A description of Taruma phonology

Serke, Anna

A DESCRIPTION OF TARUMA PHONOLOGY

Anna K. Serke

A thesis submitted in partial fulfilment of the requirements for the degree of
Master of Arts (Research) in Linguistics

Supervisor: Dr. Konrad A. Rybka

Universiteit
Leiden
The Netherlands

Abstract

Taruma is a language isolate whose last three speakers live in a Wapishana community in the south of Guyana and which is poorly described and practically undocumented. As a first step toward remedying this issue, this thesis offers a description of Taruma phonology and includes both a wordlist and a transcribed oral history in the appendices. The analysis is based on recordings made as part of a remotely conducted ethnobotanical language documentation project as well as during a month of fieldwork in Guyana. After the description of the segmental inventory and suprasegmental features, an analysis of the (morpho)phonological processes, the phonological adaptation of loanwords, and imitative sound symbolism is given. Finally, the thesis suggests an orthography for Taruma. It is apparent both in its phonology and its vocabulary that Taruma, while an isolate, does not exist in a linguistic vacuum: the lexicon contains several words which are shared with other Indigenous languages, while the speakers' bilingualism leads to the occasional transfer of two phonological processes from Wapishana to Taruma. Several avenues for future research remain, which highlights the importance of continued collaborative language description and documentation.

Keywords: *Taruma, phonology, language contact, orthography, Wapishana*

Acknowledgements

Table of Contents

Tables

Figures

List of abbreviations

1SG	first person singular
1SG.OBJ	first person singular object
1SG.POSS	first person singular possessive
2PL.POSS	second person plural possessive
2SG	second person singular
2SG.POSS	second person singular possessive
3.POSS	third person possessive
3	third person
3.OBJ	third person singular object
ANPH.LOC	anaphora locative
C	consonant
COL	collective number
DEM	demonstrative
EGIDS	Expanded Graded Intergenerational Disruption Scale
ELAR	Endangered Languages Archive
EP	epenthetic vowel
EXIST	existential
HESI	hesitation
IMP	imperative
INS	instrumental
IPA	International Phonetic Alphabet
L1	first language
L2	second language
LOC	locative
MOAA	Ministry of Amerindian Affairs
N	nasal
NEG	Negation
NS	non-subject
PL	plural
PRP	purposive
SAPhon	South American Phonological Inventory Database
SG	singular
TAM	tense / aspect / mood
V	Vowel
WLP and WWA	Wapishana Language Project, and Wapichan Wadauniinao Ati'o

1 INTRODUCTION

Taruma is a highly endangered, poorly described and until now virtually undocumented language isolate whose last speakers live in what is now Guyana. The language was long thought to be extinct (Carlin and Mans 2015) and is even absent from both the most recent volumes on the Indigenous languages spoken in the country and from the list of Indigenous languages named by the Guyanese Ministry of Amerindian Affairs (2022b, 2022c). However, there are still at least three Taruma speakers who live in Turunau, a part of the Maruranau village in the south-west of the country. In Maruranau, the primary language spoken is Wapishana[1], an Arawakan language which the Taruma also speak. They agreed to share their language and knowledge as part of a remotely-conducted interdisciplinary project aiming to document Taruma names and uses of various local plants. This thesis aims to address the gap in the literature relating to Taruma, and together with the ethnobotanical documentation project, it constitutes a basis for further descriptive research and language documentation. It is structured as follows. The remainder of this section offers background information about the research presented in this thesis and summarises the information that is currently available about the Taruma people, their language, culture, and history. In the second section, the methodology is described, introducing the project team and outlining the collection and management of the data. In section three, the phonological description and analysis is given, which details the phoneme inventory, (morpho-)phonological processes, phonotactics, stress, the phonological adaptation of loanwords, and sound symbolism. The section concludes by presenting an orthography for Taruma, as well as by comparing the two existing historical Taruma wordlists to the contemporary data. Lastly, the typological characteristics of Taruma phonology are summarised and topics for further research are indicated. In the appendices, a Taruma-English wordlist as well as a transcribed oral history are provided.

1.1 The languages of Guyana

The linguistic landscape of South America is characterised by a high degree of diversity, with 420 Indigenous languages representing 108 language families (Campbell 2012a: 59). Guyana is no exception; it is home to eleven Indigenous languages, representing two language families and two language isolates (Eberhard et al. 2022a). Table 1, based on Carlin and Mans (2015: 78) and Eberhard et al. (2022a), lists the languages and provides their vitality scores according to Lewis and Simon's (2010) EGIDS (Expanded Graded Intergenerational Disruption Scale).[2] The scale consists of thirteen levels from 0 to 10, with 6 and 8 split up into two levels a and b.

---

[1] Many Wapishana language activists prefer the spelling *Wapichan*, as it is more similar to the pronunciation of the word in their language; however, the anglicized spelling *Wapishana* is more common in the English-language literature and is also used by some Wapishana (Gomes 2022: 17). Therefore, throughout this thesis, *Wapishana* is used, in keeping with Gomes 2022: 17.

[2] It should be noted, however, that many sources listed by Eberhard et al. (2022a) were published more than fifteen years ago. Some of the languages spoken in Guyana are likely more severely endangered today.

Level 0 denotes an international language, whereas level 10 corresponds to the extinction of a given language.

Table 1. Indigenous languages spoken in Guyana (Carlin and Mans 2015: 78, Eberhard et al. 2022a).

Language family	Language	EGIDS vitality score	
Arawakan	Atorada	8a	moribund
	Arawak	8a	moribund
	Wapishana	6b	threatened
	Akawaio	6b	threatened
	Pemon	6a	vigorous
Cariban	Carib	6b	threatened
	Makushi	7	shifting
	Patamona	6b	threatened
	Waiwai	8a	moribund
isolate	Taruma	8b	nearly extinct
isolate	Warao	8a	moribund

Most of the Indigenous languages spoken in Guyana belong to one of two language families: the Arawakan family, which Atorada, Arawak, and Wapishana form part of, and the Cariban family, including Akawaio, Pemon, Carib, Makushi, Patamona, and Waiwai. The two exceptions are Taruma and Warao, both of which are language isolates. All these languages are threatened to a greater or lesser extent. Arekuna is the only language that is used across all generations and by a stable number of speakers. Akawaio, Kari'na, Patamona and Wapishana are also used by all generations, but are not passed on to the entirety of the younger generations, so that the number of speakers is decreasing. In the case of the other languages, intergenerational transmission of the language to the younger generations has all but ceased; they are now mainly spoken by older generations. Taruma is classified as nearly extinct, as the remaining speakers belong to the older generation and do not use the language actively due to a lack of opportunities to speak it, as detailed further in section 1.5.

The endangerment of the languages results from many speech communities shifting either to a more dominant Indigenous language, a creole that acts as a lingua franca or to a national language (Carlin and Mans 2015: 79). The national language of Guyana, which is a former British colony, is English; additionally, Guyanese English Creole, which Eberhard et al. (2022a) refer to as the de facto language of national identity, is spoken by 643,000 out of the approximately 743,700 inhabitants of the country (Bureau of Statistics 2021). Non-Indigenous minority languages spoken in Guyana include Portuguese, Spanish and Chinese (Gomes 2022: 19).[3] According to Carlin and Mans (2015: 79), the shift to other Indigenous languages occurs due to the movements and fusion of different ethnic groups, often resulting in multi-facetted or nested cultural identities. When these are renegotiated, the speakers' original languages are not necessarily retained as expressions of the said identities (Carlin and Mans 2015: 97-99). In the case of Wapishana, speakers shift to the national language English rather than to a different

---

[3] It is unclear which Chinese language(s) are being referred to; Eberhard et al. (2022a) also only list Chinese as a macrolanguage without providing further details.

Indigenous language (Carlin and Mans 2015: 79). This is for a variety of reasons, most prominently the fact that the language was consistently stigmatised by religious and educational institutions established during colonial times (Gomes 2022: 39-52). Presently, the institutional pressure no longer exists in this form, but some parents still consider it more advantageous for their children's future to raise them with English as their first language instead of Wapishana (Gomes 2022: 46; see §1.5).

Today, governmental institutions such as the Ministry of Amerindian Affairs (MOAA) provide support for Guyana's Indigenous languages. It was established in 2006 with the mission statement "[t]o enhance the social, economic and environmental well-being of Indigenous Peoples and their lands through collaboration, sustainable development and appropriate legislation, while at the same time ensuring the preservation of Indigenous culture and traditional knowledge" (Ministry of Amerindian Affairs 2022a). The preservation of Indigenous culture also includes the native languages spoken in Guyana, which are promoted by "creating programmes and initiatives through dictionary production, written literature and music development" (Ministry of Amerindian Affairs 2022b). However, it is unclear how much of the Ministry's annual budget is spent on language preservation efforts; for instance, the two projects named in the budget for 2022, are concerned with policy development and administration as well as with community development and empowerment (Ministry of Finance Guyana 2022a: 175). The latter lists heritage preservation as one of its objectives, but it is not detailed in the breakdown of the expenditures which steps will be taken to accomplish this objective with regard to the country's languages. For details, see also Ministry of Finance Guyana (2022b: 131-137; 2022c)). Furthermore, the list of recognised languages on the MOAA's website does not include Taruma or Atorada, presumably because neither language is spoken by an entire community anymore (Ministry of Amerindian Affairs 2022c). Thus, today, Guyana's Indigenous languages are institutionally supported to some extent. But this may not be enough to change the process of language shift observed in many speech communities, especially since it is influenced by other complex social factors connected to the speakers' attitudes and cultural identity.

1.2 Sources about the Taruma language and its genetic affiliation

The Taruma people were first mentioned by European colonizers in Portuguese texts from 1657 and were thereafter referenced occasionally over the course of the next two centuries (Carlin 2011: 228). However, more detailed information about the people, their culture and language is only found from the mid-nineteenth century onwards; the principal external historical sources on the Taruma are the ethnographic treatises by Schomburgk (1845, 1848) and Farabee (1918) as well as notes left by Fr. Cary-Elwes (see Butt Colson and Morton 1982).

