests applied to a planar structure of 21 positions, a total of 17 different layers were identified. That is to say that there were many instances where several tests fully converged in identifying the same spans of positions from the planar structure, although they did not always identify the same span.

In terms of the identified edges, the results show that, quite unsurprisingly for a mostly “suffixing” language, there is little variation on the left edge of the layers, with most tests starting either in positions 9 (person indexes) or 10 (verb core), while there is a lot of variation concerning the right edge (five tests end in position 10, five tests in position 16, three in position 18, two in position 20). This is illustrated in Figure 5, where the “left” and “right” boxes provide the left and right edges of tests, respectively. The Y axis provides the position of the template at which a test starts/ends (from 0 to 20). The X axis provides the number of tests that converge at a given edge. The main left and right edges appear in green.

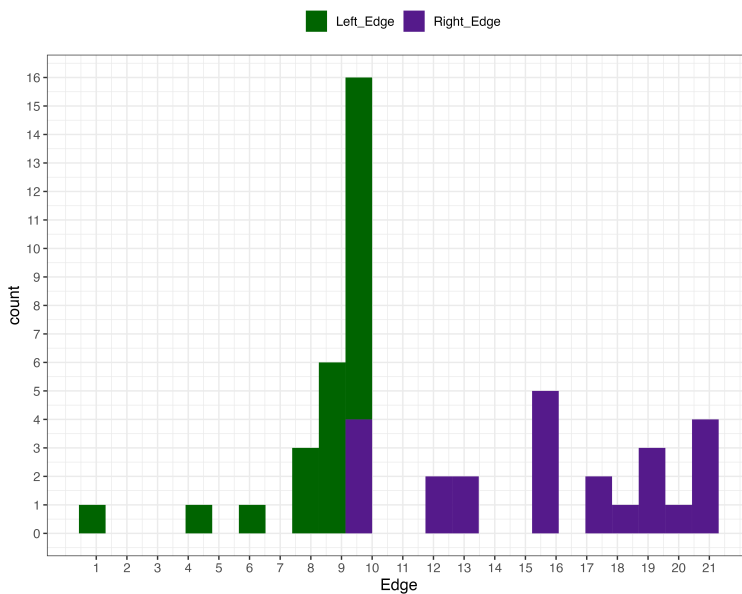


Figure 5: Left and right edges of constituency tests

What these results suggest is that the left edge of the verb complex is quite unproblematic as opposed to the right edge, which is less easily identified. From an Arawakan perspective, it is worth noting the ambiguous status of person indexes in the language, placed in position 9 right before the verb core. In Yukuna, the morphosyntactic status of person indexes is rather ambiguous, as they are included within roughly half of the spans identified by constituency tests, and excluded from the other half. Phonologically, they are bound to the core that they combine with, but less so that bound markers placed after the core, as most tonal processes apply from the core rightward (see §7.1). Syntactically, they function as pro-indexes, as they are mutually exclusive with nominal subjects and possessors, which are rigidly placed before the core as well.

The synchronic features of person indexes in Yukuna fit very well with the diachronic scenario posited by Danielsen (2011) for the southern branch of the family, wherein person indexes are said to come from a single set of free person forms (reconstructed for proto-Arawakan by Payne 1991) that grammaticalized into bound forms in the different languages, with varying degrees of morphophonological fusion with the verb core depending on the language.

Next, in terms of rates of span convergence, that is to say, the degree to which different diagnostics identified the exact same span from the planar structure, we note that in Yukuna, out of all the 17 layers identified, only three had a convergence rate above two. The three identified layers are presented in Table 2, with their size, convergence rate, and the domains of the tests that identify them (phonological or other and morphosyntactic plus indeterminate tests).

Table 2: Layers, convergence, and domains

layer	left edge	right edge	convergence	main domain
1	10	16	3 tests	PH
2	9	10	4 tests	PH
3	10	17	3 tests	MS

When we break down the convergence rate per test domain, we note that layers 1 and 2 are mostly identified by language specific phonological tests, while layer 3 is mostly identified by cross-linguistically common morphosyntactic constituency tests. At such low convergence rates (of 3-4), it is unclear whether these different layers could be analyzed in terms of a split between the phonological (layers 1-2) vs. grammatical words (layer 3). In fact, note that layers 1 and 3 are almost identical in their identified spans (layer 1 excludes the middle voice marker in 17, while layer 3 includes it), despite being identified by tests from different domains. To contrast these results with the practical definition of words in Yukuna adopted in the grammar sketch, layer 1 corresponds to what I considered to be the phonological and grammatical word in the language, with all remaining formatives of the verb complex being classified as ‘clitics’. However, as the results of this methodology have shown, there is no real reason to consider layer 1 as being a more valid word candidate than layer 3. In other words, the results do not clearly point in the direction of a bisection between phonological vs. grammatical words, with high convergence of tests within each domain but not across domains; nor do they suggest a single wordhood candidate with high convergence of tests across all domains. What the results do suggest, however, is that

there is a split between the core and its preceding formatives on one hand (layer 2), and the core and its following formatives on the other (layers 1 and 3).

