

# Chapter 13

## Graded constituency in the Araona (Takana) verb complex

Adam J. R. Tallman

Friedrich-Schiller-Universität Jena The National Museum of Folklore and Ethnography - La Paz

This paper provides a description of the verb complex in Araona. There are three layers of structure that show relatively high convergence of logically distinct constituency tests or phonological domains. It is unclear which of these layers should be regarded as the “word”, thus making it unclear whether the language should be regarded as isolating or polysynthetic (or something in between). The results of applying constituency tests following the planar-fractal method suggest a much more graded and complex situation than has been described for Takanan languages thus far.

### 1 Introduction

This chapter provides a description of the verb complex of Araona, an endangered Takanan language, spoken in the Amazonian part of the department of La Paz, Bolivia. I provide a description of the internal constituency of Araona’s verb complex by means of the methodology proposed in Tallman (2021b), the planar-fractal method.

The results of this study show some support for the notion of “word” used in Takanan studies insofar as convergences in constituency variables are assumed to be markers of candidate word constituents. However, based on the results there are actually three possible word candidates; (i) one which contains just the verb root; (ii) one which corresponds roughly to the “word” used in other descriptions; (iii) one which includes the entire predicate excluding the object NP.



Section §2 provides a brief background on the Araona speech community, the fieldwork context and the data for this chapter. Section §3 provides a description of the verbal planar structure. Section §4 discusses free occurrence tests and morphosyntactic tests. Section §5 discusses phonological domains. Some of the phonological domains could be considered morphosyntactic, an issue which I also discuss. A final section (§6) summarizes the results and contextualizes them with respect to the general Takanan literature.

## **2 Araona language, speakers and fieldwork**

### **2.1 Speakers and fieldwork**

Araona is of the Takanik branch of the Takanan language family. It is thus most closely related to Tacana and Reyesano (also known as Maropa) out of the Takanan languages (Girard 1971). Araona is spoken in 5 communities on the Manupari (literally ‘first river’) river (Palma Sola, Barero, Puerto Araona, Peñal, Baranco) and 3 communities of the Manurimi (literally ‘second river’) river (Chacra, Los Angeles, Pampa Alegre), with a total of approximately 150 speakers. The rivers are located in Iturralde province in the department of La Paz, most accessible by docking points that start from rivers in the department of Pando (via Sena or Cobija). All adult Araona speak the language fluently. Based on my own observations in the field and commentary from the Araona themselves, there is some variation in the fluency of younger generation of speakers. Some children on the Manurimi river seem to only have passive knowledge of the language, but the language is still being learnt by children on the Manupari river.

Data were gathered during three trips to Bolivia in 2016 (three months), 2019 (six months), and 2022 (4 months). Additional data have been gathered through correspondence over WhatsApp voice recordings starting in July of 2019 while I was in Germany (see Neely (to appear) for a description of methodology and workflow of online fieldwork). All data gathered from WhatsApp were rechecked with Araona speakers in person in Bolivia in 2022. Previous work on the language was done by Summer Institute of Linguistics (SIL) missionaries Donald and Mary Pitman with a few short and sketchy analyses (Pitman & Pitman 1976, Pitman 1980, Pitman 1981, Pitman & Pitman 1970). SIL publications on the language consist mostly of translations of biblical hymns and evangelical christian myths into the language as the SIL missionaries were mainly concerned with

evangelizing the Araona.<sup>1</sup> I do not have access to any audio recordings of the SIL missionaries if these exist and thus I will rely little on data from the missionaries. Hebe González published an important phonological sketch of the language (González 1997). Emkow published a dissertation length grammar of Araona (Emkow 2006, 2019), however, none of the texts on which this dissertation is based are available.

The current chapter is written in the context of an ongoing documentation project of the Araona language, funded primarily by the Endangered Language Documentation Fund, and initially by Labex ASLAN (Université de Lyon, Advanced Studies on Language Complexity). Fieldwork was conducted in the communities of Chacra, Los Angeles, Pampa Alegre, Barero and Puerto Araona. Some fieldwork sessions were also done in the towns of Rampla and Sena (department of Pano) and Riberalta (department of Beni, where the Araona frequent for political, economic and medical reasons). The documentation project currently contains about 17 hours of transcribed and translated texts (Tallman 2021a). This corresponds to about 10,000 transcribed, translated and annotated sentences of naturalistic speech. There are some short texts from the SIL missionaries which total about 600 sentences (Pitman & Pitman 1980). I do not make extensive use of these texts, because I do not have access to corresponding recordings to verify the accuracy of the transcriptions, particularly with reference to the pitch accent patterns, which are important for the current study. While the SIL missionaries did mark “stress” patterns, it is not always clear what the physical meaning is of their accentual markings.

Of potential relevance to linguistic studies on Araona is that the Araona were traditionally split into moieties, one called the Araona (/alaona/) the other called the Cabiña (/kambija/). Preferential marriage practices of the Araona were such that an Araona was always supposed to marry a Cabiña, a practice which has eroded since there are now few Cabiña left. The Araona frequently suggest that the Cabiña spoke a different variety of Araona from them, or perhaps a different language all together. Younger speakers such as Oscar Matawa do not note a significant difference between the way they talk and the contemporary Cabiña. At this point our understanding of the Araona language is not detailed enough to be able to pinpoint social variables that might be conditioning speaker variation. I limit myself to pointing out speaker variation where it exists.

---

<sup>1</sup>Note that the Araona themselves report violent confrontations with the missionaries related to the missionaries’ attempts to purge traditional beliefs and practices from the their society (Tallman 2021a). Where evangelization efforts have been more successful (on the Manurimi) there has also been a decline in linguistic vitality, as culturally relevant discourse practices have necessarily declined.

## 2.2 Araona language and data presentation

Araona has four vowels (/i, e, a, o/) and 20 consonants (/p, t, k, kʷ, b, d, ts, tʃ, dʒ, s, ç, z, w, l, m, n, j, h, j, ?/). In what follows, I generally use the Araona practical orthography. In some cases I will also refer to a more narrow phonetic transcription where the IPA will be used. In the partially Spanish based practical orthography /ç/ = <sh>; /j/ = <y>; /h/ = <j>; /tʃ/ = <ch>; /dʒ/ = <dy>. In the practical orthography the glottal stop is (often) represented by a space even if this breaks up a single morpheme (e.g. ni o ‘tapir’). I will use /?/ for the glottal stop rather than following the practical orthography so that the Leipzig glossing rules can be followed. <ni o> will thus be written as /ni?o/ in this study. Otherwise I follow the practical orthography except for when it is useful to have a surface transcription. The phoneme /l/ appears to be realized as [r] or [l], depending on the speaker, but the question requires future research.

In general, accent in Araona is predictable, falling on the second syllable of its domain of application by default (Pitman & Pitman 1976). However, as will be described below, the placement of the pitch accent depends on its intonational context. There are also some phonological and morphosyntactic environments where the pitch accent “shifts” to the first syllable.

Araona displays an ergative case marking system, realized on full NPs and in pronouns. NPs and pronouns display free constituent order in the sense that the order of A, S, or P can occur in any order in relation to the verb. However, dependent clauses with verbal predicates are always verb final (Emkow 2006, 2019)

The verbal word in Araona is described as being fairly complex (Pitman 1980) or “polysynthetic” (Emkow 2019), as it is with other Takanan languages (Guillaume forthcoming). The verb complex expresses a number of modal, expressive, tense, and various aspectual distinctions. As with other Takanan languages (Guillaume forthcoming), Araona is described as expressing a host of “lexically heavy” categories in its verbal word: associated motion, associated posture, temporal distance, time of day modifications, and manner semantics. The expression of these categories are typically described as “morphological”, but many could be seen as straddling the boundaries between morphology and syntax in interesting ways, since some of the markers are free forms and display various degrees of syntax-like variable ordering (variable ordering without conditioned scope). However, I should point out that this chapter does not provide a complete inventory and description of all the verbal modifiers, which would require a grammar length description. Rather I focus on those that are well represented in the current corpus. Because of this the analysis provided here should be regarded as preliminary.

Throughout this chapter all sentences will be marked with respect to whether they are from naturalistic speech, provided with an TXT code, whether they are from elicitation (ELIC). Sentences from elicitation can be divided into those where a speaker repeats a sentence offered to them and provides a grammaticality judgment (ELIC:ARA>SP), those translated from Spanish (ELIC:SP>ARA), those volunteered by the speaker (ELIC:ARA). The Spanish translation corresponds to the free translation provided by an Araona speaker or by me (when the example is the result of a Spanish to Araona translation) and the English corresponds closer to my interpretation. Following common practice, “\*\*” is marked on a sentence deemed unacceptable, “?” for a sentence where the speaker is unsure. Where appropriate, examples from naturalistic speech will be given. I provide examples from elicitation for expositional reasons and/or in cases where the corpus is not sufficient to make precise statements about certain syntagmatic facts of relevance.

### 3 Verbal planar structure

The verbal planar structure for Araona is provided in Table 1. For the most part I will motivate the details of the planar structure *in tandem* with a discussion of the results of the constituency variables. Below I provide some introductory comments on the verbal planar structure in Araona. While the verbal planar structures displays some overlap with the verb template provided in Pitman’s description (Pitman 1980: 108), a few extra positions and elements need to be added that the latter seems to consider outside of the verbal word. Positions where the inventory is relatively large are not filled out with any forms, but rather “...”. Instead tables are given with the relevant inventory of the most common and well understood morphemes in the discussion below. Note that NP stands for “noun phrase” as usual, PP stands for “postpositional phrase” and S stands for “Sentence” wherein another verbal planar structure can be inserted. There are no distributional differences between NPs of different grammatical roles. Araona grammatical roles are case marked. Case markers are not represented in the verbal planar structure.

The orthographic word provided in Pitman (1980: 108) and Emkow (2019) corresponds to the 4–15 span.<sup>2</sup> The 4–15 span also turns out to be the best verbal word candidate, as we will see. Pitman provides a template-like description of the ‘relative position of the verbal radicals and affixes’ (my translation)<sup>3</sup>, which

<sup>2</sup>Note that Pitman is inconsistent in this regard, as *isha* is sometimes represented as a separate word (Pitman 1980: 38).

<sup>3</sup>“Posición relativa del radical y afijos verbales” (Pitman 1980: 28)

Table 1: Verbal planar structure of Araona

Pos.	Type	Elements	Forms
(1)	slot	XP{NP, PP, S}, adverbials	...
(2)	zone	“P2”	...
(3)	zone	XP{NP, PP,S}, adverbials	
(4)	slot	“prefix”	...
(5)	slot	noun	
(6)	slot	core verb root	
(7)	zone	motion, time of day, aspectual, manner, root	- <i>pe</i> , ...
(8)	slot	affectionate, small	- <i>shodi</i> , - <i>limi</i>
(9)	slot	with difficulty, almost	- <i>sa(wa)</i>
(10)	slot	interactional	- <i>ti</i>
(11)	slot	causative, completive	- <i>eme</i> , - <i>pe</i>
(12)	slot	3A	- <i>ta</i> , - <i>me</i>
(13)	slot	finality marker	- <i>ibo</i> - <i>iba</i>
(14)	slot	tense, aspect, posture, negation, wandering, clause-linkage	...
(15)	slot	limitative, again	- <i>we</i> , - <i>isha</i>
(16)	slot	auxiliary	
(17)	slot	connector	<i>tsio</i> , <i>po</i>
(18)	zone	XP{NP, PP, Adv}	

the planar structure in Table 1 builds on. Pitman’s template is provided in Table 2. The positions of the planar structure that correspond to Pitman’s template are added underneath.

Table 2: Pitman’s analysis of Araona

Edge 1a	Root	Root	Aspect suffixes	Margin 2	Margin 3	Edge 1b
Prefix	Noun	Verb	Time, Manner, Locative, Attitude	Voice, - <i>ti</i> , - <i>ta</i>	- <i>ibo</i>	Tense, Mode
4	5	6	7–9	10, 12	13	14

There are four main differences between the verbal planar structure and Pitman’s analysis; (i) The planar structure contains positions outside the mission-

ary imposed “word” which are 1–3 and 16–18; (ii) the position for “aspectual suffixes” is split into three positions 7–9 to capture fixed ordering, not originally captured by Pitman; (iii) an extra position is added to capture the distribution of the causative *-eme* and the completive *-pe* in relation to *-ti* ‘interactional’ and *-ta* ‘third person A, third person plural A/S’; (iv) position 15 is added to capture the relative position of *-isha* and *-we*.

First, we consider the positions from 1–3 and 18. As stated in §1, Araona has ‘free’ constituent order in the sense that NPs occupying different grammatical relations can occur in any order with respect to each other and with respect to the verb (position 6). The variable ordering of NPs is derived from the verbal planar structure by virtue of positions 3 and 18. Positions 3 and 18 are zones flanking the verb that allow NPs inside of them of any grammatical relation. Position 1 is reserved for constituent interrogatives, focused NPs and coordinated/subordinated clauses (technically “subspan repetitions”). The position 1 (as distinct from position 3) is motivated by the presence of a group of Wackernagel-like (or “P2” for position 2) morphemes (*tso* ‘prior event, anterior’, *sha* ‘dubitative, interrogative’, *tokwe* ‘dubitative’ and *pa* ‘reportative’) that occur after the first NP, clause, or adverbial as in 1, but not after the second or third NP as illustrated by the ungrammaticality of 2 (where *=tokwe=pa* occurs after the A and R arguments).<sup>4</sup>

(1)	<i>pona</i>	<i>tsidi-a</i>	<i>tokwe pa</i>	<i>wada</i>	<i>todi-lipi</i>	<i>laba</i>	<i>ti</i>	<i>-ta</i>
	1	-	2	-	3	-	-	6 -12
woman little-ERG EPIS REP 3SG:GEN child-DIM cracker give -3A								
<i>-iki</i>								
<i>-14</i>								
-RECP:PST								

‘(I believe and it is said that) the small woman gave a cracker to her child.’  
Sp. ‘Creo y se dice que una mujercita dió arepa a su hijo’ (ELIC: ARA > SP)

(2)	* <i>pona</i>	<i>tsidi-a</i>	<i>wada</i>	<i>todi-lipi</i>	<i>tokwe pa</i>	<i>laba</i>	<i>ti</i>	<i>-ta</i>
	1	-	3	-	2	-	3	6 -12
woman little-ERG 3SG:GEN child-DIM EPIS REP cracker give -3A								
<i>-iki</i>								
<i>-14</i>								
-RECP:PST								

‘(I believe and it is said that) the small woman gave a cracker to her child.’  
Sp. ‘Creo y se dice que una mujercita dió arepa a su hijo.’ (ELIC: ARA > SP)

<sup>4</sup>See (Guillaume 2016) for a similar category of Wackernagel particles in Tacana.

As stated above, the P2 morphemes occur after position 1. Position 1 can be a single NP, but does not have to be filled. This is why I refer to these morphemes as ‘Wackernagel-like’. For instance, position 1 can be empty as in (3) and (3b).

- (3) a. *pa tso naeda ba -odi*  
 2 2 3 6 -14  
 REP ANT 3.PL see -FREQ  
 ‘It is said that they were looking.’  
 Sp. ‘Dice que ellas buscan.’ (TXT 1138:0047)
- b. *tso pa ena-metse mo abeta a -ta -iki*  
 2 - 3 - - 6 -12 -14  
 ANT REP water-with FOC do-twice a-3A/PL  
 ‘They were already baptized twice (lit. they already gave him with water).’  
 Sp. ‘Ya hizo la muestra con agua dos veces.’ (TXT 1549:307)

Constituent interrogatives seem to obligatorily appear in position 1 as *jico* ‘where’ in (4).

- (4) *jico tso pa neti bewewe*  
 1 2 2 6 18  
 where 2 - stand/live now  
 ‘Where do they live now?’  
 Sp. ‘¿Donde viven ahora?’ (TXT 1549:267)

The position 1 element can also be an entire clause as shown from the examples in (5a) and (5b). The P2 morphemes and their glosses are listed in Table 3.