Robert Schomburgk (1848: 47) travelled through the British colony between 1835 and 1839 as well as between 1840 and 1844 in order to survey its geography and natural history on behalf of the British Empire. Other than geographical and meteorological information, his reports contain descriptions of the Indigenous villages and the peoples whom he encountered during his travels. Among these are two Taruma settlements, in the context of which he reports on the architecture of the houses, the customs, and the language of the Taruma (Schomburgk

1845: 34-48). In addition, Schomburgk (1848: 59-60) collected small wordlists in each of the eighteen languages that he distinguished. The Taruma wordlist, later reproduced in Latham (1862) and Martius (1867), contains only eighteen lexical items: the numbers one to three, some body parts and words relating to nature (see appendix A.1.1). Importantly, Schomburgk (1845: 35) also noted that Taruma differs notably from Makushi and other Cariban languages, and considered it unrelated to any other Indigenous languages that he had encountered (see also Schomburgk 1848: 47). This assessment was based on the observation that speakers "utter[] the first syllable of each word strongly through the nose"; as Schomburgk remarks, it is "that nasal twang–that starting and raising of the voice, that renders their dialect one of the most remarkable I have hitherto visited". Further, Schomburgk (1845: 45) notes that "[t]hey possess the sound of the English th and the Greek ph". The implications of these remarks for the phonology of Taruma are discussed in §3.9.

Between 1913 and 1914, the American anthropologist William Farabee collected data on several Indigenous nations of the colony (Farabee 1918, 1924). His ethnographic publication, which was focused on the speakers of Arawakan languages, provides information on the material culture, dietary habits as well as spiritual and other cultural elements of the nations he visited (Farabee 1918). This includes the Taruma, whom he considered to be "related linguistically to the Wapisianas" (Farabee 1917: 72). Farabee (1918: 135-158) gives an overview of their history, describes the architecture of their houses and some customs relating to marriage and death. Further, he reproduces English-language translations of a few oral traditions, including those about the origin of mankind, fire and nature spirits. His work also includes a Taruma wordlist which is more extensive than that of Schomburgk. Comprising some 230 words, the wordlist contains numerals up to fifteen, terms for several animals, plants and objects, kinspeople, adjectives, and adverbs (Farabee 1918: 277-283; see appendix A.1.2).

Fr. Cuthbert Cary-Elwes was a British Roman Catholic missionary who, from 1910 to 1923, established and lived in a mission by the Takutu River near the Guyanese-Brazilian border (Butt Colson and Morton 1982: 203-206). There, he mainly worked with the Makushi, but he also travelled occasionally in order to evangelise other nations. He visited the Taruma and Waiwai a total of four times: twice in 1919, and once in 1922 and 1923 (Butt Colson and Morton 1982: 206-207). While his edited diary entries about these visits can be read in Butt Colson and Morton (1982), he also left texts and wordlists in Taruma, Makushi, Waiwai and Wapishana that remain unpublished (Butt Colson and Morton 1982: 216).

More recently, Dr Sérgio Meira made recordings with the Taruma during a brief visit in 2015, which consisted of oral traditions, songs, and elicitations of wordlists for flora and fauna, kinship terms and common adjectives and verbs. These data and the corresponding transcriptions are, as of yet, unpublished but form part of the materials on which the analysis presented in this thesis is based (see §2 for the methodology of data collection).

Other than Rivet (1924: 650), who agrees with Farabee's assessment of Taruma as an Arawakan language, and Kaufman (2007: 73, cited in Campbell 2012a: 136), who proposes a genetic affiliation with Katembrí, most authors endorse the classification of Taruma as a language isolate. As Loukotka (1949: 55) notes, the lexical similarities between Taruma and the Arawakan languages Wapishana, Atorada, and Mawayana are insufficient proof for a genetic relationship but rather point towards borrowing due to contact. Consequently, Taruma

is listed as an isolate in Carlin and Mans (2015: 77), in Campbell's (2012a: 70) classification of South American languages and on Ethnologue (Eberhard et al. 2022b).

As Taruma grammar has not been described, the information available on this topic is mainly inferential. For instance, Carlin (2011: 234) points out that according to Fr. Cary-Elwe's notes, many Taruma struggled to learn Waiwai, which could be an indication that the two languages were structurally quite distinct. However, Carlin (2011: 234) also notes that passive bilingualism is not necessarily uncommon in the Guianas; in these scenarios, speakers of different languages understand each other's languages but only use their own language when interacting. If this was the case for the Taruma and Waiwai, then the Taruma speaking Waiwai with difficulty does not provide a reliable indication as to the extent to which the two languages differ. Further, it was assumed that the Taruma phoneme inventory includes implosives (Carlin 2011: 233); however, this is not the case in the data which this thesis is based on. The misconception was presumably caused by the fact that the Taruma speakers in Maruranau primarily use Wapishana, which does have a bilabial and a dental implosive (Gomes 2022: 142-143). The phoneme inventories of the two languages do resemble each other somewhat, as will be relevant for the Taruma orthography presented in section 3.8, so that the two could easily have been confused by a researcher from outside the community.

1.3 History and language contact

The literature on the history of the Taruma people is quite limited. This subsection first describes what we know about the historical contact situations with other nations, before presenting the two competing theories as to the origin of the Taruma.

1.3.1 Attested history and contact situations

The Taruma are known to have lived and traded in the former Dutch colony near the Essequibo River, located in present-day Guyana, from 1764 until the early twentieth century (Carlin 2011: 228-30). During this time, they lived in proximity to the Mawayana, an Arawakan people, and had close trade relations with the Wapishana, the Waiwai and the Makushi from the 1830s at the latest (Carlin 2011: 230). Carlin (2011: 230) explains that they primarily traded in hunting dogs and objects such as pottery and cassava graters, which were widely sought after by other Indigenous nations, but also Western wares that were brought in from the coast (Carlin 2011: 230). Contact with Europeans is thus also attested, with additional evidence in the form of products such as glass pearls and fragments of bottles found on the archaeological sites near the Essequibo River (Boomert 1977: 16). In spite of the contact situations and the resulting intermarriages, especially with the Mawayana, the Taruma did not give up their language or identity until the 1920s (Carlin and Mans 2015: 84). However, it is unclear in how far these two languages might have influenced each other, beyond the fact that Taruma, Wapishana and Mawayana share loanwords for trade items. Carlin and Mans (2015: 84) suggest that the Taruma and the Mawayana would have either had passive knowledge of each other's languages, or used a third language known to both. From the 1920s, the Taruma began to be

absorbed by the Wapishana, Waiwai and other nations through intermarriage. As a result, by the 1960s, the Taruma language had largely disappeared and was thought to be extinct, as most speakers shifted to Waiwai and Wapishana (Carlin and Mans 2015: 85). This development is thought to have been effected partly by Fr. Cary-Elwes, who advised them to assimilate to other ethnic groups (Carlin 2011: 231).

1.3.2 Origin and migration: The Rio Negro and the Essequibo River homeland theories

The history of the Taruma preceding 1764 is debated, especially with regard to their homeland. Portuguese colonial texts from 1657 onwards make reference to a nation called *Taruma* but place it near the Rio Negro in present-day Brazil (Carlin 2011: 228). By 1770, this nation seemed to have vanished from that region and was initially thought to have dispersed (Carlin 2011: 228). But Schomburgk (1845: 35) and Farabee (1918: 135) theorised that the Rio Negro Taruma had instead migrated northward to the Essequibo (Carlin and Mans 2015: 82). This is visualised in Figure 1.
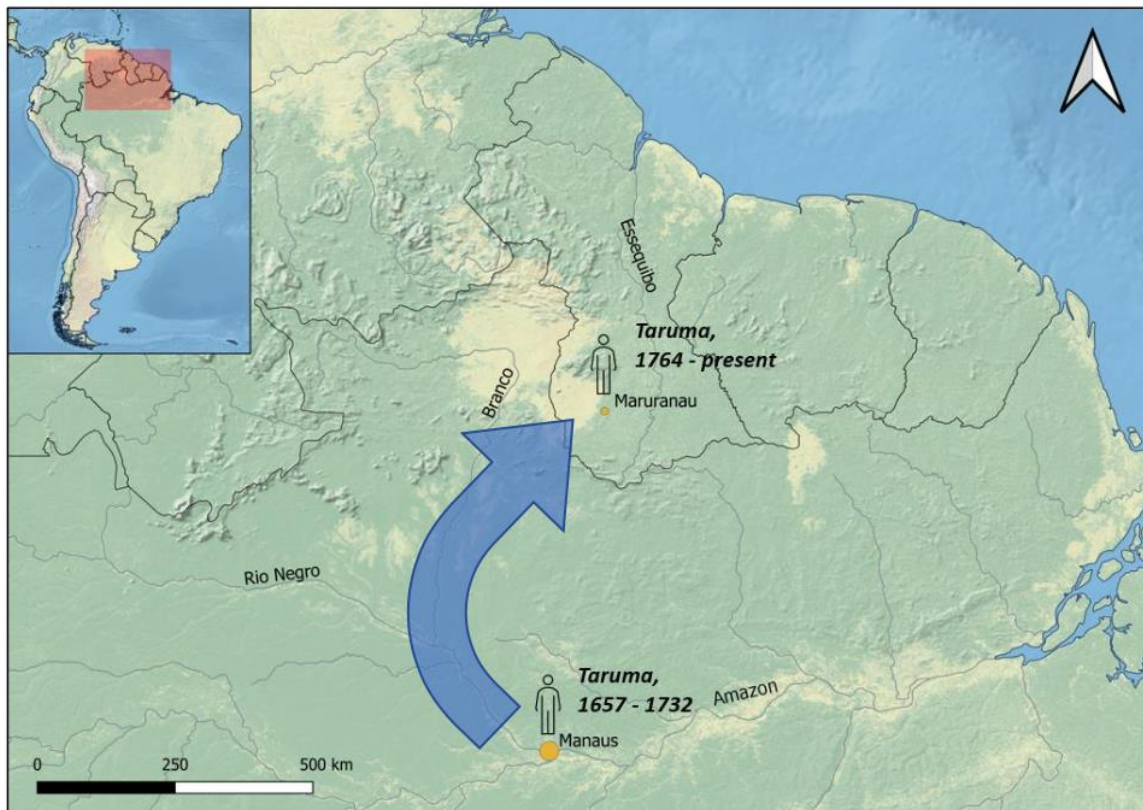
Figure 1. Hypothesized migration route of the Taruma from the lower Rio Negro in Brazil to southwest Guyana (Bredero 2021: 9).