In terms of layer size, the results show that layers in Yukuna vary from a very small layer of two positions (layer 2) including the person indexes and the verb core, to a larger layer of eight positions (layer 3), which excludes person indexes and includes postposed verb formatives up until position 17. Given the differences in layer size and low degrees of convergence, it is hard to provide a precise account of the degree of synthesis of Yukuna, but tentatively using the largest layer in Table 2 (layer 3) as the maximally inflected verb form would give us a total of 7 categories per word (counting valency markers twice, in position 11 and 17). This score places Yukuna as a language with a moderately high degree of synthesis (compared with the scales in Bickel & Nichols 2013), in contrast with languages in the higher ends of the scale (with 12 to 13 cpw). It is interesting also to contrast the maximally inflected verb form with a frequency-based definition of synthesis. Impressionistically, verbs in Yukuna show a strikingly low degree of synthesis, with most verbs showing only 2 to 3 marked categories per word (typically, person, tense and valency). On this account, Yukuna can be considered as a language with a moderately low degree of synthesis. These results fit well with the widespread idea according to which Northern Arawakan languages are less synthetic than those from the Southern branch (Aikhenvald 1999).

Finally, in terms of orthographic conventions used by Yukuna speakers, the orthographic word broadly corresponds to layer 3 (positions 9 to 20), including the verb core and all of its formatives, regardless of their phonological properties (frustrative *jlá* in position 19 is written within the same orthographic word despite having its own tone).

## Acknowledgements

I send my deepest thanks to all the Yukuna and Matapí collaborators who have patiently taught me their language, in particular, to Virgelina Matapí Yucuna. I address my sincere gratitude to Adam Tallman for his support in the application of his methodological framework to the study of Yukuna constituency. I thank ELDP and Labex ASLAN for their financial support and lastly, I wish to thank the editors of this volume for inviting us to be part of this project.

## Abbreviations

APPL	applicative	MED	medial
CAUS	causative	MID	middle
CONC	concessive	NMLZ	nominalizer
DEM	demonstrative	NVNEG	non-verbal negation
DIST	distal	PFV	perfective
F	feminine	PROH	prohibitive
FAR	far past	PROX	proximal
FRUST	frustrative	PST	past
HAB	habitual	PTCP	participle
INDF	indefinite	PURP	purposive
M	masculine	SG	singular

## References

- Aikhenvald, Alexandra Y. 1999. The Arawak language family. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian languages*, 65–102. Cambridge: Cambridge University Press.
- Bickel, Balthasar & Johanna Nichols. 2013. Chapter 22: Inflectional synthesis of the verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://wals.info/chapter/22>.
- Danielsen, Swintha. 2011. The personal paradigms in Baure and other Southern Arawakan languages. *International Journal of American Linguistics* 77(4). 495–520. DOI: 10.1086/662154.
- Fontaine, Laurent. 2001. *Paroles d'échange et règles sociales chez les Indiens yukuna d'Amazonie colombienne*. Université de la Sorbonne nouvelle - Paris III. (Thèse de doctorat). <https://tel.archives-ouvertes.fr/tel-00596637/document> (30 December, 2014).
- Hammarström, Harald, Robert Forkel & Martin Haspelmath. 2019. *Glottolog 4.1*. Jena: Max Planck Institute for the Science of Human History. <https://glottolog.org/>.
- Kawahara, Shigeto. 2015. The phonology of Japanese accent. In Haruo Kubozono (ed.), *Handbook of Japanese Phonetics and Phonology*, 445–492. De Gruyter Mouton. DOI: 10.1515/9781614511984.445.
- Lemus Serrano, Magdalena. 2016. *Observations sociolinguistiques et description phonologique du yukuna : Langue arawak de l'Amazonie colombienne*. Université Lumière Lyon 2. (Mémoire de Master 2).

- Lemus Serrano, Magdalena. 2020. *Pervasive nominalization in Yukuna, an Arawak language of Colombian Amazonia*. Université Lumière Lyon 2. (Ph.D. dissertation).
- Lemus Serrano, Magdalena, Marc Allasonnière-Tang & Dan Dediú. 2021. What conditions tone paradigms in Yukuna: Phonological and machine learning approaches. *Glossa: a journal of general linguistics* 6(1). 60. DOI: 10.5334/gjgl.1276.
- Payne, David. 1991. A classification of Maipuran (Arawakan) languages based on shared lexical retentions. In Desmond Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, vol. 3, 355–499. Berlin: Mouton de Gruyter.
- Ramirez, Henri. 2001. *Línguas Arawak da amazônia setentrional: Comparação e descrição*. Manaus: Editora de Universidade do Amazonas.
- Schauer, Junia, Stanley Schauer, Eladio Yukuna & Walter Yukuna. 2005. *Meke kemakánaka puráka'aloji: Wapura'akó chu, eyá kariwana chu*. Bogotá: Fundación para el Desarrollo de los Pueblos Marginados.
- Tallman, Adam J. R. 2021. Constituency and coincidence in Chácobo (Pano). *Studies in Language* 45(2). 321–383. DOI: 10.1075/sl.19025.tal.