- (5) a. *awada piye -ti -wiki tso tsa -tseiye -sa*  
 3 6 -7 -9 17 6 -7 -7  
 tapir shoot -go.there -going:P PRIOR look.for -all.day -FRUST  
 -ja  
 -14  
 -RECP:PST  
 ‘After shooting the tapir (which then escaped), I looked for him all day in vain.’ / Sp. ‘Después de balear el anta, lo busqué casi todo el día en vano.’ (Pitman 1980: 52)

- b. *paicho najo kwaiya -ti -(i)bo tso kwae -ti*  
     3        /       6       -7       -13     2       6       -7  
     carayana beside arrive -go.there -RELEV PRIOR -INTRC -FINAL  
     -(i)bo -me e- a -pa  
     -13 -15 4- 6 -17  
     explain -there E- AUX REP  
     ‘They arrived among the carayana, and then conversed with them.’  
     Sp. ‘Llegaron, no sé donde, ande los carayanás, y conversaron.’ (TXT  
     1817:0391)

Table 3: P2 morphemes in Araona

Position	Gloss	Free translation	Morphemes
2	so	‘so’   Sp. ‘entonces’	<i>po</i>
	because	‘because’   Sp. ‘porque’	<i>pojo</i>
	anterior	‘already’   Sp. ‘ya’	<i>tso</i>
	reportative	‘it is said ...’   Sp. ‘dice que ...’	<i>pa</i>
	epistemic	‘I believe ...’ / Sp. ‘creo que ...’	<i>tokwe</i>
	like.so	‘in this way’ / Sp. ‘así ...’	<i>dipa</i>
	conjectural	‘is it true that ...’ / Sp. ‘será que’ ...	<i>sha</i>

I have found no ordering restrictions between the P2 morphemes themselves, although there is a strong tendency for the order *tso + pa* ‘anterior + reportative’ to indicate that the sentence has an interrogative force. Data from natural speech show no clear ordering constraints between pairs of P2 elements - given two P2 elements, I have found both orders for all pairs. For instance both the orders *tso dipa* and *dipa tso* are attested, both the orders *tokwe pa* and *pa tokwe* are attested etc., illustrated in examples (6a)-(6d). Assessing more complex ordering restrictions or constraints (e.g. between three P2 elements) remains an issue for future research.

- (6) a. *kwipa po tso dipa jazeze shoe dipa-kata kwada*  
     1       2       2       2       3       -       -       3  
     how   that ANT like.so parrot hear.from.far like.so- AUG  
     *pa-ba-neti tsawa-neti*  
     4-       6       -14 6       -14  
     1PL       POST-   see -stand spy -stand  
     ‘How far one can hear the parrot from far away, let’s go up and look  
     there, lets go to spy on it.’

- Sp. ‘Porque harto se escucha gritando el loro, vamos a (subir) para mirar para allá , vamos (arriba) para espiar.’ (TXT 0047:0091)
- b. *dipa tso jana ti -me -sa e -a*  
 2 - 3 6 -11 -7 4- 6  
 like.this ANT food give -CAUS -FRUST E -AUX  
 ‘One has to share the food’
- Sp. ‘Hay que repartir toda la comida.’ (TXT 0035:0031)
- c. *aise tokwe pa e-di-a*  
 1 2 - 4-6  
 someone EPIS REP E-eat-E  
 ‘I think someone ate it already.’
- Sp. ‘Creo que ya comió alguien.’ (TXT 1109:0029)
- d. *pa tokwe ba-sa-sha Jojo.esi*  
 3 - 6-7-14 18  
 REP EPIS see-FRUST-DIST:PST Tata\_Mayari  
 ‘I think it is said that Tata Mayari was the first to see it.’
- Sp. ‘El Jojo esi él fue primero a mirar creo.’ (TXT 1535:0405)

The prefix slot of position 4 is the same as that identified by Pitman. It corresponds to Edge 1a in Pitman’s description (see §3). Table 4 provides the elements of Araona’s prefix slot. As we will see, two of the elements of the prefix slot are actually pieces of circumfixes: *pi-* ‘negative’ and *ja-* ‘interactional’. The formative *e-* can also be regarded as a prefixal component of a number of markers realized in positions 6, 7 and 14. The analysis of *e-* in Araona is somewhat complicated by phonological issues, however. It is described as ‘empty’ (glossed as *E-*) because I have not honed in on a convincing morphemic analysis (consistent, non-contradictory gloss) for the formative, perhaps because there is none (i.e. it has the status of a morpheme). The prefix *e-* is discussed in more detail in §5.3 and in Tallman & Gallinat (Accepted).

Position 5 is filled out by incorporated nouns. The incorporated noun roots do not come with additional modifiers when they occur in this position. Position 5 corresponds roughly to the first of Pitman’s root position. The noun roots that can incorporate refer to elements that are typically conceptualized as bearing a part-whole relation to another participant expressed in the clause. For instance, the nouns (*e*)*sha* ‘leaf (tree)’ and *háha* ‘fruit (of a tree/plant)’ refer to parts of trees (see Vuillermet 2014 for an analysis of analogous phenomena in Ese Ejja). Noun incorporation involves no reduction of transitivity, and the noun bears a possessed relationship with a P argument (Emkow 2019: 117). Illustrative examples are provided in (7a) and (7b).

Table 4: The prefix slot

Position	Gloss	Morpheme
4	Empty	<i>e-</i>
	Negative	<i>pi-...-ma</i>
	Interactional	<i>ja-...-ti</i>
	Slowly	<i>tsi-</i>
	Posterior	<i>pa-</i>
	Interrogative	<i>ke-</i>
	In vain	<i>noma-</i>
	Apart	<i>shoma-</i>
	Still	<i>sho-</i>

- (7) a. *wakwala-ja wada anodi shoa ?iji -(i)ki (e-shoa 'head')*  
     3       -       -       5     6   14  
     woman-ERG 3:SG:GEN daughter hair tie -RECP:PST  
     'The mother tied the hair of her daughter'  
     Sp. 'La mama amarró el cabello de su hija.' (ELIC: ARA > SP)
- b. *yama akwi-limi pa- sha ?iji (e-sha 'branch')*  
     3       -       4     5     6  
     1SG:ERG tree-DIM POST- branch tie  
     'I am going to tie up a leaf / some leaves from a small tree.'  
     Sp. 'Yo voy a amarrar una hoja de un arbol pequeño.' (ELIC: ARA > SP)

Pitman provides a list of aspectual morphemes and does not make any claims about their relative ordering. The planar structure splits his position for aspectual suffixes into a zone for adverbial suffixes (position 7), a slot for the expressive suffixes *-shodi* and *-limi* (position 8), and a slot for the frustrative *-sawa* (position 9). Position 7 is a zone because morphemes of this position can variably order with each other without variably ordering with morphemes of adjacent positions. The most frequent morphemes of this position are presented in Table 5. The *Root* is also listed in this position, an issue I will discuss in more detail in §4.3 on non-permutability. The variable ordering is illustrated in (8) and (9). However, there are few examples in naturalistic speech where there is more than one adverbial suffix.

Table 5: The adverbial suffix zone

Position	Categories	Morphemes
	Motion	- <i>shawiya</i> ‘do and go’, - <i>shana</i> ‘going’, - <i>jajo</i> ‘arrive and do’ - <i>shao</i> ‘come and do’, - <i>yoa</i> ‘wandering’, - <i>wiki</i> ‘going P’
	Time of day	- <i>sisa</i> ‘at night’, - <i>wena</i> ‘in morning/dawn’
7		- <i>tseiye</i> ‘during day’, - <i>niapona</i> ‘at dusk’
	Aspectual	- <i>jaena</i> start’, - <i>weya</i> ‘finish’ - <i>sa</i> ‘in vain’, - <i>pe</i> ‘completely, all of P’
	Manner	- <i>pasi</i> short period of time’, - <i>titi</i> ‘slowly’

- (8) *loe -shawiya -sisa -ta /loe -sisa -shawiya -ta*  
 6 -7 -7 -12 /6 -7 -7 12  
 put -do&go -at.night -3A/3PL /put -at.night -do&go -3A/3PL  
 ‘After digging the hole all/at night, s/he went.’ /Sp. ‘Después de cavar  
 toda la noche, se fue.’ (ELIC: ARA > SP)

- (9) *tsaba -tseiye -wiki /tsaba -wiki -tseiye*  
 6 -7 -7 /6 -7 -7  
 hear -all.day -going:P /hear -going:P -all.day  
 ‘Always listening everyday to something going.’ /Sp. ‘Siempre escucha  
 todo el dia.’ (ELIC: ARA > SP)

After position 7, positions 8 and 9 are for expressives *-limi* ‘affective’ and *-shodi* ‘in pain’ and the frustrative/counterfactual morpheme *-sawa* respectively. These positions are necessary because expressives always occur after motion and time of day suffixes as in (10).<sup>5</sup>

- (10) *po dipa a -ta -iki -we po -sisa -shodi*  
 2 2 6 -12 -14 -15 6 -7 -8  
 so this.way do -3A/3PL -RECP:PST -LIMIT AUX:INTR -at.night -EMOT

<sup>5</sup>Pitman (1980) contains an apparent counterexample with *jodo banalimititia* ‘He went a short distance to see for a little bit’ / Sp. ‘fue poca distancia para ver un rato’, where *-limi* is an expressive and *-titi* is an adverbial suffix expressing ‘slowly’. I do not yet have enough examples of the suffix *-titi* in my corpus to clearly position this morpheme in the planar structure. I simply note that my consultants reject this sentence (I have not been able to re-elicit it successfully).

*po* -*tseiye*

6 -7

AUX:INTR -during.day

'And so in this way (because of this), the poor one spent the whole night and the whole day (up the tree with the jasibakwa).' / Sp. 'y asi ese rato no mas de noche el pobre y todo el dia tambien' (TXT 2698:0381)

That expressives must occur after the adverbial suffixes is illustrated in the examples in (11) and (12) below.

- (11) *kwe -shana -limi -ta* (\**kwe-limi-shana-ta*)  
 6 -7 -8 -12  
 cut -going -DIM -3A/3PL (\*cut-DIM-going:S/A-3A)  
 'Cutting small things (bushes) on the way.' / Sp. 'Cortando cosas pequeñas de ida.' (ELIC: ARA > SP, Chanito Matawa)
- (12) *piso -shana -shodi -ta* (\**piso-shodi-shana-ta*)  
 6 -7 -8 -12  
 untie -going -EMOT -3A/3PL (\*untie-AFF-going-3A/3PL)  
 'On the go s/he untied him/her as a favor.' / Sp. 'Le dió un favor desatandole.' (ELIC: ARA > SP)

Similarly, associated motion morphemes must occur before the frustrative -*sawa* 'with effort' as in (13).

- (13) *loe -shawiya -sawa -ta* / (\**loe-sawa-shawiya-ta*)  
 6 7 -9 -12  
 dig.up -do&go -FRUST -3A/3PL / (\*dig.up-with.difficulty-do&go-3A/3PL)  
 'S/he dug it out with difficulty before going.' / Sp. 'Le cavó cansandose antes de ir.' (ELIC: ARA > SP)

When -*sawa* 'with difficulty' occurs, it must occur after the expressives as illustrated in (14).

- (14) *di -shodi -sawa* / \*(*di-sawa-shodi*)  
 6 -8 -9  
 eat -pity -CNTRFCT / \*(eat-CNTRFCT-pity)  
 'Poor him had wanted to eat.'  
 Sp. 'El pobre tenía ganas de comer.' (ELIC: ARA > SP)

None of the time of day suffixes nor any of the expressives can occur after the suffixal piece of the interactional marker *ja-...-ti* (of position 10). All motion suffixes must appear before the interactional marker as well, except for the morpheme *-yoa*, which can appear in position 7 or 14.

The causative can also be variably ordered with the interactional marker *-ti*. In order to capture the possibility that the causative suffix appears after the interactional *ja-...-ti*, another position for the causative suffix must be added in position 11. The positional variability of *-me* is illustrated in (15).

(15)	<i>ja-</i>	<i>bailia -me -ti</i>	/	<i>ja-</i>	<i>bailia -ti -me</i>	
	4-	6      -7      -10		4-	6      -10	-11
INTRC- greet -CAUS -INTRC / INTRC- greet -INTRC -CAUS						
'Teach to greet someone.' / Sp. 'Enseñar a alguien a saludar.' (ELIC: ARA > SP)						

In contrast, the causative *-eme* cannot variably order with the third person {A} subject marker *-ta*.<sup>6</sup> Table 6 contains the suffixes that occur in position 14. I add position 15 because *-isha* and *-we* can appear after morphemes, filling out position 14.

Table 6: The posture/tense suffix slot

Position	Categories	Morphemes
14	Temporal distance	- <i>iki</i> 'recent past 1', - <i>ja</i> 'recent past 2' - <i>asha</i> 'distant past', - <i>isa</i> 'remote past'
	Posture/Aspect	- <i>ja</i> 'lying' - <i>ani</i> 'sitting' - <i>bade</i> 'hanging', - <i>neti</i> 'standing'
	Negation	- <i>ma</i> 'negative'
	Motion	- <i>yoa</i> 'wandering'
	Modal/Mood	- <i>toa</i> 'possibility', - <i>tame</i> 'counterfactual' - <i>ke</i> 'imperative'

## 4 Morphosyntactic domains

This section provides an overview of morphosyntactic domains in Araona. Section §4.1 deals with results and fractures of free occurrence. Section §4.3 is con-

<sup>6</sup>The positions 12, 13, and 14 are the same as those of Pitman's (1980) *Edge 2*, *Edge 3* and *Edge 1b* respectively, except that the interactional is not in position 12.

cerned with domains of non-permutability. Section §4.2 is concerned with domains of non-interruptability. Section §4.4 is concerned with domains of ciscategorial selection. Section §4.5 is concerned with domains related to deviations from biuniqueness. §4.6 is concerned with recursion based constituency tests.

#### 4.1 Free occurrence (6–6, 4–17)

The FREE OCCURRENCE VARIABLE must be fractured into two subtypes: the minimal and maximal domain. The MINIMAL FREE OCCURRENCE DOMAIN refers to the smallest subspan overlapping the verb that can function as an (elliptical) utterance. In Araona, a bare verb root can occur by itself without modification (Pitman & Pitman 1976). It is the only obligatory element (by definition) of a verbal predicate construction.

For instance, the following sentence was uttered by Marta Matawa after a pot of chicha had fallen in her kitchen while we were recording. The verb *olo* 'fall' appears with no morphosyntactic elaboration. The MINIMAL FREE OCCURRENCE domain is the 6–6 span.

- (16) *olo*  
     6  
     fall  
     ‘It fell.’  
     Sp. ‘Se cayó abajo’(TXT 1081:0082, Marta Matawa)

The MAXIMAL FREE OCCURRENCE DOMAIN identifies the 4–17 span. This identifies the largest span overlapping the verb core which is a single free form. This is illustrated in (17). The morpheme *nai* 'rain' can be omitted and the sentence is grammatical. As far as I have been able to discern the morpheme *tsio* 'when' is bound: it cannot occur as an elliptical utterance.<sup>7</sup>

- (17) (*nai*) *pi-*   *olo* -*ma*   *tsio*  
     (3)   4-      6   -14   17  
     (rain) NEG- fall -NEG when  
     ‘When it (rain) does not fall (one cannot harvest).’ / Sp. ‘Y cuando la lluvia  
     no se cae.’(TXT 1139:10)

Elements outside of the 4–17 are free or else cannot surface without another free element.

---

<sup>7</sup>While the form is often translated as ‘when’ (Sp. ‘cuando’) I have never heard it being used as an elliptical question ‘when’, but only in the context of a larger utterance even if it is separated by a pause.

## 4.2 Non-interruptability (6–6,4–17)

The NON-INTERRUPTION VARIABLE refers to a domain whose elements cannot be interrupted by element *I*. The variable is fractured according to which interrupting element *I* we choose to consider.

The first fracture specifies *I* as a combination of free forms. In Araona nouns and adjectives are both free forms. When they combine to form a noun phrase, they fit the criterion for *I* where *I* is a combination of free forms. The non-interruptability by free form combinations is the 4–16 subspan. While there are numerous free elements that can interrupt the 4–16 span, there are no combinations of free forms that can. Noun phrases cannot interrupt any part of this span. In an auxiliary verb construction, the auxiliary has to be adjacent to the verb complex in the sense that it cannot be interrupted by an NP. This is illustrated by the examples in (18) and (19).