The arguments in favour of the theory identifying the Rio Negro Taruma with the Essequibo Taruma are based on lexical data and archaeological findings. Carlin (2011: 230) writes that many toponyms in southern Guyana, especially referring to bodies of water, are of Taruma origin. This is taken as an argument for the Essequibo origin theory as it indicates cultural dominance of the Taruma in the region. Additionally, based on the lexical material listed by Schomburgk (1849) and Farabee (1918), Loukotka (1949: 57) notes that the

similarities between certain words in Taruma and other languages point to borrowing due to language contact. This concerns several Arawakan and Cariban languages, but also Makú, Mura and Chibcha, which are isolates or have no sister languages in Guyana. Many of these languages were spoken south of the Essequibo territory, closer to the Rio Negro, but in many cases, the lexical correspondences are tentative at best (Loukotka 1949: 57).

Further support for the Taruma homeland near the Rio Negro is given by Boomert (1977: 16-17), who argues based on archaeological evidence. In the 1950s, the remains of several former Taruma settlements were unearthed near the Essequibo River (Boomert 1977: 15). The objects found most frequently on the sites were cassava graters and shards of pottery, the latter of which featured distinctive ornamental patterns (Boomert 1977: 16). These were made with red paint or by means of punctuating or stamping ornamentation onto the pottery, and include zig-zag or crossing lines as well as triangles (Boomert 1977: 16). Similarly decorated clay wares from the eighteenth century have been found both in the Paru savannah in Brazil and on sites in southern Suriname, in an area bordering south-eastern Guyana. The Surinamese sites were inhabited by a nation referred to in the literature by the name *Saluma* as well as several other names resembling *Taruma* (Boomert 1977: 16).[4] Boomert (1977: 16-17) assumes that the Saluma belonged to the same ethnic group as the Guyanese Taruma based on the relative geographical proximity of their settlements and the similarity of both the exonyms and the pottery. Since the shards found on the neighbouring sites, which were inhabited by other nations such as the Trio, did not feature any of the distinctive patterns described above, Boomert sees these archaeological finds as confirmation that this type of pottery is unique to the Taruma. Furthermore, similar shards were found in early nineteenth-century settlements of the Manao, an Arawakan nation. Since the settlements are located close to the Rio Negro, Boomert (1977: 17) argues that the Manao made similar pottery due to previous contact with the Taruma when the latter still lived in that region. Boomert (1977: 17) takes this as confirmation of the theory that the Rio Negro was the Taruma homeland. The problem with this interpretation, however, is that the Manao also migrated to and lived in the Essequibo region in the eighteenth century (Boomert 1977: 14; Carlin 2011: 229). Therefore, the similarity of the pottery could also result from contact that took place in the Essequibo region instead of near the Rio Negro.

The Rio Negro homeland was the accepted hypothesis until Rivière (1966: 303) called it into question, arguing that the evidence that the Essequibo Taruma are the descendants of the Taruma from the Rio Negro is poor. Alternatively, Rivière (1966: 303) suggested that the origin of the Essequibo Taruma lies at the Essequibo River, and that no large-scale migration took place. One reason for this assessment is that Rivière considers the similarity of the names of the nations living near the Rio Negro and the Essequibo River too unreliable as evidence and points to other nations with similar names. These include the nations discussed by Boomert (1977), but also the Turroomaes who lived in Suriname near the upper part of the Suriname

---

[4] The Saluma are also referred to as "Taruma, Zaruma, Tcaruma, Taroeman, Tsaruma [and] Tcharuma" in English and French sources, as "Saluma, Saloema, Salmoës, Saruma" in the Dutch and as "Charuma, Charuman, Xaruma [and] Xuruma" in the Portuguese literature (Boomert 1977: 15). They are known to have lived in the Kutari and Sipaliwini regions in the nineteenth and twentieth century (Boomert 1977: 14). After 1953, they moved to the Turunu region and no further historical information on their whereabouts or dispersal is available (Boomert 1977: 15).

River in northern Suriname (Rivière 1966: 303). Due to the resemblance of their names, these should be positively identified with the Essequibo Taruma with equal likelihood, Rivière (1966: 304) argues. His second counterargument is that the claim of migration from the Rio Negro to the Essequibo cannot be traced back to a source; according to Rivière (1966: 304), Schomburgk hypothesised about it in 1840, whereupon von Martius (1867: 567-568) treated it as fact in his publication. It should be noted here that Farabee (1918: 135) writes that the Taruma "have a tradition that they came from the south". Unfortunately, he does not provide any further details on this oral history nor on its source, so that it is difficult to judge the validity of the claim. Therefore, the possibility cannot be excluded that his report is based on Schomburgk's theory.

Bredero's (2021: 29) analysis of the Taruma plant terms recorded for the language documentation project also suggests that the Essequibo Taruma and the Rio Negro Taruma are different nations. Most of the loanwords contained in the plant data were borrowed from languages spoken near the Essequibo, mainly Wapishana, Mawayana, and Waiwai. However, one word was likely borrowed from Bahuana, a language spoken near the Rio Negro, and Bredero (2021: 29) points out that the results might be skewed to the lack of linguistic data from that region.

The attempt to reconstruct a potential migration of the Taruma is also complicated by the fact that the name *Taruma* is an exonym of unclear origin (the autonym is hodjasu [ho.ˈdʐa.sɨ] in our data). Carlin and Mans (2015: 83) point to Wapishana as a potential source, with *taruma* translating to 'stinging ant'. However, they also call this theory into question, since the name was already in use for the Rio Negro Taruma in the mid-seventeenth century, while contact between the Taruma and the Wapishana has only been attested in the Essequibo region from the eighteenth century onwards (Carlin and Mans 2015: 83). Due to the unknown source of the exonym, the possibility cannot be excluded that the similarity of the names *Taruma* and *Saluma* is accidental, which makes it difficult to prove Boomert's claim that they belonged to the same nation.

1.4 The Taruma speakers and their current sociolinguistic situation

The Taruma speakers live in Turunau, which is a part of the village Maruranau located in Region 9 in South-Western Guyana (Figure 2). The village spans an area of nearly 110 square miles in the Rupununi savannah and counts a population of approximately 800 (Ministry of Amerindian Affairs 2022d). It is one of twelve villages located in the *Wapichan Wiizi*, the Wapishana traditional lands, in which the majority of the inhabitants are Wapishana (Gomes 2022: 28). Language use in these communities can be described as diglossic (Gomes 2022: 49). While the term *diglossia* originally refers to "two distinct varieties of the same language [being] used in a community, with one regarded as a high (or H) variety and the latter (a low or L) variety" (Holmes 2013: 27), Gomes (2022: 49) proposes to extend it to denote situations in which an Indigenous and an international language are spoken. In communities such as Maruranau, Wapishana is the first language and the L variety, associated largely with informal and spoken use (Gomes 2022: 30, 49). English, on the other hand is the H variety, as it is the main language used for writing, instruction at the local school as well as formal occasions such

as visits of government officials. Due to its higher status compared with Wapishana and its association with better chances at educational success, some younger parents bring up their children with English as their first language despite being Wapishana-L1-speakers themselves (Gomes 2022: 46).[5] Some people also speak other languages, such as Waiwai, Makushi, or Portuguese, but this occurs on the level of a single person or an individual family rather than community-wide (Gomes 2022: 30).[6]
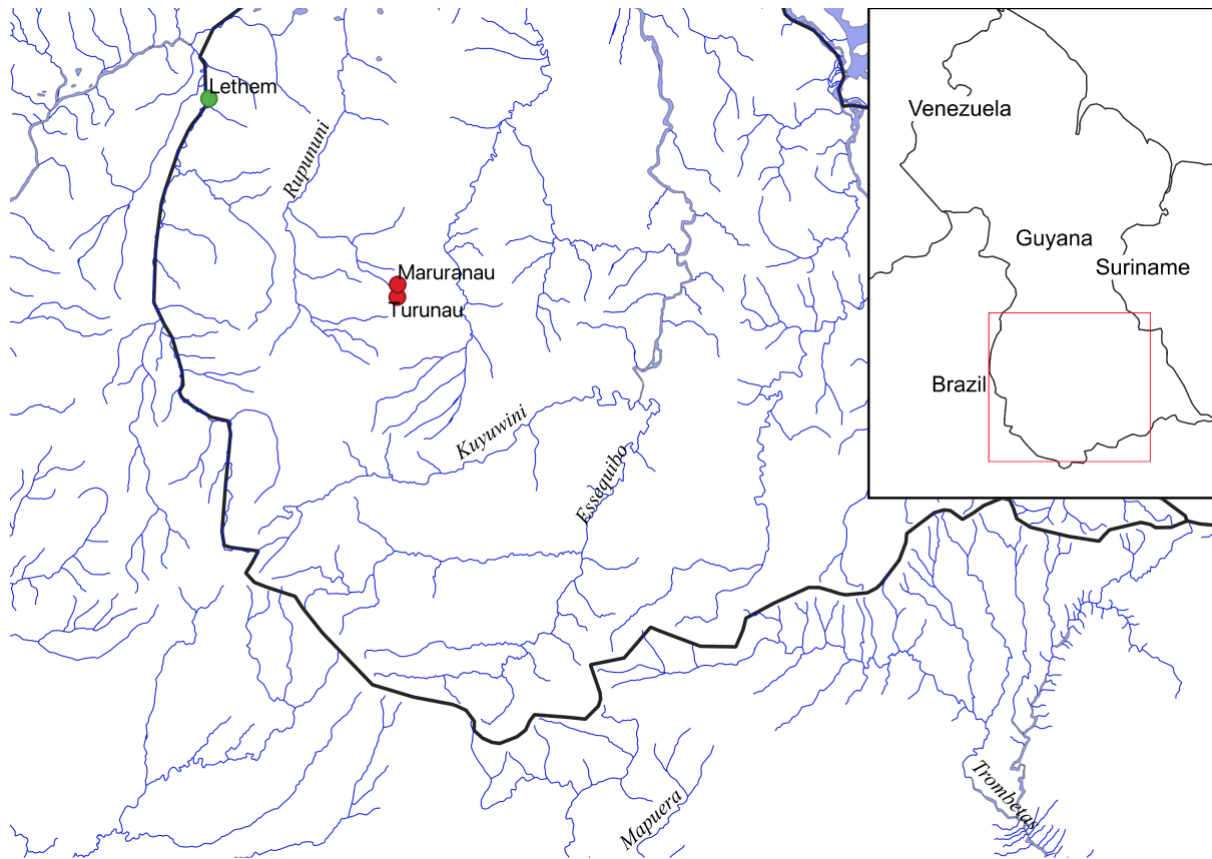
Figure 2. Location of Turunau and Maruranau, Guyana. Created with QGIS.

As the Taruma speakers live in Maruranau today, Wapishana is the principal language they use in daily communication. They are three siblings who were born in the 1950s in a village in the rainforest and later moved to Maruranau, where they have been living since. In childhood, they acquired both Taruma and Wapishana. The former was their mother's first language, whereas they learned the latter from their father. The data used for this thesis was provided by two of the siblings, Ms Irene Suttie and Mr Vincent Louis. As a child, Ms Suttie was exposed to Taruma more as she spent more time with their mother than did Mr Louis. As a result, Ms Suttie's bilingualism is balanced, whereas Wapishana is Mr Louis' dominant language. Due to marriage with Wapishana speakers, this language is used to communicate not only with other

---

[5] Efforts are made to combat this attitude, promote Wapishana literacy and establish it more firmly as a written medium, such as by means of a Wapishana-English dictionary, which was published by the Wapishana Language Project in 2000, as well as through various types of workshops; Gomes (2022) also advocates for the expansion of bilingual education programmes in Wapishana schools.

[6] For more information on the Wapishana communities in Guyana, see Gomes (2022).

members of the community but also within their wider family. As a result, they have few opportunities to speak Taruma and often speak Wapishana even to each other. Both speak some English as well.

For Mr Louis' two daughters, Ms Elizabeth and Ms Nita Louis, Wapishana is their first language. It is the main language spoken in their speech community, in which they were born in the 1990s and have lived for most of their lives. Furthermore, since their mother is a Wapishana-L1 speaker, it is also the language that their parents share and is therefore the one used in their family. Both Ms Elizabeth and Ms Nita Louis also speak English, which they acquired at school from age six onwards. In their daily lives, they speak Wapishana in most of their social interactions, whereas they use English principally with non-Wapishana speakers, usually people from outside Maruranau. Ms Elizabeth Louis also speaks Portuguese due to having lived in Brazil for a year. Both have learned some Taruma words and expressions from Mr Louis and Ms Suttie and are interested in learning more of their heritage language as well as in documenting it.

1.5 The culture of the Taruma

This subsection lays out the elements of Taruma culture that are described in the literature, principally in the sources from the nineteenth and early twentieth century, beginning with architecture and material culture, food production, and finally customs and oral traditions.

The houses of the Taruma were described by Schomburgk (1845: 34) as being dome-shaped or, on one occasion, in the shape of a pyramid, with roofs made of palm leaves. Farabee (1918: 136) wrote that they resembled Wapishana houses in so far that the walls were similarly thatched. But in contrast to the Wapishana, the Taruma did not seem to add mortar to the walls (Farabee 1918: 136). In terms of material culture, their cassava graters receive special attention in Schomburgk's (1845: 38) account: made from a plank of soft wood into which small rocks were hammered and secured with tree resin, the graters were traded with the Wapishana. However, according to Farabee (1918: 21), the Taruma did not make the graters themselves but received them from the Waiwai. Canoes, bow and arrow as well as hammocks were indispensible in everyday life, but are merely mentioned in passing in the historical sources without details as to the materials or techniques used to make them. Nor is the above-mentioned pottery described by either ethnographer. Other objects that are not described in the historical sources include aprons made by means of loom beadwork, necklaces made from glass beads or seeds as well as plaited baskets that are dyed in geometric patterns. However, such Taruma objects were collected by Alpheus Hyatt Verrill (1871-1954) and are currently held by the National Museum of the American Indian in Washington, D.C. and New York, N.Y. (Smithsonian Institution 2021).

With regard to food production, we know that the Taruma cultivated cassava, sugar cane and pineapples (Schomburgk 1845: 34, 36, 42), as well as corn (Farabee 1918: 136). Farabee (1918: 136) considered them the "greatest fish eaters of all tribes" and in addition, they were skilled hunters (Schomburgk 1848: 40, 34-35). The Taruma also kept poultry at the time of Schomburgk's (1845: 36) visit, namely trumpeter and curassow birds. They were also experts

in training dogs for hunting, which is why their dogs were so popular with their trading partners (Schomburgk 1845: 36).

We know very little about Taruma customs, which are mainly mentioned anecdotally. Farabee (1918: 141) reports that the dead were cremated along with personal belongings, with the exception of healers, who were buried. Schomburgk (1845: 45, 44) further notes that intermarriage with other nations was not approved of, and that on festive occasions, circular group dances were performed, accompanied by singing and music made with rattles. Records of Taruma narratives are slightly more extensive, as Farabee (1918) recounts several origin stories and oral traditions about animals. Among the former are the stories of how the first humans, two brothers, found a woman, how this woman was the source of fire and how she taught agriculture to the men (Farabee 1918: 143-149). The latter include narratives about an anthropomorphic anaconda, parrots, and tortoises (Farabee 1918: 149-150, 152-157). Finally, there are two narratives warning of dangerous nature spirits residing in pools and in rocks, respectively (Farabee 1918: 151-152, 157-158).

2 METHODOLOGY

This section describes the methodology adopted for this thesis as part of the ethnobotanical language documentation project. The aim of the project was to document the Taruma language and plant-related knowledge, in the form of photographs of the plants and corresponding audio recordings. First, the project team is introduced, followed by a discussion of the methods of data collection, which initially took place remotely and later through on-site fieldwork. Finally, the data management is described, including the archiving of the files. The specific roles of the researchers working on the project are listed in table 2.

Table 2. The members of the research team and their tasks.

Team	Member	Tasks
on-site	Vincent Louis	identifying Taruma taxa in the field and describing their uses
	Irene Suttie	cross-checking Taruma taxa from photos and describing their uses
	Elizabeth Louis	creating photo vouchers, recording plant uses, translating
	Nita Louis	creating photo vouchers, recording plant uses, translating
	Adrian Gomes	coordinating activities in Maruranau, budget-keeping
	Konaukii Gomes	sending data via WhatsApp, facilitating communication
off-site	Robin Bredero	processing data, species identification, linguistic comparison
	Anna Serke	transcribing Taruma and Wapishana recordings
	Anne-Marie Holt	evaluating the remote research methodology
	Tinde van Andel	supervising ethnobotanical research
	Konrad Rybka	supervising linguistic research, coordinating the project

The on-site team in Maruranau consisted of the Taruma speakers, their relatives, and Dr Adrian Gomes and Ms Konaukii Gomes. Mr Louis and Ms Suttie are not only among the last remaining speakers of Taruma, but also have a vast cultural knowledge of the various nutritional and medicinal uses of the local plants. They have farms on which they cultivate several types of

plants for consumption but also to manufacture objects from them, and Mr Louis is also a skilled healer. As Wapishana was the most suitable working language for the project, the translations into English provided by Ms Elizabeth and Ms Nita Louis were indispensable. Dr Adrian Gomes and Ms Konaukii Gomes constituted the link between the researchers on-site and off-site. Due to the interdisciplinary nature of the project, the off-site team members, who worked on the data remotely, specialise in botany (Prof dr van Andel, Mr Robin Bredero, Ms Anne-Marie Holt) and linguistics (Dr Konrad Rybka, myself).

2.1 Data collection

2.1.1 Remote data collection

The remote part of the project took place from April to October 2021. During this time period, Ms Nita and Ms Elizabeth Louis used a smartphone in order to take pictures of plants in their entirety as well as of the individual parts, such as leaves, fruit and flowers (Figure 3). They also made accompanying audio recordings with Mr Louis and Ms Suttie. Ms Gomes then sent the data to the off-site team members via WhatsApp using a WIFI hotspot in the village. The individual voice messages were then concatenated by Mr Bredero, so that the audio snippets from a given day formed a single longer audio file. As a basis for phonological analysis, the Taruma and Wapishana data was transcribed in ELAN, and PRAAT was used to draw the pitch contours of different intonational patterns. Based on the photographs and transcriptions, Mr Bredero identified the plant species and compared their names to the corresponding plant terms in other Indigenous languages in order to identify possible borrowings, while Ms Holt assessed the merits of the remote research methodology.