- (18) *dea* *esi-po*            *oto*      *e-po*  
      3       -                    6        16  
man old-REL/NMLZ cough E-AUXINTR  
‘The old man coughed.’ / Sp. ‘El hombre viejo tosió.’      (ELIC: ARA > SP)
- (19) \**oto*    *dea* *esi-po*            *e-po*  
      6       -       -                    16  
cough man old-REL/NMLZ E-AUX:INTR  
Intended: ‘The old man coughed.’ / Sp. ‘El hombre viejo tosió.’      (ELIC:  
ARA > SP)

I have no obvious cases of NPs interrupting verbs and auxiliaries in my corpus. Furthermore, no examples of full NPs interrupting verb auxiliary combinations appear in any other published source as far as I am aware (Pitman 1980, Emkow 2019).<sup>8</sup>

The second version of the test would be NON-INTERRUPTION BY A SINGLE FREE FORM. This variable identifies a 6–6 span. Position 5 is fitted out by noun roots, which are free in Araona.

## 4.3 Non-permutability (6–6, 4–6)

Non-permutability is based on the often made claim that word or phrase constituents do not display variable ordering - their elements cannot permute. There

<sup>8</sup>Note that Emkow (2019: 88) considers the combination of the main verb with an auxiliary as a grammatical word in Araona. Unfortunately she does not explain why she thinks this, but non-interruptability could be rallied to support this claim.

are two interpretations present in the literature of this claim. One could be called “strict” – no variable ordering is allowed in any circumstances (Dixon & Aikhenvald 2002). Another could be called “flexible” – elements are in a fixed order or they can be variably ordered but with an obligatory scope difference (Anderson 2005).

In Araona, a further complication arises because of noun incorporation. The question is whether we should consider an incorporated noun root in position 5 to be an instance of a head noun from a noun phrase from positions 1, 3, 18 or a distinct morpheme. To appreciate the problem consider the following sentences, which both have the same meaning. At face value one might argue that the noun *wátsi* ‘foot’ can permute with prefixes of position 4.

- (20) *pi-*    *watsi-* *iji* -*ma* / *watsi pi-*    *iji* -*ma*  
       4-      5          6 -14 / 3      4-      6 -14  
       NEG- foot    tie -NEG / foot  NEG- tie -NEG  
       ‘One cannot tie its foot.’ / Sp. ‘No debe amarrar su pie.’

However it is not clear whether an incorporated noun of position 5 should be treated as the same noun which heads a full noun phrase occupying positions 1, 3 or 18, or whether it should be treated as a lexically and diachronically related element. Theoretical models of noun incorporation are likewise split on the issue. Syntactic approaches tend to assume identity (Baker 1988, Sadock 1991), whereas lexicalist approaches tend not to (Rosen 1989, Anderson 2005). In a lexicalist approach *wátsi* ‘foot’ might not be viewed as permuting with *pi-* ‘negation’, because the incorporated *wátsi* ‘foot’ is a different sort of element from the *wátsi* ‘foot’ of the non-incorporated example – the similarity between the forms being a fact about diachrony. A syntactic approach might assume that *wátsi* does ‘permute’ with *pi-* ‘negative’ – a noun root can fit out more than one structural position.

The problem is that this is not an all or none issue. In Araona, the reason to assume that incorporated and unincorporated *wátsi* ‘foot’ are the same elements in different positions is that incorporated and unincorporated forms are the same phonologically. As far I have been able to discern the range of senses of the incorporated and unincorporated nouns are also the same, thus providing apparent evidence for a syntactic approach. On the other hand, incorporated nouns cannot fit out position 5 with any accompanying modifiers. Incorporated nouns appear to not be referential compared to unincorporated counterparts (Tallman & Grossman 2022), and the set of incorporable nouns is a closed class. The methodology employed here reports results from both analyses. More technically, we “fracture” according to analysis according to the logic of “full reporting”.

On the assumption that incorporated nouns are distinct elements from non-incorporated variants, the strict non-permutability domain is 4–6. Otherwise the strict non-permutability domain is 6–6, i.e. there would be no fixedness in the verb complex.

#### 4.4 Ciscategorial selection (6–13, 4–15, 1–17)

The CISCATEGORIAL SELECTION VARIABLE refers to domains identified by spans of ciscategorial elements, elements which are specific to the part of speech class of a specific planar template. Ciscategorial selection captures the idea that affixes are more selective than other elements.

Noun roots are transcategorial. They can combine with other nouns in noun-noun combinations as in the forms in (21).

- (21) a. *zotó-wi* ‘jaguar nose’ (*zoto* ‘jaguar’; *éwi* ‘nose’)  
b. *zotó-tsoa* ‘jaguar bone’ (*zoto* ‘jaguar’; *etsoa* ‘bone’)  
c. *tsokwé-kwe* ‘toucan beak’ (*tsokwe* ‘toucan’; *ekwe* ‘beak’)  
d. *babá-tae* ‘Shaman house’ (*baba* ‘God’; *etae* ‘house’)  
e. *tseyé-na* ‘long river’ (*tseye* ‘day’; *ena* ‘agua’)  
f. *akwí-ça* ‘tree branch’ (*akwi* ‘tree’; *eça* ‘branch’)  
g. *çoá-íya* ‘hair (on head)’ (*eçoa* ‘head’; *eiya* ‘hair, leaf’)  
h. *toá-na* ‘tear’ (*etoa* ‘eye’; *ena* ‘water’)  
i. *mé-shokwe* ‘thumb’ (*eme* ‘hand’; *shokwe* ‘stump’)  
j. *ziki-tsoa* ‘sternum’ (*ziki* ‘chest’; *etsoa* ‘bone’) (reelicited from Pitman 1980: 223)

The morpheme *-odi* ‘always, only, repeatedly, just’ is transcategorial. It can combine with verbs as in (22a) and with nouns as in (22b) and (22c).

- (22) a. *araona dea kana ja da kabiña pona kana jemi -odi*  
1 - - - 3 - - - 6 -14  
Araona man PL ERG that cabiña woman PL grab/marry **-always**  
'The Araona men always married the Cabiña women.' / Sp. 'Los  
hombres araonas juntaban siempre con las mujeres cabiñas.' (TXT  
0700:0112)
- b. *kwama mimi metse -odi pewe*  
1PL:GEN speech with **only** completely  
'Only our language.' / Sp. 'Puro nuestro idioma no más.' (TXT  
1739:0007)

- c. *wada -odi elo -ta -neti*  
 3 - 6- 12 -14  
 3SG:ERG only catch-3A-stand  
 'He was the only one that was fishing.' / Sp. 'era la unica que está acabando pescado.' (TXT 0067:0009)

The ciscategorial variable in Araona has three interpretations. The MINIMAL CISCATEGORIAL DOMAIN identifies a subspan overlapping the verb core which contains positions which can only contain ciscategorial elements: the latter domain is the 6–14 subspan. All of the elements within this domain are verb ciscategorial. The noun root in position 5 and the element *-odi* 'always, only, repeatedly' of position 14 are plausibly transcategorial.

There is a narrow and broad way of defining a verb-ciscategorial element. On the narrow definition, a verb-ciscategorial element can only surface if the verb core is filled out. On the broad definition of ciscategorial an element is verb-ciscategorial if it cannot combine with any *other* lexical part of speech classes except the verb. The broad definition allows such ciscategorial elements to appear in nonverbal predicate construction. The maximal ciscategorial domain can be fractured according to the narrow versus broad interpretations of ciscategoriality.

The NARROW MAXIMAL CISCATEGORIAL SELECTION DOMAIN identifies the 4–15 subspan. All position 4 prefixes are verb ciscategorial. Position 15 contains at least one ciscategorial element *-lelajai* 'habitually', which is not attested in combination with nouns. While transcategorial elements can fill out position 15, the maximal domain is defined based on the presence of ciscategorial elements in positions rather than the total absence of transcategorial elements across its span.

The BROAD MAXIMAL CISCATEGORIAL SELECTION DOMAIN identifies the 1–17 span (the whole verbal planar structure). If elements are ciscategorial just because they do not combine with nouns or adjectives, then we would include position 1 adverbials and the connector *tsio* 'while'. Notice these elements can occur in utterances where a verb is not present, in nonverbal predicate constructions, as in examples (23a) and (23b).

- (23) a. *ajalili ado-eje tsio ...*  
 large macho when ...  
 'When I was fat and large.' / Sp. 'Cuando yo estaba gordo y macho.'  
 (TXT 1535:521)

- b. *kwipá tso pa eshai kana*  
when ANT REP spirit PL  
'How were the spirits?' / Sp. 'Como eran los duendes?' (TXT  
0049:0001)

#### 4.5 Extended exponence (4–10, 4–14)

For many morphologists deviations from biuniqueness signal morphological relations (Tallman & Epps 2020 for review and criticism). Across Takanan languages a number of verbal markers are realized with extended exponents or circumfixes (Guillaume forthcoming, see also Harris (2017) for typological overview and relevant terminology). I will assume that the positions fitted out by extended exponence correspond to the boundaries of various extended exponent domains in Araona (Tallman 2021b). This section discusses two categories which are expressed through extended exponents, specifically circumfixes: (i) the interaction marker *ha-...-ti* and (ii) the negative marker *pi-...-ma*. The posture suffixes could also be argued to be extended exponents in Araona, because they come with a prefix *e-* obligatorily.<sup>9</sup> However, a description of the distribution of the prefix *e-* requires consideration of phonological issues. For this and other reasons I will leave a discussion of the morphosyntax and/or phonology of posture suffixes to §5.3.<sup>10</sup>

The interactional marker displays extended exponence with a prefixal piece *ha-* and a suffixal piece *-ti*. Under certain circumstances the prefixal piece can be drop, but these circumstances are not yet well understood. The most common functions of the marker are as reciprocal as in (25) or as a middle or middle-like meaning as in (24). The prefixal part of the interactional marker *ha- ... -ti(me)* occurs in position 4. The suffixal piece *-ti* 'interactional' (24). One can see from this example that the suffixal piece comes after the adverbial suffix zone.

(24)	<i>wada beipa</i>	<i>ja-</i>	<i>ba</i>	<i>-weiya</i>	<i>-ti</i>	<i>-me</i>	<i>-tso</i>	...	<i>kwichai</i>
	1	3	4	6	-7	-10	-11	14	...

3SG not.know INTRC- feel -end -INTRC -CAUS -PRIOR ... spirit  
*jo-batae*

this-like

'She went unconscious because of the spirit / It made her go unconscious,

<sup>9</sup>Strictly speaking this is not true. The prefix only surfaces when the right-adjacent element (a noun or verb root) is vowel initial. When the element which is right-adjacent to the noun/verb, a L+H\* accent appears on the first syllable. This has been analyzed as a case where the prefix *e-* 'drops' after the L+H\* accent has been assigned (Pitman & Pitman 1976, Pitman 1980).

<sup>10</sup>This is not the same as saying that the distribution of *e-* is purely phonological.

it was a spirit like this one (that did it)' / Sp. 'Ella, no sé como, se puso inconsciente, y era espíritu así como eso.' (TXT 0049:0148)

One can see from Example (25) that the suffixal piece comes before the finality marker.

- (25) *ma ja- meya meya -ti -ibo pochi yama dia*  
 1 4- 6 6 -10 -13 - - - -  
 that INTRC leave leave -INTRC FINAL this 1.SG:ERG like.so  
*a-ja-ba-ja*  
 -  
 say-DUR-VIS-PST

'They (the young couples) are always leaving other, that's why I said it like this.' / Sp. 'Lo que están entre sí están dejándose (las parejas) yo dije.' (TXT 1447:0047)

The alternative order (*-ibo-ti*) is judged ungrammatical by speakers of the language. Morphemes of position 8/9 do not co-occur with *ja....ti(me)* in naturalistic speech, but the combination is easily elicitable. The following data show that the interactional marker occurs after the emotive *-shodi* and the frustrative *-sawa*. The reverse order is not permitted. This is illustrated in (26a)-(26d).

- (26) a. *ja- zamojo -shodi -ti (\*jazamojotishodi)*  
 4- 6 -8 -10  
 INTRC- hug -EMOT -INTRC  
 'They hug each other (with emotional pain).' / Sp. 'Ellos se abrazan de despedida.' (ELIC, ARA>SP)
- b. *ja- zamojo -sawa -ti (\*jazamojotisawa)*  
 4- 6- 9- 10  
 INTRC- hug -FRUST -INTRC  
 'They almost hug each other.' / Sp. 'Ellos casi se abrazan.' (ELIC, ARA > ESP)
- c. *ja- ba -weiya -shodi -ti (\*jabaweiyatishodi)*  
 4- 6 -7 -8 -10  
 INTRC- see/feel -stop -EMOT -INTRC  
 'S/he (the poor one) lost all feeling.' / Sp. 'Casi está por morirse el pobre.' (ELIC, ARA > ESP)

- d. *ja- ba- weiya- sawa- ti* (\**jabaweyatisawa*)  
 4- 6 -7 -9 -10  
 INTRC- see/feel -stop -FRUST -INTRC  
 ‘S/he almost lost all feeling.’ / Sp. ‘Estaba borracho, casi no sintió  
 nada.’ (ELIC, ARA > ESP)

The negative marker *pi-* ... -(m)a is also realized as a circumfix. The left edge is position 4 because this is the position occupied by the prefixal piece *pi-*. The right edge of the domain of negative exponence is more difficult to determine. The suffixal piece clearly appears before the interactional suffixal piece *-ti* as illustrated in example (27).<sup>11</sup>

- (27) *Baiwipi Nali-nae amigo batawe pojo mo pi- meya -ti -ma*  
 - - - - 2 - 4 6 -10 -14  
 Baipiwi Nali-with friend like so FOC NEG leave -INTRC -NEG  
 ‘Baipiwi and Nali were like friends, that it why they never left one  
 another.’ / Sp. ‘Baipiwi y Nali era amigos andaba juntos, por eso no se  
 larga.’ (TXT 0056:0149)

The suffixal piece of the negative marker also occurs after the finality marker *-ibo*, as illustrated in (28).

- (28) *dos mil veinte pewe pa e- izoa -ta -ni ba -asha dos*  
 1 - - 2 2 4- 6 -12 -14 6 -14 1  
 two thousand twenty no.more REP E- wait -3.A/3.PL -sit see -DIST:PST two  
*mil tsio pi- po -be -ibo -ma po -tso*  
 2 4- 6 -6 -13 -14 6 -14  
 thousand when NEG- go -come -FINAL -NEG AUX:INTR -PRIOR ...  
 ‘We will wait until 2020, he never came in 2000.’ / Sp. ‘Vamos a esperar a  
 2020, y no venía en 2000.’ (TXT 1549:0314)

I have posited that the suffixal piece of the negative marker is *-ma* and that it is in position 14. This is because it displays mutual exclusivity with all other morphemes from this position in the same verb complex, including posture suffixes and tense markers (see §3 for more details). The negative marker *-ma* does not co-occur with the marker *isha* ‘again’ in my corpus, of position 15. But it does

---

<sup>11</sup>One might think that the *-ti* in this case is the AM marker *-ti* ‘do and go’. Note that speakers consider *pi-meya-ti-ma* to be the negative version of the reciprocal construction *ja-meya-ti* INTRC-leave-INTRC ‘leave each other’ and ascribe it the same meaning that does not involve motion.

co-occur with the morpheme of *-we* ‘still’ of the same position as in (29). Thus we can infer that the negative marker is not in position 15.

- (29) *da-batae bamewe pi- a -ma -we*  
 1           6       4-   6 -14 -15  
 this-similar now   NEG- a -NEG -still

‘We will wait until 2020, he never came in 2000.’ / Sp. ‘Como por ejemplo,  
 ellos no sabian.’  
 (TXT 1737:0150)

Note that the negative marker can co-occur with morphemes of position 14 as long as these are not within the same verb complex. A verb can host one morpheme of position 14 and an auxiliary can host another. Thus (30a) is ungrammatical but (30b) is accepted.

- (30) a. *\*pi- dyi -ma -odi*  
 4     6 -14 -  
 NEG- eat -NEG -FREQ  
 ‘He never eats.’ / Sp. ‘Siempre no come.’
- b. *pi- dyi -ma a -odi*  
 4     6 -14 6 -14  
 NEG- eat -NEG AUX:TR -FREQ  
 ‘He never eats.’ / Sp. ‘Siempre no come.’

Thus, the domain of interactional exponence is 4–10 and the domain of negative exponence is 4–14.

#### 4.6 Subspan repetition

Subspan repetition refers to constructions or structures that can be analyzed as repeating spans of structure of the verbal planar structure. We could also call these “recursion-based diagnostics” as long as by recursion we mean self-similar iteration or self-similar embedding (the difference is not important for constituency in my view). A SUBSPAN REPETITION VARIABLE renders some recursion based diagnostic into a test result or domain with respect to the planar-fractal method.