The photographed plants included crops that the speakers cultivate in their gardens or on their farms, but also plants that grow wild in the nearby rainforest. The content of the corresponding recordings consisted of the names of the plants and plant parts, which the speakers repeated at least three times in each of the three languages. This way, the words occurred often enough in the data to mitigate small differences in pronunciation. Additionally, Mr Louis and Ms Suttie recorded descriptive sentences in Taruma, which they then translated into Wapishana. The Wapishana sentences, in turn, were translated into English by Ms Elizabeth and Ms Nita Louis.

Figure 3. Photographs taken by the in-situ researchers of the banana plant (Musa sp.). Images from top left clockwise: whole plant, stem, leaf, fruit, and flower (Holt 2021: 17). [7]

---

[7] The person in the photo asked for their face not to be published.

In the longer recordings, the speakers explained the uses of the plants as food, medicine, or materials from which to manufacture objects, and provided descriptions of how to do so. The information also included terms for different species of a given plant, such as various types of cassava, as well as the names of tools that are essential for the preparation of food, such as cassava sifters and graters. Recordings sent in the later stages of the remotely conducted project included data that were not directly related to plants, such as geographical and meteorological vocabulary, as well as the personal history of Ms Suttie and her ancestors. These were again translated by Ms Nita and Ms Elizabeth Louis. The benefit of the Taruma speakers closely working together with the heritage speakers in this way was its revitalising effect. Before the beginning of the project, the heritage speakers had already been interested in expanding their knowledge of Taruma. Since they heard the plant names being repeated several times and translated the words and sentences into English, the project provided additional opportunities for intergenerational transmission to take place.

The remote method of data collection had two additional advantages: firstly, it was a good solution for a time in which in-person fieldwork was not possible due to Corona-related travel restrictions. Secondly, the speakers were able to prepare for the recordings and curate what they wanted to share, and do so when it fit in best with their daily activities. This way, they were not affected by demands on their time at specific moments that might have interfered with their schedules, which can be the case during in-person fieldwork. However, there were also some problems that arose due to this method; the principal issue was the rather unstable internet connection in Maruranau, especially when the wifi hotspot ceased to function and remained broken for months. This hampered the progress of the project considerably, as the data reception in and around the village is insufficient to send larger files. The next closest access to wifi is located several miles away in Shea, a different village, and it was not viable for the consultants to travel there frequently. As a result, it was difficult for the off-site team to ask questions about the recordings and receive clarifications from the on-site team. When the possibility for in-person fieldwork arose despite the COVID-19 pandemic, it was beneficial for the linguistic analysis that we had worked on the files before. The initial remote research was helpful for planning which kind of data to elicit, which questions to ask and which aspects of the recordings to clarify more.[8]

2.1.2 On-site data collection

In December 2021, it was possible to travel to the village for a month of on-site fieldwork in Maruranau. During this time, Dr Rybka and I made several additional audio recordings with the speakers. They consisted mainly of further elicitations concerned with more botanical material and also provided the opportunity to better parse the Taruma and Wapishana sentences and perfect the translations. Due to the structural and lexico-semantic differences between Taruma, Wapishana and English, some English translations did not contain all the information present in the Taruma and Wapishana sentences. We also recorded greetings and other conversational expressions in Taruma and elicited some grammar, for instance by means of the

---

[8] For a more detailed discussion of the remote data collection methodology, especially with regard to ethnobotanical research, see Holt (2021).

stimuli from Bowerman et al.'s (2004) put project.[9] The speakers also granted us access to the recordings they had made with Dr Meira in 2015. Further, we recorded sessions in which the Taruma translated Ms Suttie's personal history, the oral histories from the 2015 data, and another oral history that Ms Suttie told spontaneously. These translation sessions were a useful source of additional data as they shed light on additional vocabulary, grammatical structures and sound symbolism. The speakers were compensated for their time with 1000 Guyanese dollars per hour, which is equivalent to a teacher's salary in the region.

2.2 Data management and archiving

At the end of the fieldwork, the speakers were given copies of the audio and video recordings in digital form, and they will receive the updated ELAN files when the transcription and analysis are completed. Additionally, there are two types of output for the community. The first is a booklet containing the collected pictures, names, and uses of the plants in Taruma, Wapishana and English, which could find use at the local school, for instance. The second is a series of videos that Dr Rybka created by adding the plant pictures to the audio recordings in ELAN (Figure 4). As a result, the videos show the plants, while the audio plays and the transcriptions appear on the video comparably to subtitles.
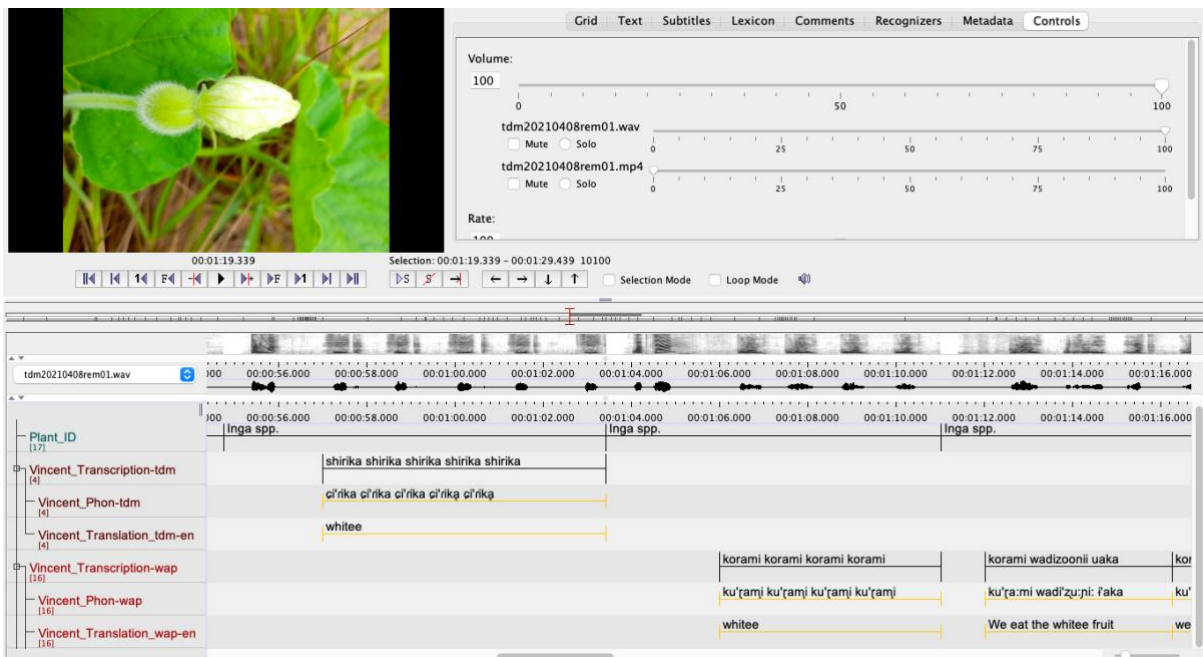
Figure 4. Screenshot of a recording transcribed in ELAN.

---

[9] The stimuli consist of short videos of a person performing different actions of putting an object somewhere or taking it from somewhere; these are designed to ascertain if a given language has lexical or morpho-syntactic elements such as verbs, prepositions or morphology, to distinguish between different types of actions involving objects and locations.

In the interest of open science, the files will be archived with the Endangered Languages Archive (ELAR). Most of the data will be accessible to registered users of the archive.[10] The files will include both subsets of the data, as the speakers and Dr Meira gave their permission to archive the earlier recordings along with those from the ethnobotanical project. The sessions in both subparts of the deposit are organised by date and then by genre: remote recording (rem), narrative (nar), songs (song), elicitation (eli), or translation (transl). The files from 2015 have been renamed to conform to the file naming conventions adopted for the project, which consists of the ISO-code for Taruma, the date of recording, the genre and a number representing the order in which recordings of that genre were made on that date (for instance, tdm20211002nar01). The file names of the plant pictures consist of the date on which they were taken, a number and the depicted plant part; the numbers correspond to the order in which the plants were photographed, named and described on a given day (e.g., 20210410_01_leaf.jpeg). This is also the order in which the plants appear in the ELAN files, in which there is a separate tier containing the file names of the corresponding photographs. The transcriptions in some of the ELAN files are incomplete at the time of writing but will be updated after the completion of this thesis.

3 TARUMA PHONOLOGY

This section is a description of Taruma phonology, beginning with the inventory of the phonemes and their allophones (§3.1). Next, I describe the phonological processes in Taruma, which include both language-internal processes and processes resulting from contact with Wapishana (§3.2). After analysing the phonotactics and prosodic features in Taruma (§3.3, §3.4), I examine the characteristics of sound-symbolic words (§3.5). Thereafter, I describe the types of adaptation processes at play in loanwords (§3.6) and show two morphophonological processes (§3.7). In §3.8, I provide background information about developing an orthography, before comparing the contemporary data to the transcriptions in Schomburgk's and Farabee's historical wordlists (§3.9).

3.1 The Taruma phoneme inventory

The Taruma consonant inventory consists of 17 phonemes (table 3): two nasal stops, seven plain oral plosives, four fricatives, a tap, and a set of labialised consonants including two plosives and a fricative. In contrast to previous reports, according to my analysis, the inventory does not contain implosive consonants. Unusual features of the Taruma consonant inventory include the absence of a lateral and the presence of a set of labialised consonants spanning the

---

[10] In ELAR, there are three levels determining who can access archived files (Endangered Language Archive 2022). Thus, the speakers can decide to make their data openly available to either everyone browsing ELAR (O), registered users of the archive (U) or registered users who have additionally requested and obtained permission from the speakers (S). The majority of the files will be archived as U, with only certain files containing personal information marked as S.

categories of both plosives and fricatives. While voicing is the contrastive feature in the pairs of bilabial, alveolar, and velar plosives, there is some asymmetry with regard to voicing in fricatives. The voiced retroflex fricative does not have a voiceless counterpart, and none of the voiceless fricatives have a voiced counterpart. The Taruma vowel inventory contains eight phonemic vowels, including four nasal vowels (table 4). In the analysis presented here, the approximants [j] and [w] are not phonemes but rather allophones of /i/ and /o/, which combine with the other vowels to form nine different diphthongs as well as one triphthong (see §3.1.2.3).