There are four recursive structures in Araona that can be formulated into variables in this fashion; (i) auxiliary verb constructions; (ii) clause linkage via *tso* ‘prior’ or *tsio* ‘while’. Another could be formulated with the clause linker / relativizer *po*, however, I do not yet have enough data to determine the relevant facts in this case. We could also consider cases of nominalization with the suffix *-hi* as relevant, but I do not yet have enough data to assess such cases.

#### 4.6.1 Auxiliary verb construction (6–6, 4–15)

The auxiliary verb fills out position 16 in the verbal planar structure. It repeats certain positions of the verbal complex as a whole which overlap the verb core, which is why it is placed in a single position. As the structure of the auxiliary verb can be Auxiliary morphemes (*po* ‘do’ and *a* ‘do’) can also fill out the verb core position of the main verb. Pitman (1980: 71) describes two types of auxiliary verbs. The verb *a* ‘do, say, (transitive)’ and the verb *po* ‘do, be (at), say, go (intransitive).'<sup>12</sup>

The auxiliary can be used to accommodate more than one marker of the same slot in the verb complex. For instance, the prefixal piece of the negative *pi-* and the prefixal piece of the interactional *ha-* cannot co-occur in the verb complex as they occupy a single slot. They can appear together in the verb complex if an auxiliary is present, because one can fill out the prefixal slot of the auxiliary. The combination of negative and interactional marking is illustrated in (31a) and (31b) below. These sentences have the same meaning.

- (31) a. *pi- ti -a -ma ja- ?- a -ti*  
 4- 6 - -14 4- - 6 -10  
 NEG- give -E -NEG INTRC- E- AUX:TR -INTRC  
 ‘to not share/ give together’  
 Sp. ‘no dar juntos, no compartir’ (ELIC: ARA > ESP)
- b. *ja- ti -a -ti pi- a -ma*  
 4- 6 - -10 4- 6 -14  
 INTRC- gave -E -INTRC NEG- AUX:TR -NEG  
 ‘to not share/ give together’  
 Sp. ‘no dar juntos, no compartir’ (ELIC: ARA > ESP)

In an auxiliary verb construction the morpheme *-ta* ‘third person A, third person plural’ can appear with the main verb as in (32a) or the auxiliary verb as in (32b).

- (32) a. *ema nio-wa tsoi -ta pi- a -ma*  
 1 3 6 -12 4 6 -14  
 1SG:ABS dog-ERG bite -3A NEG AUX:TR -NEG  
 ‘The dog didn’t bite me.’  
 Sp. ‘El perro no me mordió.’ (ELIC, ARA > ESP)

<sup>12</sup> According to (Pitman 1980), his *Edge 2*, *Edge 3* and *Edge 1b* suffixes can occur on auxiliaries. These correspond to position 4, position 10, position 12, and position 14.

- b. *ema nio-wa pi- tsoi -ma a -ta*  
 1 3 4- 6 -14 6 -12  
 1SG:ABS dog-ERG NEG- bite -NEG AUX:TR -3A  
 'The dog didn't bite me.'  
 Sp. 'El perro no me mordió.' (ELIC, ARA > ESP)

According to Pitman "few other of the verbal suffixes attach to the auxiliaries" (Pitman 1980: 71), apart from those morphemes of positions 4, 10, 12 and 14. Pitman does not give a specific example of which verbal suffixes cannot appear on the auxiliary. Time of day suffixes (position 7) *can* appear on the auxiliary. Position 15 markers can also appear on the auxiliary. These facts are illustrated from (33) and (34) respectively, both sentences from Pitman's sketch. I re-elicited these forms to corroborate Pitman's data.

- (33) *wada ema tawi -zowi a -sisa -ta*  
 1 3 6 -7 6 -7 -12  
 3SG:ERG 1SG:ABS dream -interrupt AUX:TR -at.night -3A  
 'He interfered with my sleep at night.' / Sp. 'él interfió con mi sueño durante la noche.' (Pitman 1980: 41)
- (34) *zia shoma- a -me a-lelajai*  
 3 4- 6 -14 6 14  
 maiz apart- do -CAUS AUX:TR -HAB  
 'By custom we prepare the maiz seperately' / Sp. 'Por costumbre preparamos el maíz aparte.' (Pitman 1980: 31)

I have not been able to corroborate Pitman's claim regarding restrictions on the appearance of verbal suffixes on auxiliaries. Pitman provides no more details except to state that some restrictions exist. Auxiliaries cannot take NP arguments distinct from the main verb, nor any of the Wackernagel-like morphemes of position 2.

The MAXIMAL AUXILIARY SUBSPAN REPETITION DOMAIN corresponds to the largest span that can be filled out by the auxiliary. This domain identifies a 4–15 span with the auxiliary construction.

The MINIMAL AUXILIARY SUBSPAN REPETITION DOMAIN is the subspan of structure containing positions whose elements cannot display wide scope over the main verb and the auxiliary in the auxiliary verb construction. There are no elements that satisfy this condition in Araona, and, thus, the domain identifies a 6–6 span. There are no morphemes that *do not* scope over the auxiliary and the verb

when they occur on either one of them. In fact, there are no attested cases of a morpheme that can appear on both the main verb and the auxiliary at the same time, and switching a morpheme position from the main verb to the auxiliary or *vice versa* makes no difference to the meaning of an expression. Illustrative examples are provided in (35a) and (35b) below with the interrogative prefix *ke-*. I have not been able to find any cases where varying the position of a suffix with respect to whether its in the main versus auxiliary verb conditions a difference in meaning.

- (35) a. *midya ke- oto e- po*  
 3 4- 6 4- 6  
 2SG INTER- cough E- AUX:INTR  
 ‘Had you been coughing?’ / Sp. ‘¿Estabas tosiendo antes?’
- b. *midya e- oto ke- po*  
 3 4- 6 4- 6  
 2SG E- cough INTER- AUX:INTR  
 ‘Had you been coughing?’ / Sp. ‘¿Estabas tosiendo antes?’

#### 4.6.2 -*tso*-marked clause combination (4–14,1–17)

The marker *tso* appears in position 14 of a verb complex expressing an event immediately prior to the event expressed by the following main clause. These clauses can share the same subject as in (36a) and (36b).

- (36) a. *jae lale -tso jelo -ta -iki*  
 3 6 - 6 -12 -14  
 fish roast -PRIOR:COMPL eat -3A -RECP:PST  
 ‘After roasting the fish, he ate all of it.’  
 Sp. ‘Después de asar el pescado, lo comió todo.’ (Pitman 1980: 102)
- b. *awada piye -ti -wiki -tso tsa -tseiye -sa*  
 3 6 -9 -7 - 6 -7 -8  
 tapir shoot -go&do -going:P -PRIOR:COMPL look.for -all.day -FRUST  
 -ja  
 -14  
 -RECP:PST  
 ‘After shooting the tapir, he looked all day for it in vain.’  
 Sp. ‘Después de balear el anta, lo busqué casi todo el día en vano.’  
 (Pitman 1980: 52)

The MAXIMAL TSO-MARKED SUBSPAN REPETITION DOMAIN is the 1–17 span. This domain refers to the size of the *tso*-clause as judged by the positions that can be filled out in this clause independently from the main clause. I will give a brief justification of the maximal domain, before discussing the minimal one.

Pitman claims that clauses marked off with *-tso* ‘prior, anterior’ *must* have the same subjects (Pitman 1980) as the following main clause, as in (36a) and (36b). However there appear to be exceptions to this as in the sentences in (37a) and (37b) from naturalistic speech. It is also not hard to elicit a sentence where the A argument of the *tso* marked clause is coreferential with the P of the main clause as in (38).

- (37) a. *didia -tso aleokata ... po -ana -odi kwizi-sawa-po daesha*  
       6     -           18     .... 6   -6   -15   18           -  
       eat   -PRIOR:A/S quickly ... do -leave -FREQ fart-FRUST-NMLZ like.so  
       ‘Every time he eats quickly farts come out, that’s how it is.’ / Sp. ‘Cada vez que come rápido no más salen los pedos así es.’ (TXT 0098:0098)
- b. *palma.sola me- jemi -isha -tso-dada zai -ki*  
       3       4-    6   -15   -       6       -7  
       Palma.sola hand- grab -again -PRIOR:COMPL-only be.angry -come  
       -(i)bo ba-(a)sha           naeda  
       13      16       18  
       -PFV AUX:VIS-DIST:PST 3PL  
       ‘When we took Palma Sola and lower again, then they (the carayana)  
       got angry.’ / Sp. ‘Cuando agarramos palma sola y más abajo de nuevo,  
       de allí ellos se enojaron.’ (TXT 1739:0077)
- (38) *yama zoto pisa -wiki -tso shipa wada ema tsoi*  
       1       3    6    7   -       1    3       -    6  
       1SG:ERG jaguar shoot -going:P -PRIOR:COMPL more 3SG:ERG 1SG:ABS bite  
       -ta  
       -12  
       -3A  
       ‘When I shot the jaguar, it bit me afterwards.’ / Sp. ‘Cuando yo chumbié al  
       tigre, él después me mordió.’ (ELIC: ESP > ARA)

Example (38) shows that both clauses have a position 3 available to them and that each can have their own core arguments (they “project” these positions independent of the main clause). The *tso*-marked clause and the main clause can each project position 1 and the Wackernagel-like position 2. This is illustrated in (39).

(39)	<i>jae (pa) tso lake -pe -tso</i>	... <i>bisha tokwe pa kwawea di</i>
1	(2) 2 6 -11 -	... 1 2 2 3 6
	fish REP ANT roast -COMPL -PRIOR:COMPL	... again EPIS REP yuca eat
	<i>-ta</i>	
	<i>-12</i>	
	<i>-3A</i>	

'First, he roasted all of the fish and then he ate yuca.' / Sp. 'Primero asó todo el pescado, y después comió yuca.' (ELIC: ESP > ARA)

Both verb complexes can have position 4 and 14 filled out as in the example in (40), where both clauses are marked with negation.

(40)	<i>dea-ja kwawea pi- kwawi -ma a -tso pi- jelo</i>
1	3 4- 6 -14 -16 -
	man-ERG yuca NEG- roast -NEG AUX:TR -PRIOR:COMPL NEG- devour
	<i>-ma</i>
	<i>-14</i>
	<i>-NEG</i>

'The man didn't roast the yuca and didn't eat it.' / Sp. 'El hombre ni asó la yuca, y ni comió (\*el hombre ni asó, ni comió la yuca).' (ESP > ARA)

The minimal domain refers to the span of structure whose elements cannot display wide-scope in over both verbal complexes. NPs can always scope over both clauses as in (36a) and (36b). The markers of position 2 can scope over both of the clauses which is illustrated in (39). The sentence means the same regardless of whether *pa* 'reportative' is removed or not in the *tso*-marked clause. Position 4 and position 14 elements cannot scope over both clauses as illustrated in (41).

(41)	<i>dea-ja kwawea kwawi -tso pi- jelo -ma</i>
1	3 6 -
	man-ERG yuca roast -PRIOR:COMPL NEG- devour -NEG
	'The man roasted the yuca, but didn't eat it (*the man neither roasted nor ate the yuca).' / Sp. 'El hombre asó, pero no comió la yuca.'
	*'El hombre ni asó, ni comió la yuca').' (ESP > ARA)

Elements between the negative circumfix cannot scope over the verb conjoined complexes. For instance, *-tseiye* 'all day' cannot scope over both clauses in (36b). The time of day marker only scopes locally. With a few marginal exceptions, elements from 4–14 cannot scope over two clauses combined with *tso*. The MINIMAL TSO-MARKED SUBSPAN REPETITION variable thus identifies the 4–14 span.

Because I have found no convincing evidence that elements from position 4 through 14 can scope over two clauses marked by *tso* or *tsio*, the minimal domain is 4–14. The maximal domain is 1–17. This includes all positions except the position for post-verbal NPs, position 18. As stated in §3, arguments cannot occur post-verbally in dependent clauses, i.e. those marked by *tso* or *tsio*. The maximal domain of subspan repetition is therefore 1–17.

## 5 Phonological domains

This section is concerned with domains that can be identified based on the (non)-application of phonological processes. Section §5.1 is concerned with pitch-accent assignment domains. §5.2 is concerned with domains of vowel coalescence. Leading off from the discussion of morphosyntax, §5.3 is concerned with domains that can be identified based on the distribution of the dummy prefix *e-*. In these sections I will also include an extra line for phonetic transcription in order to be able to more easily describe the results of phonological processes.

### 5.1 Pitch-accent domains (6–11, 4–17)

Araona has (at least) four tones; LH\*, H%, L% and (L)HL%. None of these tones are contrastive. They can all be regarded as “post-lexical” (Ladd 2008). The notation and terminology here follows Pierrehumbert & Beckman (1986) and Gussenhoven & Bruce (1999). The “%” refers to an intonational tone docked to the edge of an utterance, the “\*\*” refers to the fact that a tone docks to a stressed syllable. The L+H\* pitch-accent is docked to the second syllable of a span of structure whose left edge is 4. Usually, the L is realized on the first syllable and the H on the second. I will refer to this rule as the *pitch-accent rule*. It is illustrated in example (42), the pitch track for the sentence is presented in Figure 1. All elements in the noun phrase, except a few suffixes receive an L+H\* on their second syllable. An illustrative example is presented in (43), for which a pitch track is provided in Figure 2. In each of these examples the intonational marker L% is docked to the final syllable of the utterance phrase.<sup>13</sup>

---

<sup>13</sup>Pitman & Pitman (1976) and Pitman (1980) describe a default “stress” rule that places stress on the second syllable in Araona, agreeing with my L+H\* superficially. However, we do not agree on the exceptions. While it might be possible to partially reconstruct the missionaries’ analysis into something meaningful, their comments on how one should go about identifying stress and the fact that they left no accessible recordings makes it impossible for future researchers to corroborate or test any of their claims. They explicitly state that pitch is not necessarily involved, but their comments about the phonetic interpretation of “stress” are too vague to engage with scientifically (Pitman & Pitman 1976: 10).

- (42) [sipó pisáiki ↓]  
*sipo pisa -iki*  
 3 6 -14  
 squirrel kill -RECP:PST  
 'S/he killed a squirrel.' / Sp. 'Mató una ardilla recién.' (ELIC)

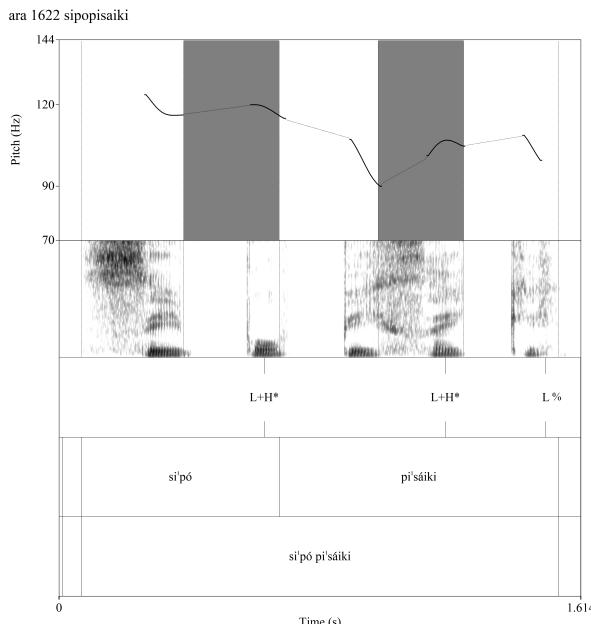


Figure 1: Pitch track for the sentence *sipo pisaiki*

- (43) [jamá <sup>n</sup>datí kwa<sup>n</sup>déwaha ↓]  
*yama dati kwadewa -ja*  
 1 3 6 -14  
 1SG:ERG tortoise chase -RECP:PST  
 'I chased the tortoise.' / Sp. 'Yo correteé a la peta.' (ELIC)

The prefix is counted as part of the domain of pitch accent assignment. If a prefix occurs (position 4), the pitch accent will occur on the first syllable of the verb root (assuming there is no incorporated noun). This fact is illustrated with the example in (44) with its corresponding pitch track in 3.

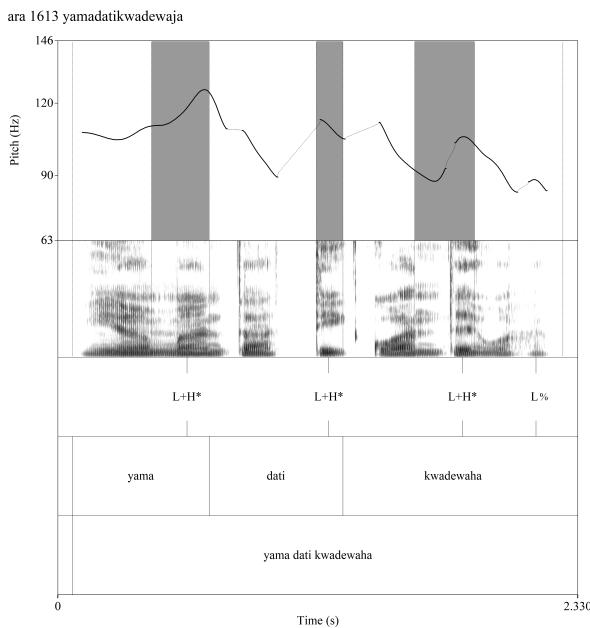


Figure 2: Pitch track of the sentence *yama dati kwadewaha*

- (44) [wàlípa pìbábama tsio ↑↓]  
*walipa pi- baba -ma tsio*  
 3 4 6 14 17  
 chicken NEG- know -NEG still/when  
 ‘When we still didn’t know about chickens’  
 Sp. ‘Cuando todavía no conociamos el gallo.’ (TXT 1535:0306)

Thus the left edge of the LH\* DOMAIN is position 4. The right edge of the domain corresponds to a position prior to the next possible LH\* pitch, where the pitch assignment rule restarts. Note that the L+H\* pitch accent *can* occur on position 14 morphemes. For instance, the H\* will dock to the first syllable of *-lelajai* ‘habitual’ in *po-lelajai* producing /polé.la.hai/. This shows that elements of 14 are in the domain. It is somewhat less clear whether elements of position 15 should be included in the LH\* DOMAIN. The morpheme *-we* ‘still’ is only in this position because it can collocate with a negative marker *-ma* as in (29). The morpheme *ishá* ‘again’ occurs with a L+H\* docked to the second syllable. If the rest of the verb complex is minimally bisyllabic, then the verb complex will contain two L+H\* tones; *ishá* ‘again’ will be realized with two L+H\* pitch independent of that from the host. This can be seen from comparing the examples such as *iji-isha*

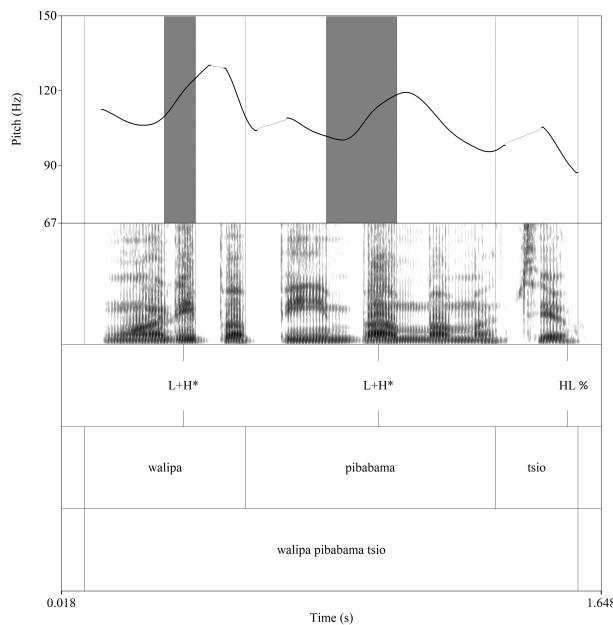


Figure 3: Pitch track of the sentence *walipa pibabama tsio*

‘tie-again’ and *po-isha* ‘do-again’, tokens of these examples appear in Figures 4 and 5.

- (45) [ihí içá ↓]  
*iji ishá*  
 6 15  
 tie again  
 ‘S/he tied again.’  
 Sp. ‘Amarró otra vez.’ (ELIC)

- (46) po içá ↓  
*po isha*  
 6 15  
 do again  
 ‘S/he did it/so again.’  
 Sp. ‘Hizo otra vez.’ (ELIC)

The MAXIMAL LH\* DOMAIN thus contains position 15.

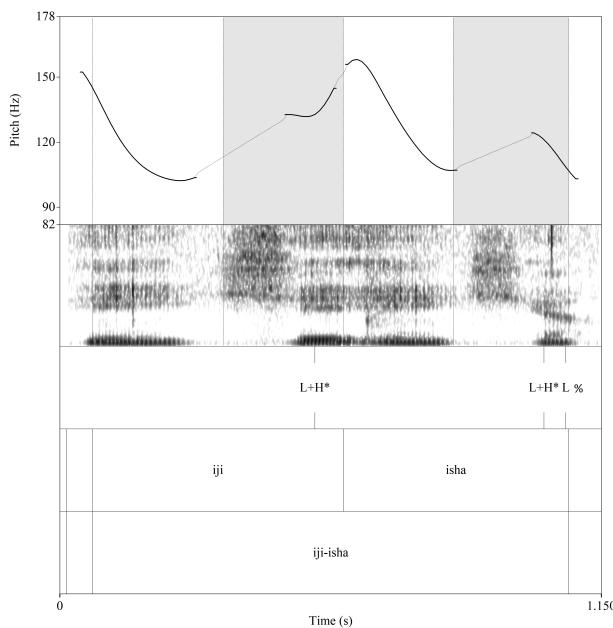


Figure 4: Pitch track for the sentence *iji-ishā*

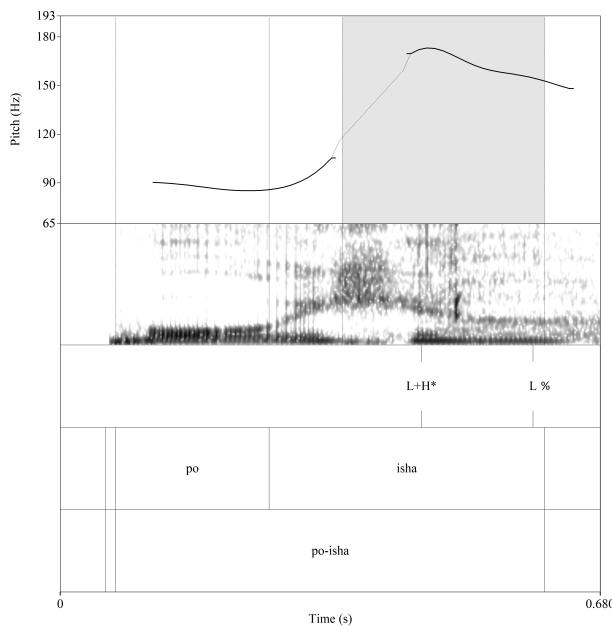


Figure 5: Pitch track for the sentence *po-ishā*

The marker *tsio* of position 17 also does not seem to receive its own pitch accent as in (44). After this position we have full noun phrases, and thus, the L+H\* pitch rule necessarily restarts. The MAXIMAL LH\* DOMAIN is thus 4–17.

The minimal domain corresponds to the span of structure overlapping the verb core where one could never find another L+H\*. The left edge of this domain would be 6. The reason is that position 5 can be occupied by a noun which takes its own L+H\*, independent of the verb. This is illustrated in Example (47) and the corresponding pitch track provided in Figure 6.

- (47) pi watsí ihí ma ↓  
*pi*    *watsi iji* *ma*  
4      5      6    -14  
NEG- foot tie -NEG  
'S/he didn't tie the foot.' / Sp. 'No amarró la pie.'                         (ELIC)

Thus position 5 cannot be in the MINIMAL LH DOMAIN. The right edge of the minimal domain can be determined by finding the position to the right of the verb core, closest to the verb core, where an element can be found that has an L+H\* independent of the L+H\* that occurs due to the presence of the verb (i.e. that occurs because of the LH domain projected from the verb). This is position 12, which contains the marker *-ta* which can take its own L+H\* independent of that associated with the verb core. An example can be found in (48) illustrated with the pitch track in Figure 7. We can see that the verb form has two L+H\* pitch accents.<sup>14</sup>

- (48) [çáma tá ibo yoá ↓]  
(e-) *shama -ta            -ibo        -yoá*  
4    6        -12        -13        -14  
E    see        -3A/3PL -FINAL -wandering  
'S/he went visiting them.'  
Sp. 'Ellos se fueron visitando.'     (ELIC)

The reader will have noticed in the examples above that the height reached by the H component of the pitch accents is different when there is more than one L+H\* present in the same example. In each case the first pitch accent reaches a higher peak than the second. One might argue that the examples in (47) and (48) contain one *main* pitch accent and that another domain should be posited to

<sup>14</sup>Note that in the following example, a pitch accent occurs on the first syllable. This is because of a rule in Araona that deletes the prefix *e-*. The issue is discussed in §5.3.

13 *Graded constituency in the Araona (Takana) verb complex*

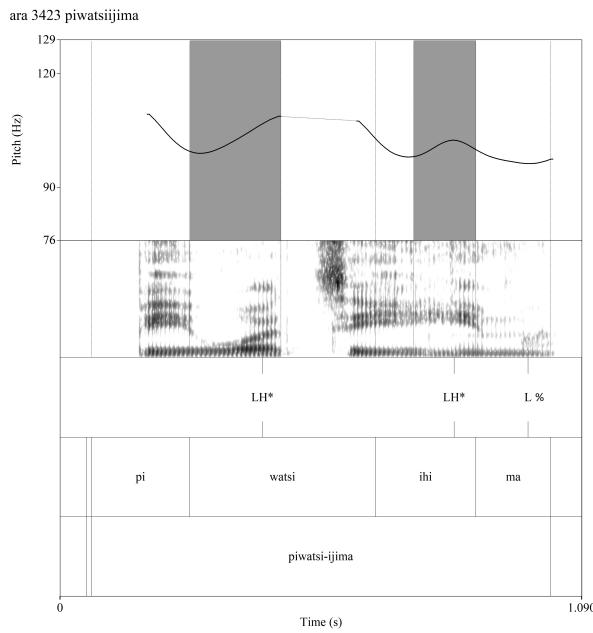


Figure 6: Pitch track for the sentence *pi-watsi-iji-ma*

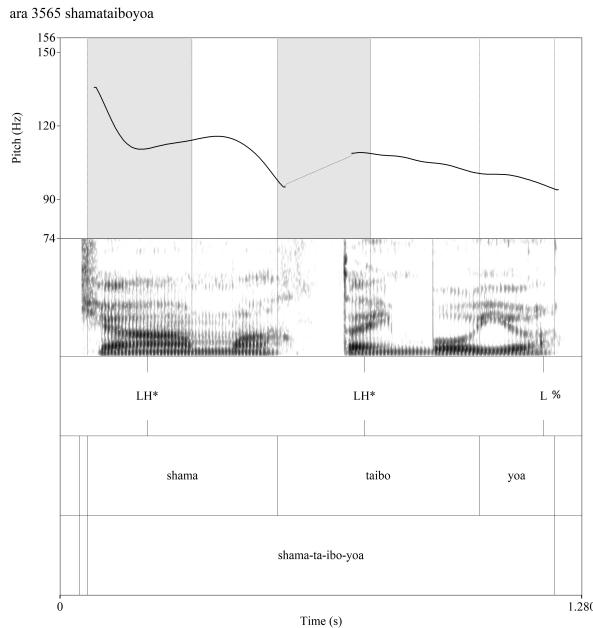


Figure 7: Pitch track for the sentence *e-shama-ta-ibo-yoa*

account for the position of the second. Or similarly, perhaps the first pitch accent corresponds to “primary stress” and the second one to “secondary stress” and that the domain should be formulated only in terms of primary stress (Emkow 2006). These are reasonable criticisms which I do not have space to fully test at this point. The reason I consider these cases to have multiple L+H\* tones of the same type is that the difference in pitch height can be understood as resulting from a declination throughout the utterance, which from the utterances I have observed produces the same effect throughout the whole sentence. I consider this to be the most reasonable hypothesis at this point given the ubiquity of declination cross-linguistically (Ladd 2008), but future research may require us to posit different types of pitch accent with different relative heights (in terms of F0), perhaps requiring an accentual domain larger than the LH DOMAIN.

This section has also presented a simplified view of L+H\* accent assignment in Araona. The assignment rule also interacts with utterance level intonational markers such as L% (Tallman & Gallate Accepted). This issue was not discussed because it is not important for stating the domain of application of the LH\* rule.

## 5.2 Vowel syncope/synaeresis domains (6–6, 6–14)

When two vowels occur adjacent to one another at a position juncture there are three possibilities in Araona; (i) deletion: one of the vowels deletes; (ii) synaeresis: a diphthong or phonetic long vowel is created; (iii) an insertion of a glottal stop between the vowels. All three of these processes apply depending on the juncture.

Table 7 summarizes the vowel deletions or diphthongizations that occur in Araona. Only five morphemes are involved in such deletion or diphthongization operations; -eme (position ‘causative’; -ibo ‘perfective’; -iki ‘recent past’; -asha ‘distant past’; -ani ‘sit’.

Table 7: Vowel combinations (deletions and combinations)

_i (-ibo, -iba -iki)		_e (-eme)	_a (-asha)	_a (ani)
i_	i	i	ia	ia~a
e_	ei	e	ea	ea~a
o_	oi	oe	oa	oa~o
a_	ai	ae	a	aa~a

The main generalization that emerges from the combinations is that when two identical vowel phonemes occur at a juncture, one deletes. Additionally /ie/

is disallowed. The morpheme *-ibo~iba* ‘finally’ occurs in position 13. The deletion rule can be seen as operative when these are adjacent to verb roots of position 6, as in the example *nawi-ibo* swim-FINAL → [nawí<sup>m</sup>bo]. Vowel deletion can be seen as operative where *shodi* ‘emotive’ (position 8) is left-adjacent to *-ibo~iba* as in *di-shodi-ibo* ‘eat-EMOT-FINAL → [<sup>n</sup>diçó<sup>n</sup>di<sup>m</sup>bo]. Vowel deletion can be seen as operative when *-ti* of position 10 is left-adjacent to *ibo~iba* as in *ja-zamojo-ti-ibo* ‘INTRC-hug-INTRC-FINAL’ → [hazámohoti<sup>m</sup>bo].

Similarly the morpheme *-iki* ‘recent past’ participates in this vowel deletion rule when it combines with any of the same morphemes described above. The morpheme *-eme* ‘causative’ which I have placed in position 7 and 11 can be seen as realized as *-me* whenever it is right-adjacent to any morpheme which ends in /e/. For instance, *kwe-eme* ‘cut-CAUS’ is realized as [kwéme]. The /e/ is also lost if the morpheme is right-adjacent to /i/. For instance, *ja-ba-ti-eme* ‘INTRC-see-INTRC-CAUS’ is realized as [ha<sup>m</sup>bátíme]. Finally, both *-ani* ‘sitting, future’ and *-asha* ‘distant past’ of position 14 lose their first phoneme when they are right-adjacent to a morpheme with /a/. In the case of *-asha* ‘distant past’ the reduction is obligatorily, as in *ba-asha* ‘see-DIST:PST’ → [<sup>m</sup>báça].

I will assume that there is a general vowel deletion rule operative across these cases that is responsible for the allomorphy that we find with the suffixes:  $V_i \# V_i \rightarrow V_i$  where # is a juncture between positions 6–14. The rule should read as follows: if two vowels adjacent to one another are of the same quality, delete one of them.

Outside of the 6–14 span, glottal stops are inserted between vowels flanking a juncture between positions.<sup>15</sup> For instance, the *i* of *isha* ‘again’ is not subjected to the vowel deletion process ever, as illustrated in (49) below.

- (49) [i.hí.?*i*.ça ↓]  
*iji -isha*  
 6 -15  
 tie -again  
 ‘Tie again.’ / Sp. ‘Amarrar otra vez.’ (ARA > SP)

The left boundary of this domain is 6. When an element from positions of the span 1–5 occurs with a final vowel adjacent to the first vowel of the verb core (position 6) a glottal stop is inserted as illustrated in (50a) and (50b) below.

<sup>15</sup>For most speakers of Araona ‘glottal stops’ are realized as creaky voice rather than a complete closure in the vocal tract (Gordon & Ladefoged 2001, Garellek 2013), but the phonetic realization of glottal stops across Araona speakers requires more research.