Table 3. Taruma consonant inventory.

	Bilabial		**Alveolar**		**Retroflex**	**Palatal**	**Velar**		**Glottal**
Nasal		m		n					
Plosive	p	b	t	d			k	g	ʔ
Labialised plosive							kʷ	gʷ	
Fricative			s		ʐ	ɕ			h
Labialised fricative	ɸʷ								
Tap				ɾ					

Table 4. Taruma vowel phoneme inventory.

	Front	**Central**	**Back**
Close	i	ɨ	
	ĩ	ɨ̃	
Close-mid			o
			õ
Open		a	
		ã	

The following table gives an overview of the different allophones and the graphemes used to represent them in this thesis. For a more detailed discussion of the orthography, see §3.8.

Table 5. Taruma phonemes and allophones with corresponding graphemes.

Phoneme	Allophone	Grapheme	Example (phonetic)	Example (orthographic)	Meaning
/a/	[a, ɛ]	\<a\>	[ˈa.tɕi]	\<achi\>	face, eye
/ã/	[ã, ãN]	\<ã\>	[ka.ˈɾã̃ŋ.gʷa]	\<karãgwa\>	otter
/b/	[b]	\<b\>	[aˈba.ɾa]	\<abara\>	back
	[w, β]	\<w\>	[ˈwo.βo]	\<wowo\>	termite
/d/	[d]	\<d\>	[ˈda.nɨ]	\<danu\>	jaguar
	[dʑ]	\<dj\>	[dʑa.ˈʔa.ra]	\<dja'ara\>	pet

/ɸʷ/	[ɸʷ]	<fw>	[ɸʷa]	<fwa>	fire
/g/	[g, gʲ]	<g>	[go.'mi.tɕi]	<gomichi>	woman
/gʷ/	[gʷ]	<gw>	['gʷi]	<gwi>	dig
/h/	[h]	<h>	['ha.ma]	<hama>	I, my
/i/	[i, e]	<i>	[bi.'kʲe.ɾu]	<bikiro>	balata tree
	[j]	<y>	['jo.ɾo]	<yoro>	slow
/ĩ/	[ĩ, ĩN]	<ĩ>	['hĩn.dʑa]	<'hĩdja>	no
/k/	[k, kʰ, ʰk, kʲ]	<k>	['ko.ba]	<koba>	arrow
/kʷ/	[kʷ]	<kw>	['kʷa.sɨ]	<kwasu>	person
/m/	[m]	<m>	[ma.'ɾa.gʷi]	<maragwi>	stingray, firefan
/n/	[n, nʲ]	<n>	[na.'ka.nʲi]	<nakani>	know
/o/	[o, u]	<o>	[o.ɾo.'ko.da]	<orokoda>	knee
	[w]	<w>	['wa.ɾa]	<wara>	put
/õ/	[õ, õN]	<õ>	['a.ɸõ]	<afõ>	hand
/p/	[p, pʰ, pʲ]	<p>	['pa.na]	<pana>	dawn
	[ɸ]	<f>	['ɸo.na]	<fona>	wasp
/ɾ/	[ɾ]	<r>	[ɾi.'kʲi.nʲiu]	<rikinyo>	cricket
/s/	[s]	<s>	['sɨ.wi]	<suwi>	throw
/ɕ/	[ɕ]	<sh>	[ɕo.'no.ɾa]	<shonora>	crane species
/t/	[t, tʰ]	<t>	['ta.kʲi]	<taki>	ask
	[tɕ, ʰtɕ]	<ch>	[tɕi.'na.ɾi]	<chinari>	shin
/ɨ/	[ɨ]	<u>	[a'dɨku]	<aduko>	hair
	[ɛ]	<e>	['hɛ.ɾa]	<hera>	arrive
/ɨ̃/	[ɨ̃]	<ũ>	[nɨ̃]	<nũ>	bad, ugly
	[ɛ̃]	<ẽ>	['ɛ̃.hɛ̃]	<ẽhẽ>	flesh
/z̺/	[z̺]	<z>	['ho.z̺a]	<hoza>	rain
	[dz̺]	<dz>	['dz̺ɨ.dz̺ɨ]	<dzudzu>	language
/ʔ/	[ʔ]	<'>	['ɾa.ʔwa]	<ra'wa>	give birth

3.1.1 Consonants

The following set of minimal pairs and near-minimal pairs illustrate the contrasts between the Taruma consonants (table 6). The remainder of this section is a description of the phonemes and their allophones. Since Taruma does not permit consonant clusters, consonants do not influence each other: there are no assimilation or dissimilation processes triggered by

consonantal contexts. Instead, allophonic variation is conditioned by the neighbouring vowels. For this reason, when discussing consonantal phonemes and their allophones, I give tables listing examples of the consonant phonemes surrounded by different oral vowels. Nasal vowels do not occur frequently enough in the data to establish if they trigger the same phonological processes in consonants as oral vowels and are therefore not included in the tables. All consonants appear in both word-initial and word-medial onsets but not in the coda except for /n/, which is the only phonemic consonant found in codas.

Table 6. Minimal pairs showcasing the contrasts between Taruma consonant phonemes.

Contrast	Phonetic	Phonemic	Meaning
/m/ - /n/	['ma.kʲi]	/'ma.ki/	3.POSS
	['na.kʲi]	/'na.ki/	trumpeter bird
/p/ - /b/	[a.'pa.ɾa]	/a.'pa.ɾa/	bury
	[a.'ba.ɾa]	/a.'ba.ɾa/	back
/m/ - /p/	['ɖʐɨ.mi]	/'zɨ.mi/	disappear
	['ɖʐɨ.pi]	/'zɨ.pi/	baby sling
/m/ - /b/	['ma.kʲi]	/'ma.ki/	3.POSS
	['ba.kʲi]	/'ba.ki/	tapir
/p/ - /ɸʷ/	['a.pʲe]	/'a.pi/	stone
	['a.ɸʷi]	/'a.ɸʷi/	walk
/p/ - /h/	['pa.na]	/'pa.na/	daybreak
	['ha.na]	/'ha.na/	fly
/t/ - /d/	[tɨ]	/tɨ/	agouti species
	[dɨ]	/dɨ/	here
/t/ - /n/	['ta.kʲi]	/'ta.ki/	ask
	['na.kʲi]	/'na.ki/	trumpeter bird
/d/ - /n/	[dɨ]	/dɨ/	here
	[nɨ̃]	/nɨ̃/	sad
/ɾ/ - /t/	['ka.ɾo]	/'ka.ɾo/	make
	['ka.to]	/'ka.to/	soul
/ɾ/ - /d/	['ɾu.mi]	/'ɾo.mi/	spider monkey species
	['do.me]	/'do.mi/	strong
/k/ - /g/	['ko]	/'ko/	eat
	['go]	/'go/	neck
/k/ - /kʷ/	['ka.sɨ]	/'ka.sɨ/	ring-tailed coati
	['kʷa.sɨ]	/'kʷa.sɨ/	person
/g/ - /gʷ/	['gi]	/'gi/	say
	['gʷi]	/'gʷi/	dig
/kʷ/ - /gʷ/	[kʷa]	/kʷa/	NEG.
	['gʷa]	/'gʷa/	ant
/ʔ/ - /h/	['ba.ʔe]	/'ba.ʔi/	collared peccary
	['ba.hɨ̃]	/'ba.hɨ̃/	hit
/ɸʷ/ - /h/	['hi]	/'hi/	bush dog
	['ɸʷi]	/'ɸʷi/	dead
/s/ - /ʐ/	['hi.su]	/'hi.su/	stomach
	['hi.ʐu]	/'hi.ʐu/	white-lipped peccary

3.1.1.1 Nasal stops

The bilabial nasal /m/ is realised as [m] in every environment (table 7). Similarly, the alveolar nasal /n/ is realised as [n] in all environments, except when preceding the high front vowel /i/, or its allophone [e], in which case it is obligatorily realized as a palatal nasal [nʲ] (table 8).

Table 7. The phoneme /m/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	[ma.'ɾa.gʷi]	firefan	['ha.ma]	I
i	[mi'.gʷi]	put	[dʑi.'mi.kɨ]	tasty
ɨ	[mɨ.'hʷã]	a while ago	[dʑɨ.'mɨ.sɨ]	forget
o	['mu]	howler monkey	['na.do naj.ko.'no.mo]	pepper pot

Table 8. The phoneme /n/ in different vocalic environments.

vowel	word-initial onset		word-medial onset		coda	
	phonetic	meaning	phonetic	meaning	phonetic	meaning
a	[na.'ka.nʲi]	know	['a.na]	you (plural)	[wu.wu.'wan]	waterfall
i	[nʲi.tu]	bitter cassava	[gʷi.'nʲi]	to bathe		
ɨ	[nĩ]	sad	[ɾi.ku.'nɨ.ku]	porridge		
o	[nu.'ku.da]	forest	[ɕo.'no.ɾa]	crane	[pɛ.ɾi.'ton]	plant sp.

The obligatory allophonic palatalisation of /n/ is conditioned only by the vowel following the nasal, either /i/ or its allophone /e/, but not by the vowel preceding it. This can be seen, firstly, in the word *chamani* [tɕa.'ma.ɲi] 'bacaba palm' (*Oenocarpus* species), in which the palatalisation is triggered by /i/ following /n/, irrespective of the vowel before /n/. Secondly, in *gigina* ['gi.gi.na] 'wash' and *adjino* [a.'dʑi.no] 'backyard', /n/ is not palatalised as it precedes low vowel /a/ in the former, and back vowel /o/ in the latter case, while the high front vowel before the nasal in either example does not result in palatalisation. This process also applies to the voiceless bilabial plosive /p/ (§3.1.1.2.1), the alveolar plosives (§3.1.1.2.2), and the voiceless velar plosive (§3.1.1.2.3).