The vowels have to be of the same quality otherwise the rule applies differently according to the prefix.<sup>16</sup>

- (50) a. [wa.tsí.?i.hí ↓]  
*watsi iji*  
5 6  
foot tie  
'Foot-tie.' / Sp. 'Amarrar pie.' (ARA > SP)
- b. [pi.?í.hi.ma]  
*pi- iji -ma*  
4- 6 -14  
NEG- tie -NEG  
'S/he does not tie it.' / Sp. 'El/ella no lo amarra.' (ARA > SP)

The MINIMAL  $V_i \# V_i \rightarrow V_i / V_j \# V_i \rightarrow V_j V_i$  DOMAIN refers to the span overlapping the verb core that contains only positive evidence for the vowel deletion rule.<sup>17</sup> This identifies the 6–6 span. We have no evidence for the application of vowel deletion in position 7 either way.

The MAXIMAL  $i \# V_i \rightarrow V_i / V_j \# V_i \rightarrow V_j V_i$  DOMAIN refers to the span overlapping the verb core where we have no negative evidence against its application. For the maximal domain we assume that the process is applying “vacuously” across junctures where its phonological preconditions are never met. This domain identifies a 6–14 span, because outside this structure we can find junctures flanking vowels, which introduce a glottal stop to block the adjacent vowels.

### 5.3 E-selection / initial L+H\* domain (4–15)

This section identifies the domains for the phonological and/or morphosyntactic rules that account for the distribution of the prefix *e*- . The analogous and cognate morph of other Takanan languages is described as an inflectional prefix, whose distribution is determined by the presence or absence of other inflectional morphemes (Vuillermet 2012, Guillaume 2008, forthcoming). In this section I describe the prefix *e*- in terms of insertion rules that make reference to the presence or

<sup>16</sup>The prefixes seem vary in terms of how the rule of glottal insertion operates. For instance, the posterior/future *pa-* and the intensifier *tsi-* will always come with a glottal stop if it is right-adjacent to a vowel. The interactional marker *ja-* and the negative *pi-* will only insert a glottal stop if the vowel has the same quality. The prefix *e*- never occurs with a glottal stop to its right, but this can be attributed to the fact that there are no verb roots that begin with /e/ in Araona.

<sup>17</sup>Note that the disjunctive rule that I have stated has one exception. A combination /ie/ is realized as /i/ as in *nawi-me* 'bathe-CAUS'.

absence of certain suffixes and phonological context. The two styles of analysis are not incompatible in general: Araona *e-* could be described as an inflectional prefix whose distribution is partially phonologically determined.<sup>18</sup>

When the prefix *e-* appears on verbs in Araona it has been described as coding a specific meaning (Emkow 2006, 2019, Pitman & Pitman 1976, Pitman 1980). The two authors who have written on the topic (myself excluded) do not have consistent glosses of the morpheme. Emkow glosses the morph ‘declarative’ in Emkow (2006: 114, 123) and Emkow (2019: 356, 365), as ‘directional’ in Emkow (2006: 106) and Emkow (2019: 272), and ‘resultative’ in Emkow (2019: 209) and as a second person singular absolute marker in Emkow (2019: 318). Unfortunately, Emkow does not provide any evidence for these glosses (and see Tallman & Gallinat Accepted for specific counterexamples to each of them).

Pitman & Pitman (1976) and Pitman (1980) provide the gloss ‘affirmative’ (Sp. ‘*afirmativo*’). They never define this term. Pitman (1980) claims that *e-* also marks ‘narrative past’, a notion which is never defined nor defended (there are many verbs in the past in narratives that do not occur with the prefix, so it is unclear what the empirical force of the claim is, cf. Tallman 2021a). Pitman & Pitman (1976: 16) state “the full significance of presence and absence of this prefix in relation to the discourse structure of Araona has not yet to be determined”. The statement might be misleading because it implies that some clarification of the ‘significance’ of the morph had been given, where none had been and never has in any of the missionaries’ sources to my knowledge.

I have not found any specific meaning for the prefix *e-*. It is possible that future research will uncover a meaning for it, but no one has presented any convincing evidence thus far. I do not understand what the purpose is (unless it is obfuscation of one’s current state of knowledge) in proposing a gloss for a morph, which cannot be verified – I therefore, gloss the prefix *e-* as E-. What is clear, however, is that there are specific morphosyntactic and phonological environments where *e-* must occur, cannot occur and can occur optionally. These morphosyntactic and phonological contexts can be translated into spans of structure. It is less clear whether these span results should be regarded as phonological or morphosyntactic domains.

---

<sup>18</sup>I do not understand what the benefit of a using the concept of “inflection” is in any Takanan language, which is why I avoid the term. I have two reasons for this: (i) the notion of inflection is applied to an arbitrary set of morphemes in the languages that do not share any jointly sufficient and necessary properties; (ii) the one criterion that is brought up for identifying inflection is “obligatoriness” (Guillaume 2008: 179–181), but the definition of the word “obligatory” has to change in order to fit the Takanan data, as so-called “obligatory” slots need not be filled in naturalistic speech.

So much for the semantics/pragmatics of *e*-. The (morpho)phonology of *e*- provides a problem which I regard as expositional that needs to be addressed. In Araona there are a number of contexts where the pitch accent rule described in 5.1 is violated and the L+H\* appears on the first syllable. In the same contexts for L+H\* appears on the first syllable, the prefix *e*- would appear if it were not for the verb root being consonant initial. A proposal to make irregularity in the stress rules disappear seems motivated from the coincidence. One can assume that there is an underlying *e*- wherever L+H\* occurs on the first syllable on the surface, it is actually the result of docking to the second syllable on an ‘underlying form’ followed by the subsequent deletion of the prefix (the first syllable) *e*. The default L+H\* rule applies before a *e*- deletion rule. On this analysis, Araona’s pitch accent rule is perfectly regular (Pitman & Pitman 1976). Pitman & Pitman (1976) further justify the rule on the grounds that the *e*- in such cases can be found in cognate forms in other Takanan languages (e.g. *pona* ‘woman’ is *e-púna* in Maropa, Tacana, and Cavineña and *e-póna* in Ese Ejja).

The way the analysis works for verbs is illustrated below in example (51). First the ‘underlying form’ *e-tawi-ani* receives pitch accent assignment on the second syllable; /e-táwi-ani/. Then the prefix deletes because it is before a consonant: /táwi-ani/. The LH\* accent rule is thus not violated. The L+H\* of the first syllable can be realized as a relatively higher pitch or as a rising pitch as in Figure 7. The ordering of the rules is represented as proceeding from bottom to top starting with the underlying morphemic analysis.

- (51) [táwi ani ↓]  
Delete *e*- before C: *e*- táwi -ani  
LH\* assignment: *e*- táwi -ani  
*e- tawi -ani*  
E- sleep -sitting  
'S/he is sleeping' / Sp. 'Está durmiendo.'

Independent evidence for the underlying *e*- in such cases comes from the fact that when we replace *tawi* ‘sleep’ with a vowel initial verb, the prefix *e*- is realized (not deleted) as in the example in (52) below.

- (52) [eoló ani ↓]  
Delete *e*- before C: *e*- oló -ani  
LH\* assignment: *e*- oló -ani  
*e- olo -ani*  
E- fall -sitting  
'S/he/it is falling (in a sitting position)' / Sp. 'Está cayendo de sentado.'

One could posit an underlying *e-* with rule ordering or one could simply state that posture suffixes require a *e-* prefix or a L+H\* on the first syllable in case the verb is consonant initial.<sup>19</sup> For the purposes of the description presented here, it does not matter which one of these alternatives is chosen. The important point is that there are domains that condition the appearance or suppression of the *e-* prefix and that the presence of the prefix can also be cued by an initial LH\* pitch accent in certain phonological environments.

### 5.3.1 E-#L+H conditioning suffixes

The presence or absence of the prefix *e-* is conditioned by which suffixes occur after the verb complex up to position 17. They can divided into three types with respect to how they interact with *e-*.

- (53) a. *E-selecting* suffixes: They require the presence of the prefix *e-*, if the verb is vowel initial, otherwise they shift the pitch accent to the first syllable in the 4–17.
- b. *E-suppressing* suffixes: They ban the presence of the prefix *e-*.
- c. *E-neutral* suffixes: They are neutral with respect to the prefix *e-*. The prefix could appear or not.

An example of an *e-selecting* suffix is the posture suffix *-ani* ‘sitting’, illustrated in (52). Removing the prefix in this example is deemed ungrammatical by speakers; *olo-ani* ‘fall-sit’ is rejected. Another example from naturalistic speech is provided in (54) with the verb *otso* ‘burn, blossom’ and the *e-selecting* posture suffix *-neti*. *otsoneti* is not grammatical.

- (54) wéi<sup>m</sup>ba po mo <sup>m</sup>boisi eotsóneti <sup>m</sup>bá<sup>n</sup>di ↓  
we -iba po mo boisi e- otso -neti badi  
bloom -FINAL REL FOC mapajo E- blossom -standing moon/month  
'The month where the mapajo bloomed mapajo leaves blossom (June).' /  
Sp. 'En la mes cuando retoña las hojas de mapajo.' (TXT 1535:0008)

If the verb is consonant initial there will be a L+H\* tone on the first syllable if there is a *e-selecting* suffix in the verb complex. This is illustrated with *didiani*

---

<sup>19</sup>The rule is not quite this simple because *e-* can surface before a consonant if the verb root is monosyllabic with a certain set of suffixes. Posture suffixes on the other hand always disallow *e-* before a consonant initial verb. This is explained below.

as in (55), which contains the *e*-selecting suffix *-ani*. That the L+H\* pitch accent has shifted to the first syllable can be observed from Figure 8.<sup>20</sup>

- (55) [tsekwá esía díidianí po<sup>m</sup>bíshahaha ↓]  
*tsekwa esi-a di-di-ani pobishajaja*  
 vagina old-ERG E-eat-eat-sitting already  
 ‘The God of harvest (lit. old vagina) is already eating the offerings.’ /  
 Sp.‘El Dios (la vagina vieja) ya está comiendo.’ (TXT 0603:0228)

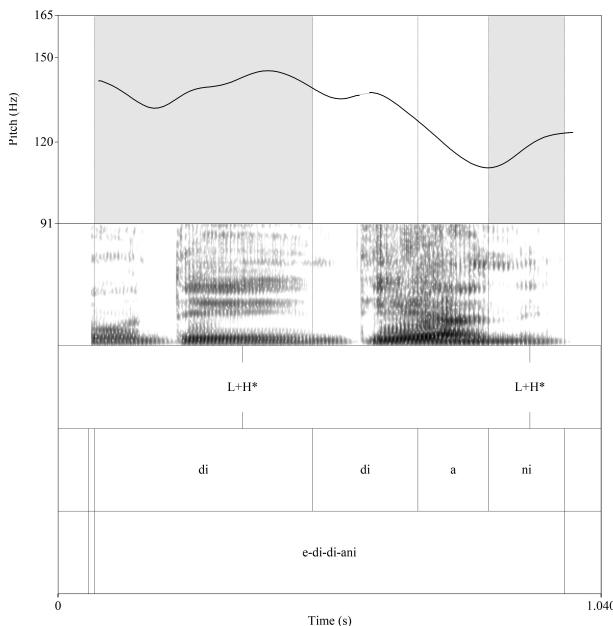


Figure 8: Pitch track for the sentence *e-di-di-ani*

*E*-banning suffixes do not allow the position 4 *e*- to surface. The L+H\* pitch accent always will occur on the second syllable in the presence of *e*-suppressing suffixes, following the default rule. An example of an *e*-suppressing suffix is *-tso* ‘prior, anterior’ of position 17.

<sup>20</sup>Note that this verb form also has a pitch accent on the final syllable as well. Final TAME markers and posture suffixes receive their own pitch accent sentence internally - in isolation their final pitch accent is blocked from appearing because of the intonation level tones. I do not yet know whether the L+H\* pitch accents always occur with such forms.

- (56) [<sup>m</sup>ba<sup>m</sup>bátaezo anítso pa ↓... <sup>m</sup>ba<sup>m</sup>ba kwéleani pa ↓]  
*babá-tae-zo*      *aní -tso*    *pa* ... *baba*    (*e*)- *kwéle*  
 god/spirit-house-SPAT sit -PRIOR REP ... god/spirit (E)- perform.ritual  
*-ani*    *pa*  
*-sitting*  
 'Inside the God house, he was doing a ritual.' / Sp.'Adentro del templo sagrado, estaba sentado. Estaba haciendo ritual.' (TXT 1549:0456-0457)

If the verb root is consonant initial the L+H\* accent will always fall on the second syllable as in (57). Figure 9 contains the form *mimi-tso* extracted from the example in (57), showing the L+H\* realized on the second syllable.

- (57) [pónae má <sup>n</sup>do báti mimítso ... anátiaça ↓]  
*po-nae*   *yama*   *do*   *ba-nati*    (*e*)-*mimi-tso*  
 that-with 1SG:ERG that see-go&do (E)-speak-PRIOR  
*a-nati-asha*  
 do/say-go&do-DIST:PST  
 'I went with them to see and converse, I went and said (something).' /  
 Sp.'Con ellos yo fui a visitar y conversé.' (TXT 0049:0227)

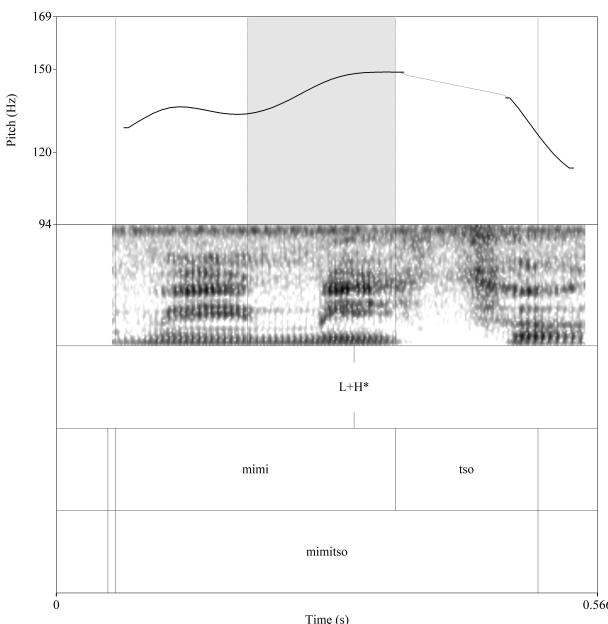


Figure 9: Pitch track for the verb form *mimi-tso*

Finally, there are *e*-neutral suffixes. In the context of such suffixes the appearance of *e*- is optional. For instance the suffix *-ibo* ‘finally’ is neutral with respect to whether the verb root receives a prefix *e*- or not, as illustrated with the examples in (58a) and (58b).

- (58) a. [ihíbo ↓]  
*iji-ibo*  
tie-FINAL  
'He finally tied it (at last he tied it).' / Sp. 'Por fin, lo amarró.'  
b. [eihíbo ↓]  
*e-iji-ibo*  
E-tie-FINAL  
'He finally tied it (at last he tied it).' / Sp. 'Por fin, lo amarró.'

For verb roots that are not monosyllabic *e*- is also always optional. Thus, the verb root *iboeta* ‘forget’ can be realized as *eiboeta* or *iboeta* when there are no other suffixes. The same is true of the initial L+H\* shift. The verb *hododo* ‘run’ can be realized as *hó<sup>n</sup>do<sup>n</sup>do* or *ho<sup>n</sup>dó<sup>n</sup>do*.<sup>21</sup>

In Araona, *e*-selecting, *e*-suppressing and *e*-neutral suffixes can appear in the same verb complex. When an *e*-neutral suffix co-occurs with an *e*-suppressing or an *e*-selecting suffix, the *e*-neutral suffix is ‘overruled’ by the selectional requirements of the others. For instance, if the *e*-suppressing morpheme *-tso* appears with the *e*-neutral morpheme *-ibo*, *e*- is suppressed, rather than being optional. If the *e*-selecting morpheme *-ani* co-occurs with the *e*-neutral morpheme *-ibo*, the prefix *e*- must occur. The basic pattern is illustrated in (59a) and (59b).

- (59) a. [eihíboani ↓] (\*ihíboani)  
*e-iji-ibo*  
E-tie-FINAL  
'He is finally tying it (at last he is tying it while seated).' / Sp. 'Por fin, lo está amarrando.'  
b. [ihibotso ↓] (\*eihíboani)  
*iji-ibo-tso*  
E-tie-FINAL-ANT  
'When he finally tied it.' / Sp. 'Cuando por fin amarró.'

That the *e*-suppressing and *e*-selecting morphemes take precedence regardless of the relative syntagmatic order of the suffixes can be seen from the following

---

<sup>21</sup>Pitman & Pitman (1976) also noted that that ‘stress’ could vary on bare verbs.

example. The morpheme *-ta* is *e*-suppressing and occurs before the morpheme *-ibo*. If these morphemes co-occur, then the *e*- (or initial L+H\*) is suppressed following the priorities of the morpheme *-ta*. For example, the initial L+H\* falls on the second syllable in the example in (60).

- (60) [pa tikwáijatai<sup>m</sup>bo édia ↓]  
*pa ti-kwaiya-ta-ibo e-di-a*  
 REP give-arrive-3A/3PL-FINAL E-eat-E  
 ‘When hee arrived he finally gave it (the heart of the peccary) to him (his brother) and he ate it (and his brother transformed into a peccary).’ / Sp.  
 ‘Cuando llegó lo dió, y lo comió.’ (TXT 2698:0565)

The *e*-suppressing suffixes are as follows: (i) *ta* ‘third person singular A, or third person plural S/A; (ii) *-iki* ‘recent past’; (*h*)*a* ‘recent past II’; *aça* ‘distant past’; *-tso* ‘prior, anterior’; *-ke* ‘imperative’; *-lelahai* ‘habitual’; *?odi* ‘frequentive’. All of these morphemes occur in position 14, except *-ta*, which occurs in position 12.

When an *e*-suppressing and an *e*-selecting suffix occur together, the *e*-selecting suffix wins out. An example of this is illustrated in (61). While *-ta* is *e*-suppressing, as can be seen from *ti-kwaiya-ta-ibo* in (60), the *e*-selecting morpheme *-ani* ‘sitting, progressive’ is also present. The result is that *e*- surfaces as in *e-olo-eme-ta-ibo-ani* ‘drop something down once and for all’, overriding the *e*-suppressing properties of *-ta*, which is an *e*-suppressing morpheme.

- (61) [<sup>m</sup>bo<sup>m</sup>bá pá eloemétai<sup>m</sup>boani jidyo ló Ángele kwi amohídžakweshodi do  
 aníme <sup>n</sup>dipa kana tsio ema pátsoa apétai<sup>m</sup>ba ↓]  
*boba pa e-olo-eme-ta-ibo-ani jidyo lo Angele kwi*  
 bomb REP E-drop-CAUS-3A/3PL-FINAL-standing here Los Angeles so  
*a-moiji-dya-kwe-shodi po dipa do ani-me kana*  
 ADJ-dangerous-INTENS-AUG-EMOT REL in.this.way that sit-there 3PL  
*tsio ema paichoa a-pe-ta-iba*  
 when 1SG carai-ERG say-COMPL-3A/3PL-FINAL:PST  
 ‘They are going to throw bombs here on Los Angeles it is said, so it is dangerous, very dangerous (it gives me pain), while others just stay there (stay in one spot, they are not worried), a carai (non-indigenous bolivian) told me everything.’ / Sp. ‘Va a largar bombas aquí a los Angeles, como es muy peligroso pero otros están tranquilos un carai me contó todo.’ TXT 1882:0060

Thus the rule for predicting *e*- or initial L+H\* shift follows a hierarchy; *e*-selecting > *e*-suppressing > *e*-neutral. This should be read as the suffix to the left of the hierarchy wins out in terms of the appearance of absence of *e*.<sup>22</sup>

If we assume that we can identify a domain of *e*-selection based on which elements participate in predicting the absence or presence of *e*, the domain would cover a 4–14 span. The whole process restarts at the auxiliary – a suffix on an auxiliary does not have any determining role in predicting whether *e*- appears on the main verb. The prefix is included because the presence of a prefix always blocks *e*- from occurring.

### 5.3.2 “Surface” *e*-deletion domain (4–6)

The deletion of *e*- resulting in a L+H\* on the surface is determined by the phonological structure of elements in positions 5 and 6. For *e*- to delete it must appear before a consonant and the span from 5–6 must contain at least two consonants. The prefix *e*- will always delete on the surface before verb roots such *tsaba* ‘hear’, *zewi* ‘write’, *lokwakwa* ‘boil’ or *hododo* ‘run’, because they begin with a consonant and contain at least two consonants. If there is a noun root which is consonant-initial, the *e*- will always delete as well; e.g. *e-nala-seo* ‘cut throat’ is realized as [nálaseo]. The *e*- deletion rule identifies a domain from 4–6.

### 5.3.3 E-minimality domain (4–15)

If there is only a single consonant in the positions 5–6, which can only occur if position 5 for incorporated nouns is empty, then *e*- *must* surface. It does not matter in this case where we say that *e*- is present underlyingly, or whether we claim that *e*- only appears on the surface. The point is that there is a constraint on the suppression of *e*- that is not predicted by the presence of *e*-suppressing suffixes, but **only** predicted by the phonological content of right-adjacent material in 5 and 6. The examples in (62a–62f) illustrate the constraint against monosyllabic forms.

- (62) a. ékwe ↓(\*kwe)  
      *e-kwe*  
      (E-)cut  
      ‘S/he cut it.’

<sup>22</sup>We could also state that when *e*-selecting suffixes and *e*-suppressing suffixes are in competition, the *e*-selecting or *e*-suppressing suffix furthest to the right wins out. It just so happens to be the case that there are no cases, as far as I know, where an *e*-suppressing suffix can follow an *e*-selecting suffix. The only *e*-selecting suffixes are posture verb suffixes, which are either in the same slot as or occur after all *e*-suppressing suffixes.

- b. ékwae ↓(\*kwaе)  
*e-kwae*  
 (E-)explain  
 'S/he explained.'
- c. étsoi ↓  
*e-tsoi*  
 (E-)bite  
 'S/he bit it.'
- d. épа ↓  
*e-pa*  
 (E-)cry  
 'S/he cried.'
- e. épо ↓  
*e-po*  
 (E-)do  
 'S/he did it/went.'
- f. ewi ↓  
*e-wi*  
 (E-)urinate  
 'S/he urinated.'

Note that as soon as a suffix is added to such forms, the *e-* is either banned from appearing or can delete. But when suppressing the *e-* would result in a monosyllabic root, *e-* must appear. All speakers reject *pa* as a free form for 'cry', for example. There are two exceptions to this rule: the verbs *ti* 'give' and *di* 'eat' can occur as isolated forms.

In Cavineña, the suffix *-u* is inserted on monosyllabic verbs with no other morphology. Guillaume (2008: 41) attributes the insertion rule to a minimality condition that forces phonological words to be bisyllabic. In Araona, the verb root *do* 'carry, manage, drive' similarly receives the epenthetic 'suffix' *-ho* to avoid being a monosyllabic form. Or stated alternatively – there are two forms of *do*, one which requires additional morphological material to surface and another one (*doho*) which can surface without it. The form *do* is ungrammatical without extra suffixes from positions 4 to 15. Instead *doho* must be used. No other morphemes display this specific (Cavineña-like) rule of *-ho* insertion. However, note the distribution of *do* vs. *doho* is identical to that of the alternation of *kwe* vs. *ekwe*. One might plausibly claim that Araona uses *e-* to avoid subminimal verbs just as

Cavineña uses *-u*. One difference is that there are two exceptions in Araona (*ti* ‘give’ and *di* ‘eat’) and *do* ‘carry’, which has a long form so that the *e-do* is not necessary (and, in fact, is considered ungrammatical). Minimality is rarely an all or nothing process, however, so this does not count as an argument against minimality conditioning the presence of *e*- (Garrett 1999: 69-70).

Another hypothesis might attribute the obligatory *e*- in the cases above to the idea that *e*- is an inflectional element. Since inflection is obligatory, *e*- is inserted for the purposes of making a complete ‘word’ (Guillaume 2008, forthcoming). The problem with this view is that it does not explain why *e*- does not obligatorily surface on verbs such as *iji* ‘tie’, *olo* ‘drop’ and *iboeta* ‘forget’. Speakers accept these forms in isolation and they appear in isolation in naturalistic speech. But if *e*- appears as a matter of making a “complete” word, it should surface on these forms. Even so, claiming that *e*- is an inflectional prefix is not mutually inconsistent with the idea that the element also bears a phonological function. A similar case appears in Bantu languages. Morphs which have no clear meaning but are co-opted for phonological purposes are described in Bantu languages as “stabilizers” (Gowlett 2007), and these morphs often seem to fill out positions considered to be ‘inflectional’. One can claim that *e*- is an inflectional element, but to capture its distribution we would still need to say that it is an inflectional prefix that’s distribution is partially conditioned by minimality. There is no contradiction in this claim unless one believes that language subsystems are necessarily self-contained and only interact via highly constrained interfaces without the possibility of bleeding into one another.

There are more details worth mentioning in relation to *e*- on monosyllabic forms. If the verb root or the incorporated noun and verb root contain more than one consonant then *e*- must delete. If the verb root is monosyllabic and any *e*-neutral elements from positions 7 through 15 are present, then *e*- deletion is optional (as it is for vowel initial forms).

- (63) a. [pá.ça.na ↓] ~[epá.ça.na ↓] ~[pa.çá.na ↓]  
(*e*-) *pa* -*shana*  
(4-) 6 -7  
(E-) cry -going  
'I/you cry while going.' (ARA > ESP, 2021-09-03)
- b. [kwé.si.sa ↓] ~[kwé.si.sa ↓] ~[kwe.sí.sa ↓]  
(*e*-) *kwe* -*sisa*  
(4-) 6 -7  
(E-) cut -at.night  
'I/you cut it at night.' (ARA > ESP, 2021-09-03)

The minimality domain extends to elements of position 15 on the right side as illustrated in the example in (64a). Note that this form does not obligatorily suppress *e-* as illustrated in (64b).

- (64) a. ipáke a?íça açéwe ↓  
*iya-ke a-isha ashewe*  
 grab-IMP do-FREQ still  
 ‘Grab it and keep doing so (holding the rope).’ / Sp. ‘Agarra pues!  
 todavía no lo largue!’ (TXT 1442:0012)
- b. pomoke kwitſi "da tsakata ea?íça <sup>m</sup>baçilio wana?íça ↓  
*pomoke kwichi da tsa-kata e-a-isha ba-shili-o*  
 This.way so this hard/difficult-AUG E-do-FREQ VIS-DEPREC-LIMIT  
*wana-isha*  
 go-FREQ  
 ‘First they did it (delimiting the territory) in the harsh way, after that  
 they did it again, and he did it again dammit, and went again (to the  
 government in order to advocate for a territory).’ / Sp. ‘Primero hizo  
 grave una cosa, así hizo pues.’ (TXT 1739:0216)

The minimality domain is the domain where if no affixes are present in this domain *e-* must insert on monosyllabic forms or if a *e*-neutral affix is present it does not need to block the insertion of *e-* monosyllabic forms. The domain is 4–15.

## 6 Summary and discussion

This section provides a brief summary and discussion of the results of applying the constituency/wordhood tests over the available Araona data. I will start with the phonological domains and then move to the morphosyntactic ones.

There are at least two main difficulties in assessing the issue of phonological vs. morphosyntactic wordhood in Araona. The first is that approximately a third of the domains that we have identified are actually “indeterminate” with regards to whether they should be considered phonological or morphosyntactic domains. There are a few reasons to classify a domain as indeterminate with respect to the morphosyntax-phonology division. The identification of the domain could involve the combination of phonological and morphosyntactic considerations. All domains that fall under the category of deviations from biuniqueness are accordingly indeterminate. Thus, extended exponence in Araona, which identifies three domains (Sections 4.5 and 5.3) are indeterminate. It could also be that

the available linguistic literature provides competing phonological and syntactic accounts of the phenomena which is responsible for defining the domain. This is true of all minimal subspan repetition domains, since these could be accounted for as conditions on ellipsis, which may be thought of as an operation that makes reference to phonological structure (Szczegielniak 2018). Free occurrence is also classified as indeterminate as authors vary in terms of whether they consider it a morphosyntactic test (Haspelmath 2011) or a phonological one (Zingler 2020).

When we are assessing convergence in phonological versus morphosyntactic domains, it matters whether we classify such domains as phonological or morphosyntactic.

The second problem in assessing how domain convergence relates to phonological and morphosyntactic wordhood in Araona is that there are competing converging domains in the language. If wordhood is marked off by high convergences, then its not clear which of these converging domains to choose as the word.

The purely phonological domains, displayed in Figure 10, show no span convergences, but do show convergences around the left and right edges. If the phonological domains are combined with indeterminate domains, then there is perhaps a domain of convergence on the 4–14 span. This depends on us assuming that the extended exponence and minimal subspan repetition are phonological diagnostics. A convergence strip plot which combines phonological and indeterminate domains is provided in Figure 11.

The pitch accent domain does not converge with any other domains. A methodological point is in order here. In the literature on Araona, syllabification rules are described as operating within the phonological word – one could infer from the discussions that syllabification is supposed to align with “stress” assignment. However, I have left syllabification out. The reason is that there is no known empirical consequence of syllabification except that some account of how adjacent vowels combine to make docking points for L+H\* is necessary to account for where L+H\* lands. For instance, in the form *e-ilo-wiki* ‘E-send-going:P’ one must state that /ei/ forms a syllable in order to capture the fact that L+H\* lands on the second syllable. However, syllabification is not domain independent of L+H\* assignment because it has no independent empirical effect from what I can observe. We are interested in identifying logically independent domains and syllabification in Araona does not have this status. Rather it is an analytic or ex-positional tool that helps describe the rule of L+H\* assignment more succinctly. Future research might find that there are independent effects of syllabification in Araona, which could change the picture presented here. I suspect that phonetic

13 *Graded constituency in the Araona (Takana) verb complex*

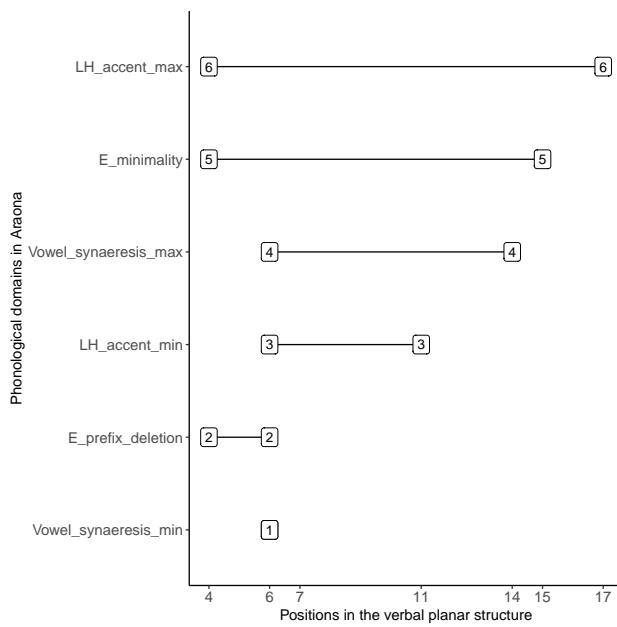


Figure 10: Phonological domains in Araona

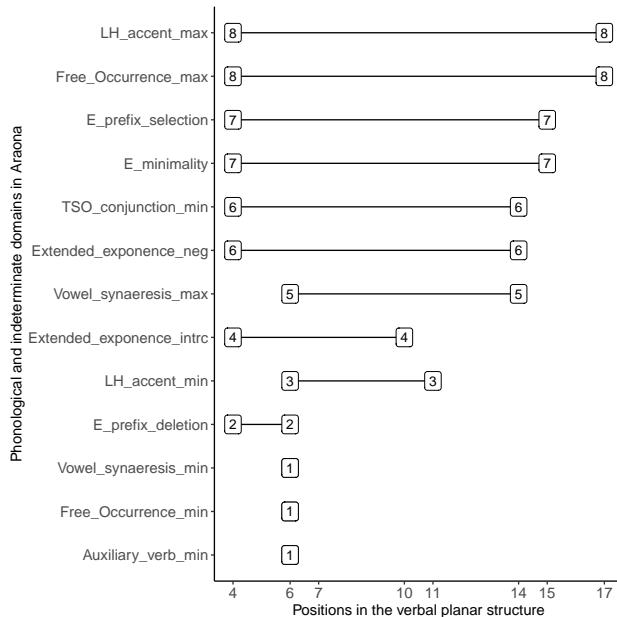


Figure 11: Phonological and indeterminate domains in Araona

studies will be necessary to determine whether syllabification domains can really be regarded as independent from L+H\* assignment (Fougeron 1999).