The phoneme /n/ also appears in the coda, albeit quite rarely. In our data, there are only three examples, none of which are typical lexical items of indisputable Taruma origin. The first is *pishan* [pi.'ɕan] 'cat', most likely a loan from Wapishana *pishan*, a language that has nasals in the coda position. The second, *wowoan* [wu.wu.'wan] 'waterfall', is onomatopoeic. The third example is *kaon* ['ka.on], a word with an unclear meaning. In one, /n/ occurs in the coda more as an exception than as the rule; further, no other consonantal phonemes have been found in the coda position. However, bilabial, alveolar, and velar nasals appear in the coda position as homorganic allophones of nasal vowels preceding bilabial, alveolar, and velar stops, respectively (see §3.1.2).

20

3.1.1.2 Oral stops

3.1.1.2.1 Bilabial stops

Taruma has both a voiceless and a voiced bilabial stop. The voiceless bilabial stop /p/ is realised as [p] in most environments but it is subject to three phonological processes (table 9). Firstly, in all environments before /i/ except in unstressed word-initial onsets, /p/ is obligatorily palatalised and realised as [pʲ]: compare *djipi* [ˈdʑi.pʲi] 'name' with *piroka* [pi.ˈɾu.ka] 'banana'. There is currently no example of a stressed word-initial syllable formed of /p/ and /i/, but the palatalisation of /p/ is similar to that of /k/ insofar that /p/ does not palatalise in word-initial unstressed /pi/ sequences. Therefore, it can be assumed that /p/, like /k/, also palatalises in stressed word-initial /pi/ sequences. Secondly, before /o/, /p/ is occasionally realised as [p], but more often as either the voiceless bilabial fricative [ɸ] or the voiceless glottal fricative [h], such as in *fofone* [ɸo.ˈɸo.nɛ, ho.ˈɸo.nɛ] 'gras'. In word-initial onsets before /o/, there is some variation between [ɸ] and [h], whereas in word-medial onsets before /o/, /p/ usually surfaces as [ɸ]. Thirdly, most realisations of /p/ are unaspirated, but there are also occasional examples of [pʰ], for instance, in *pana* [ˈpʰa.na] 'daybreak'. However, this word is pronounced without aspiration in other instances; therefore, aspiration is not phonemically contrastive. Instead, [pʰ] occurs largely in free variation with [p], often in stressed syllables.

Table 9. The phoneme /p/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	[pa.ˈˣkʷa.ta]	medicine	[a.ˈpa.ɾa]	bury
i	[pi.ˈɾu.ka]	banana	[ˈdʑi.pʲi]	name
ɨ	[pɨ.ɾa.ta]	money	[ˌma.pɨ.ˈma.pɨ]	Wapishana
o	[po.ˈɾo.ɾi, ho.ˈɾo.ɾi]	millipede	[ki.ˈɾa.pu,	thorn
	[ɸo.ˈɸo.nɛ, ho.ˈɸo.nɛ]	gras	ki.ˈɾa.ɸu]	

The voiced bilabial /b/ stop is realised as [b] in most environments, but similarly to /p/, /b/ often undergoes lenition before /o/ (table 10). With the exception of *shaboro* [ɕa.ˈbo.ɾo], a type of fermented drink made from cassava, in which /bo/ is always realised as [bo], /b/ preceding /o/ in stressed syllables is realised either as a labial-velar approximant [w] or as a voiced bilabial fricative [β]. This can be seen in the example of *wowo* [ˈwo.βo, ˈβo.βo] 'termite'. In unstressed syllables, there is some variation between her [b] and [β] when preceding /o/, as /ˈma.bo.ma/ 'I return' is realized as realised as either [ˈma.bo.ma] or [ˈma.βo.ma]. Furthermore, there is variation between the speakers. The word *wabo* /ˈwa.bo/ 'manicole palm' is realised with a fricative, as [ˈwa.βu], by Ms Suttie. In contrast, Mr Louis consistently pronounces it with a stop. Mr Louis' realisation of /b/ as [b] in this word might be Wapishana influence: Taruma *wabo* is borrowed from Wapishana *wabo*, itself a borrowing from Cariban languages. In Wapishana, this word is pronounced with the bilabial implosive /ɓ/, which does not undergo lenition. Replacing it with a plain bilabial plosive in Taruma is thus presumably a strategy to adapt this loanword to Taruma phonology.

Table 10. The phoneme /b/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	[ba.ˈhõ]	snake species	[dʑi.ˈba.ba]	early
i	[bi.ˈnʲi.nʲi]	crab species	[kʲi.ˈbi.kɨ]	glue
ɨ	[ˈbɨ.ʔa]	run	[ˈdʑa.bɨ]	older sister
o	[ˈwo.βo, ˈβo.βo]	termite species	[ɕa.ˈbo.ɾo]	cashiri drink
			[ˈwa.bu, ˈwa.βu]	manicole palm

3.1.1.2.2 Alveolar stops

The alveolar voiceless stop /t/ is realised as [t] except before /i/, where it obligatorily palatalises and surfaces as the voiceless alveolo-palatal affricate [tɕ] (table 11). In this analysis, [tɕ] is treated as an allophone of /t/ here because [t] and [tɕ] appear in complementary distribution. The allophone [t] does not occur before /i/ in the data other than in *kamoti* [ka.ˈmo.ti] 'clay pot' and [a.ˈti:] 'until', both of which are listed in the Wapishana dictionary compiled by the Wapishana Language Project and Wapichan Wadauniinao Ati'o, henceforth WLP and WWA (2000: 137, 16). Furthermore, there are no examples of [tɕ] before /ɨ/ other than in *chupon* [tɕɨ.ˈpun] 'spoon', which is also borrowed from Wapishana *chupoon* (WLP and WWA 2000: 23). When analysing [tɕ], diachronically, as a combination of /t/ and /i/, this pattern can be explained based on the constraints in diphthong formation. As /iɨ/ diphthongs are not found in the data, this suggests that there are no underlying /tiɨ/ sequences which would surface as [tɕɨ].

Further, similarly to /p/, /t/ is sometimes realised with optional aspiration; this does not constitute a phonemic contrast with the unaspirated stop, but rather free variation. The aspirated [tʰ] appears in stressed syllables, for instance, in *tara* [ˈtʰa.ɾa] 'bowl' and *tania* [ˈtʰa.nʲia] 'forehead', when words are repeated several times during wordlist elicitation. However, the same words are realised without aspiration in longer utterances in more naturalistic discourse. The aspiration during the elicitation sessions could therefore be due to emphasis or fatigue.

Table 11. The phoneme /t/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	[ˈta.ɾa]	bowl	[pa.ˈˣkʷa.ta]	medicine
i	[tɕi.ˈna.ɾi]	shin	[ˈhi.tɕi]	gray brocket deer
ɨ	[tɨ.ˈna.bu]	belt	[ga.ˈtɨ.ka]	always
o	[to.ˈbe.to]	mushroom	[ko.ˈɾo.to]	hole

The voiced alveolar stop /d/ is realised as [d] except when preceding /i/, in which case it palatalises and is obligatorily realised as the voiced alveolo-palatal affricate [dʑ] (table 12). The argument for considering [dʑ] an allophone of /d/ is identical to the reasoning concerning its voiceless counterpart: [d] and [dʑ] appear in complementary distribution before /i/ and /ɨ/,

which can be explained by analysing [dʑ] as /di/. The absence of /iɨ/ diphthongs in Taruma means that there are no underlying /diɨ/ sequences which would be realised as [dʑɨ].

Table 12. The phoneme /d/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	[da.ʔa.ˈkʷi]	come	[ˈa.da]	head
i	[ˈdʑi.ba]	afternoon	[wa.ˈʔa.dʑi]	talk
ɨ	[dɨ]	here	[ˈba.dɨ]	axe
o	[do.ˈna.ha]	far	[ˈdu.du]	earth

3.1.1.2.2 Plain velar and glottal stops

The voiceless velar plosive /k/ is realised as [k] in all environments except when preceding /i/, where it is palatalised as [kʲ] (table 13).

Table 13. The phoneme /k/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	[ˈka.ɾo]	make	[a.ˈka.ɾa]	chicken
i	[ˈkʲi.na]	piranha	[ˈwi.ᶜkʲi]	opossum
ɨ	[kɨ.ˈɾɨ.ɾi]	rub	[he.ˈbɨ.kɨ]	salty
o	[ko]	eat	[no.ˈˣko.da]	forest

Similarly to the other consonants that are palatalised in this environment, only the following vowel conditions this allophonic variation, as the example of *gika* [ˈgi.ka] 'man' shows. Similarly to /p/, /k/ palatalises obligatorily except in unstressed syllables in word-initial onsets: compare *kina* [ˈkʲi.na] 'piranha' and *kirafo* [ki.ˈɾa.ɸu] 'thorn'. As with [pʰ] and [tʰ], there are occasional non-phonemic instances of aspiration, such as in *karo* [ˈkʰa.ro] 'make' and *kato* [ˈkʰa.tu] 'soul', which are found in stressed syllables when words are repeated during elicitation recordings but not in longer utterances. Thus, similarly to to the other voiceless stops, [kʰ] occurs in free variation with [k]. Additionally, like [tɕ], /k/ can be optionally pre-aspirated, such as in *wiki* [ˈwi.ᶜkʲi] 'opossum' or in *nokoda* [no.ˈˣko.da] 'forest'. The process is discussed further in §3.2.1.