We also observed in §5.1 that the L+H\* tones hit different heights. Not all instances of L+H\* have the same phonetic realization. I attributed this to an automatic declination process in the language, not to some rule of downstep operating over a higher prosodic domain. Since I view automatic downstep as a product of pausing, and I am currently skeptical that pausing could be or should be coded in the planar-fractal method, these processes have not resulted in another domain. Future research might reveal that there is something like a phonological phrase in the language that results in different pitch heights associated with the peaks of L+H pitch accents. We might also figure out a way of coding pausing in the planar-fractal method.

If one considers “pure” morphosyntactic domains (i.e. those that would normally not be considered phonological), then there are two convergences in Araona as illustrated by the strip plot in Figure 12. Most constituency tests identified in the literature are vague with respect to which level they identify, but this is likely not the case with maximal subspan repetition domains. In the maximal subspan repetition domain for clause combination with *tso* and the maximal ciscategorial selection domain simply identify full clauses (the 1–17 span contains full NPs). The only other convergence is at position 6 in Araona. We would be forced to conclude based on these results that Araona is basically an isolating language, if “indeterminate” tests were not rallied.

The strip plot in Figure 13 displays morphosyntactic and indeterminate domains together. In this case the 4–14 span comes out as a possible wordhood candidate, but the smaller 6–6 span comes out somewhat stronger.

A pooling of the results is displayed in Figure 14. The overall picture is that there are three important layers/constituents below the sentence; (i) a small “stem” constituent that consists of just the verb core; (ii) a larger constituent which corresponds to the word in previous work on Araona that spans from “inflectional” prefixes to “inflectional” tense, aspect and posture morphemes; (iii) a post-word like constituent which contains the auxiliary and some clause-linkage morphemes.

When we ignore span convergence, as perhaps we should, and simply look at edges we find that the strongest structural edges in Araona are 4 on the left edge and 6 on the right edge.

Still the situation is much more complex and the facts of constituency appear to be much richer than what is typically described for Takanan languages, even if the results are vaguely in agreement with current descriptions. Guided by “Basic

13 *Graded constituency in the Araona (Takana) verb complex*

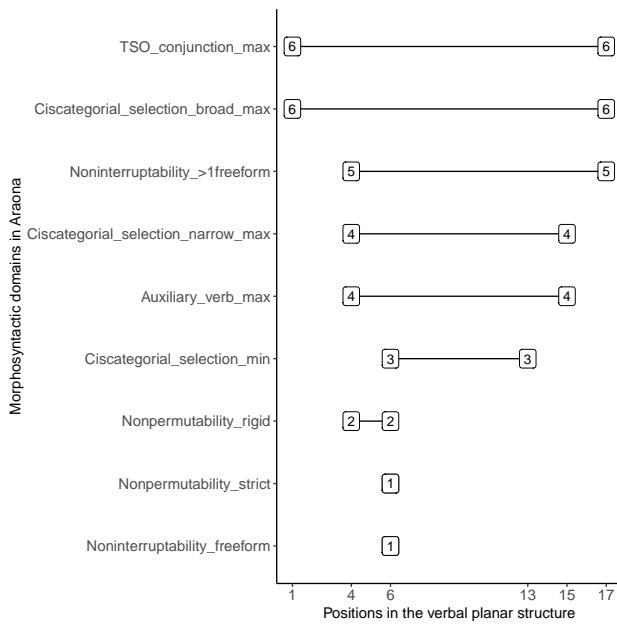


Figure 12: Pure morphosyntactic domains in Araona

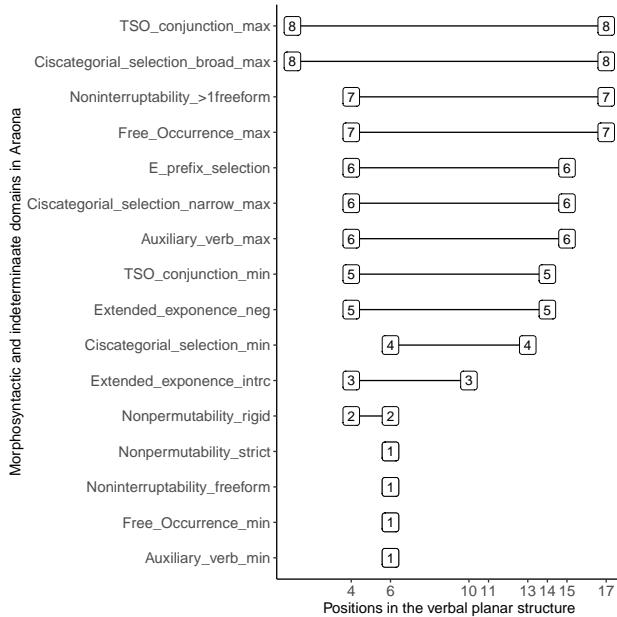


Figure 13: Morphosyntactic and indeterminate domains in Araona

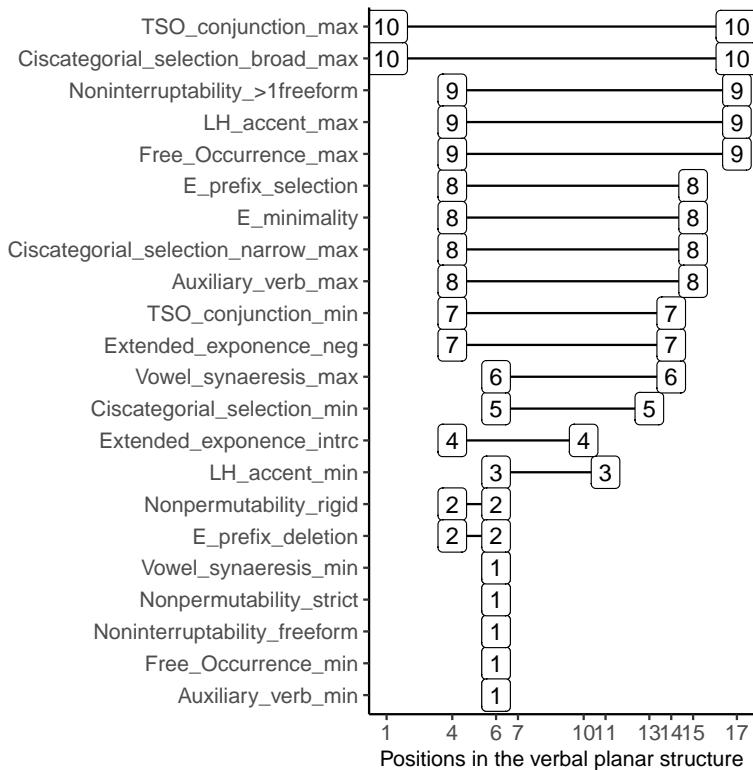


Figure 14: Domains in Araona (morphosyntax, phonology and indeterminacy pooled)

Linguistic Theory” (BLT), which assumes the universality (and the universal comparability) of phonological and morphosyntactic words, modern descriptions of Takanan languages tend to assume without argumentation that phonological domains align around a discrete and abstract “phonological word” (Guillaume 2008, Vuillermet 2012, Emkow 2019). Guillaume (2008) does not consider the possibility that minimality effects and phonological phrasing in Cavineña might not refer to the same domains of structure. The morphosyntactic word may “misalign”, but internally it is a consistent structure. A division between words and phrases (phonological and morphosyntactic) is made, but the potential for identifying intermediate domains or the possibility that divergences between available diagnostics might not align is not considered. However, the patterning of observable phonological domains at any given point plausibly reflects smaller piecemeal changes at specific junctures of structure. What we observe may not be sculpted

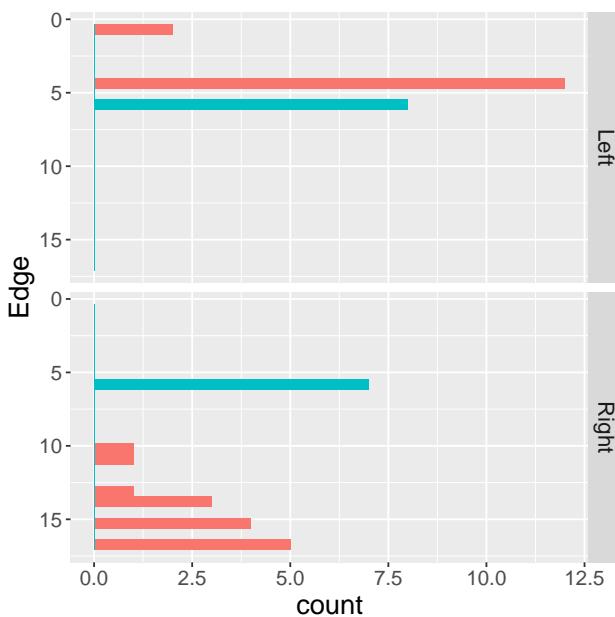


Figure 15: Edges of the constituency test results in Araona

out of universal abstract “phonological words”. I suspect that descriptions in the service of finding support for the formal categories presupposed in BLT misrepresent the degree to which constituency in Takanan languages is a gradient and indeterminate phenomena (Bybee 2001, 2010). I would suggest that the fact that more fine grained descriptions that seek to find convergences only do so sometimes, suggests that language specific history might be playing a more determinative role (Blevins 2004), than the abstract constituent structures assumed in BLT.

## Acknowledgements

I would like to thank Chanito Matawa, Oscar Matawa, Mateo Matawa, Nilda Washima, Eliane Matawa and Freddy Matawa for working with me on elicited Araona examples. I would like to thank the Araona community for welcoming me into their communities, teaching me their language, and telling me their stories. Antoine Guillaume, Gabriel Gallinat, Hiroto Uchihara and Sandra Auderset provided helpful comments on an earlier version of this chapter. All mistakes are my own.

## Abbreviations

A	agent	FOC	focus
ABS	absolutive	FREQ	frequentive
ADJ	adjective	FRUST	frustrative
AFF	affective	INTENS	intensifier
ANT	anterior	INTER	interrogative
AUG	augmentative	INTR	intransitive
CNTRFCT	counterfactual	INTRC	interactional
COMPL	completive	LIMIT	limitative
DEPREC	depreciative	P	patient
DIST	distal	PFV	perfective
DUR	durative	POST	posterior
E	e	REL	relative
EMOT	emotive	RELEV	relevance
EPIS	epistemic	SPAT	spatial
ERG	ergative	TR	transitive
FINAL	final	VIS	visual

## References

- Anderson, Stephen R. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.
- Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
- Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Bybee, Joan. 2001. Frequency effects on French liaison. In Joan Bybee & Paul Hooper (eds.), *Frequency and the emergence of linguistic structure*, 337–359. Amsterdam: John Benjamins Publishing Company.
- Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. & Alexandra Y. Aikhenvald. 2002. Word: A typological framework. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A Cross-linguistic Typology*, 1–41. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511486241.002.
- Emkow, Carola. 2006. *A grammar of Araona, an Amazonian language of North-western Bolivia*. Melbourne: LaTrobe University. (Doctoral dissertation).

- Emkow, Carola. 2019. *A grammar of Araona* (Outstanding grammars from Australia 19). Munich: LINCOM.
- Fougeron, Cécile. 1999. Prosodically conditioned articulatory variants: A review. *U.C.L.A Working Papers in Phonetics* 97. 1–74.
- Garellek, Marc. 2013. *Production and perception of glottal stops*. Los Angeles. (Doctoral dissertation).
- Garrett, Edward. 1999. Minimal words aren't minimal feet. *UCLA Working Papers in Linguistics* 1. 68–105.
- Girard, Victor. 1971. *Proto-takanan phonology*. Berkeley: University of California Press.
- González, Hebe. 1997. *L'araona (langue takana de Bolivie). Eléments de sociolinguistique, phonétique et phonologie*. Lyon: Université Lumière Lyon II.
- Gordon, Matthew K. & Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29(4). 383–406.
- Gowlett, D. F. 2007. Stabilization in Bantu. *African Studies* 2(43). 187–203.
- Guillaume, Antoine. 2008. *A Grammar of Cavineña*. Berlin & New York: Mouton de Gruyter.
- Guillaume, Antoine. 2016. A preliminary investigation of epistemic particles in Tacana (takanan family, Amazonian Bolivia). Laboratoire DDL, Lyon, 21 octobre 2016.
- Guillaume, Antoine. Forthcoming. Takanan languages. In Patience Epps & Lev Michael (eds.), *Amazonian languages: An international handbook*. Berlin: De Gruyter Mouton.
- Gussenhoven, Carlos & Gösta Bruce. 1999. Word prosody and intonation. In Harry van der Hulst (ed.), *Word prosodic systems in the languages of Europe*, 233–271. Berlin: Mouton de Gruyter.
- Harris, Alice. 2017. *Multiple exponence*. Oxford: Oxford University Press.
- Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80. DOI: 10.1515/flin.2011.002.
- Ladd, D. Robert. 2008. *Intonational phonology*. 2nd edn. Cambridge: Cambridge University Press.
- Neely, Kelsey. to appear. Doing language work via social media voice messages: Some tips for workflows and data management. *Special edition on remote methods in language documentation, Language Documentation and Conservation*.
- Pierrehumbert, Janet & Mary Beckman. 1986. *Japanese tone structure*. Cambridge: MIT Press.
- Pitman, Donald. 1980. *Bosquejo de la gramática araona* (Notas Lingüísticas de Bolivia 9). Cochabamba: Summer Institute of Linguistics.

- Pitman, Donald & Mary Pitman. 1970. *La jerarquia fonologica en araona*. Riberalta: Summer Institute of Linguistics.
- Pitman, Donald & Mary Pitman. 1976. Araona syllable and stress. In Dennis Newton, Donald Pitman, Mary Pitman, Martin Krüsi, Elizabeth L. Camp Dorothee Krüsi & Millicent R. Liccardi (eds.), *Work Papers of the Summer Institute of Linguistics*, 1–17. Riberalta: Summer Institute of Linguistics.
- Pitman, Donald & Mary Pitman. 1980. *Historia araona narrada por Bani Huali*. Riberalta: Instituto Lingüístico de Verano.
- Pitman, Mary. 1981. *Diccionario araona y castellano*. Riberalta: Instituto Lingüístico de Verano.
- Rosen, Sara Thomas. 1989. Two types of noun incorporation: A lexical analysis. *Language* 65(2). 294–317.
- Sadock, Jerrold M. 1991. *Autolexical syntax: A theory of parallel grammatical representations*. Chicago: University of Chicago Press.
- Szczegielniak, Adam. 2018. Ellipsis. In *Oxford bibliographies*. Oxford: Oxford University Press. DOI: 10.1093/obo/9780199772810-0218.
- Tallman, Adam J. R. 2021a. *An ethnographically based linguistic documentation of Araona: A takanan language of Bolivia*. <https://elar.soas.ac.uk/Collection/MPI1207738>.
- Tallman, Adam J. R. 2021b. Constituency and coincidence in Chácobo (Pano). *Studies in Language* 45(2). 321–383. DOI: 10.1075/sl.19025.tal.
- Tallman, Adam J. R. & Patience Epps. 2020. Morphological complexity, autonomy, and areality in western Amazonia. In Francesco Gardani & Peter Arkadiev (eds.), *The complexities of morphology*, 230–264. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198861287.003.0009.
- Tallman, Adam J. R. & Gabriel Gallinat. Accepted. The prefix *e-* in Araona (Takana). *Amerindia*.
- Tallman, Adam J. R. & Eitan Grossman. 2022. Body-part incorporation in Chácobo and Araona: Phrasal movement, root merger, or none of the above? SSILA presentation, 21 January 2022.
- Vuillermet, Marine. 2012. *A grammar of Ese ejja, a takanan language of the bolivian amazon*. Lyon. (Doctoral dissertation).
- Vuillermet, Marine. 2014. Two types of incorporation in Ese Ejja. In Swintha Danielsen, Katja Hannß & Fernando Zúñiga (eds.), *Word formation in South American Languages*, 113–142. Amsterdam: John Benjamins.
- Zingler, Tim. 2020. *Wordhood issues: Typology and grammaticalization*. The University of New Mexico. (Doctoral dissertation).