The voiced velar stop /g/ is usually realised as [g] (table 14). In two examples, *gibi* [ˈgi.be, ˈgʲe.be] 'hit' and *gima* [ˈgi.ma, ˈgʲe.ma] 'I say', /g/ is optionally palatalised and realised as [gʲ]. The phoneme does not occur frequently overall and is not found before /ɨ/. Nevertheless, phonemic status can be assumed for /g/ due to the existence of the minimal pairs shown in table 6. Furthermore, /g/ does not alternate with any other sounds and appears in core vocabulary items that are unlikely to be borrowed, such as *gika* [ˈgi.ka] 'man' and *gomichi* [go.ˈmi.tɕi] 'woman'.

Table 14. The phoneme /g/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	['ga.na]	thing	['ho.ga]	bacaba palm
i	['gi.ka]	man	['gi.gi.na]	wash
ɨ				
o	[go]	neck	[ha.'go.ri]	corpse

Intervocalically, the voiceless glottal stop /ʔ/ is realised as [ʔ] (Table 15). In this environment, it is contrastive, as the example of *noa* /'no.a/ 'electric eel' and *no'a* /'no.ʔa/ 'long ago' shows. Phonetically, glottal stops are also often found word-initially in words which begin with a vowel. However, there is no evidence as of yet that this constitutes a phonemic contrast, as there are no minimal pairs distinguished by the presence of an initial glottal stop. Therefore, initial glottal stops are presumably inserted to mark the word boundaries of words which are underlyingly vowel-initial, such as in *hama ada* ['ha.ma 'ʔa.da] 'my head'.

Table 15. The phoneme /ʔ/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a			[da.'ʔa.kʷi]	come
i			[wa.'ʔi.ɾi]	gourd
ɨ			[ˌno.'ʔo.kʷa.'ko.ro]	story
o			['(d)ʑi.ku ʑɨ.'ʔɨɾi]	toucan

3.1.1.2.2 Labialised velar stops

The voiceless labialised velar stop /kʷ/ is realised as [kʷ] in all environments (table 16). Similarly to /k/, /kʷ/ also optionally occurs with pre-aspiration, for instance in *akwa* ['a.ˣkʷa] 'arm' (see §3.2.1). It is noticeable that there are no word-initial instances of /kʷo/ and only one example of word-medial /kʷo/. This suggests that the sound developed from /ko/ sequences and became reanalysed as a phoneme. However, its phonemic status, as opposed to an underlying /koV/ diphthong, can be assumed for three reasons. Firstly, the existence of *djikwo* ['dʑi.kʷo] 'boil', which would have /koo/ as its underlying structure. This would be highly unusual, as Taruma does not have many words with a length distinction. Secondly, there are no /oɨ/ diphthongs in the data except in the context of what is here analysed as phonemic labialised stops, /kʷɨ/ and /gʷɨ/ (as well as one instance of /pʷɨ/, which is discussed further below). They could derive from a structure consisting of two syllables, /ko.ɨ/ and /go.ɨ/ which were collapsed into one when /koV/ and /goV/ became reanalysed as phonemes. Otherwise, if the labialised consonants were not phonemes, then these sequences would underlyingly be /koɨ/ and /goɨ/, and there would be no explanation as to why the two velar stops are the only sounds that precede this diphthong. Furthermore, if /kʷV/, /gʷV/ and /ɸʷV/ were plain consonants preceding rising diphthongs whose first segment is /o/, then we could also expect to see other

consonants, such as /t/, /d/, /s/, or /ɕ/, preceding such diphthongs. This distributional difference suggests that the labialised consonants are the result of a change that did not affect other sounds, such as the contraction of two syllables into one in instances where the second syllable was vowel-initial.

Table 16. The phoneme /kʷ/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	[kʷa.ˣkʷa.nʲi]	paddle	[ˈha.kʷa]	1SG.OBJ
i	[kʷi.ˈɾi.kɨ]	medicine-man	[ˈtɕi.kʷi]	fill
ɨ	[kʷɨ]	here	[ˈdɨ.kʷɨ]	black paint
o			[ˈdʑi.kʷo]	boil

The phoneme /gʷ/ is always realized as [gʷ] and appears in the data with low frequency, similarly to /g/. The lack of [gʷo] sequences suggests that similarly to [kʷ], [gʷ] developed from /go/ before being reanalysed as a phoneme. The arguments for the phonemic status of /gʷ/ are similar to those concerning its voiceless counterpart: it occurs before /ɨ/, while /oɨ/ diphthongs are not found otherwise, and other consonants should also be found in /CoV/ sequences if that was the underlying structure of [gʷV].

Table 17. The phoneme /gʷ/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	[gʷa.sɨ.ˈki]	right away	[ka.ˈɾã.gʷa]	otter
i	[gʷi.ˈnʲi]	bathe	[ma.ˈɾa.gʷi]	stingray, firefan
ɨ	[ˈgʷɨ.na]	swallow	[dʑa.ˈgʷɨ.na]	visit
o				

A possible diagnostic to determine the underlying structure of [kʷ], [gʷ], and [ɸʷ] could be stress reassignment. As is laid out in section 3.4.1, the first-person suffix *-ma* causes stress reassignment to the penultimate syllable in disyllabic words in which the second syllable is heavier. Applied to the words *akwa* [ˈa.kʷa] 'arm' and *bafwa* [ˈba.ɸʷa], this means that if the stress is reassigned to remain penultimate ([a.ˈkʷa.ma, ba.ˈɸʷa.ma]), this would be an indication that the syllable [kʷa] is heavier and therefore consists of two vowels, /koa/. However, if the first syllable bears the stress ([ˈa.kʷa.ma, ˈba.ɸʷa.ma]), that would suggest that both syllables are light and that [kʷa] is underlyingly /kʷa/. This question will remain unanswered for the time being, since neither word has appeared in this inflected form.

3.1.1.3 Fricatives

3.1.1.3.1 Sibilant fricatives

The alveolar fricative /s/ is realised as [s] in all vocalic environments (Table 18). Notably, there are only three words containing a sequence of [s] before [i]: *siriri* [si.ˈɾi.ɾi] 'bat', *sidja* [ˈsi.dʑa]

'cotton', and *asidjari* [a.si'.dʑa.ɾi], the name of a lizard species. Other than in these exceptions, the only sibilant that does occur before /i/ is the voiceless alveolo-palatal fricative /ɕ/. Conversely, there are numerous instances of [sɨ], but only one word containing [ɕɨ], *ushukuwi* [ɨ.ɕɨ.'kɨ.wi] 'touch'. Given the palatalisation of several other consonants before /i/ and the similarity of this distribution to that of /t/ and /d/, I assume that /ɕ/ was initially a palatal allophone of /s/ in the past. However, the number of exceptions to the largely complimentary distribution of [s] and [ɕ] when preceding /i/ and /ɨ/ suggests that, at a later point /ɕ/ was reanalysed as a separate phoneme and is treated as such here.

Table 18. The phoneme /s/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	['sa.ɾi]	tomorrow	[a.'sa.dʑi]	basket type
i	['si.'ɾi.ɾi]	bat	[a.si'.dʑa.ɾi]	lizard species
ɨ	[sɨ'katɕu]	night	[dʐɨ.'mɨ.sɨ]	forget
o	['so.do]	strap	['su:.su]	flower

The voiceless alveolo-palatal fricative /ɕ/ is realised as [ɕ] in all environments (table 19). The overall distribution of this phoneme before different vowels provides additional support for the hypothesis that it developed from an allophone of /s/. Out of twenty-three words containing this sound, in addition to the above-mentioned instance of [ɕɨ], there are only four instances of [ɕa] and six of [ɕu]. The other twelve examples of [ɕ] precede /i/, so that [ɕi] sequences are by far the most common. Furthermore, the dearth of [ɕɨ] sequences follows from /ɕ/ having previously been an allophone of /s/: sequences of /ɕ/ and /a/ or /o/ were initially underlyingly /sia/ or /sio/, but since /i/ and /ɨ/ do not form a diphthong /iɨ/ (see §3.1.2.3), there were no /siɨ/ sequences that could have been reanalysed as /ɕɨ/ except for [ɨ.ɕɨ.'kɨ.wi].

Table 19. The phoneme /ɕ/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	[ɕa.'bo.ɾo]	drink type	[ba.'ɕa.da]	drink type
i	[ɕi.'ɾi.ka]	inga species	[a.'ɕi.ɕi]	dirt
ɨ			[ɨ.ɕɨ.'kɨ.wi]	touch
o	[ɕo.'no.ɾa]	crane species	[a.ˌba.ɕo.'ro]	then

The Taruma phoneme inventory contains a third sibilant, the voiced retroflex fricative /ʐ/, which is often realised as a voiced retroflex affricate [dʐ] in word-initial position (table 20). The affricate is analysed as an allophone of the fricative because its distribution is more limited, both with regard to its position within words and the vowels it precedes. The affricate ['dʐ] is found almost exclusively in word-initial onsets, with the exception of *dzudzu* ['dʐɨ.ʐɨ]

'language', in which it also appears word-medially as [ˈdʐɨ.dʐɨ].[11] The fricative also appears word-initially, albeit in only one example *zu'i* [ˈʐɨ.ʔi], the name of a tree whose flammable resin is used to start the fire (*Protium sp.*). Moreover, [dʐ] occurs only before /ɨ/, with the exception of *dza* [dʐa̰] 'water' and words evidently derived from it, such as *dzakuri* [dʐa̰.ˈkɨ.ɾi] 'thirsty'. In contrast, the fricative appears before /ɨ/, /a/, and /o/. Since /ʐ/ is found in more contexts, it is more likely to be the underlying form. Further observations about the phoneme /ʐ/ include its absence before /i/ and that in word-medial onsets, the retroflex fricative occasionally alternates with the voiced alveolar fricative, such as in *hoza* [ˈho.ʐa̰] 'rain' and *hizo* [ˈhi.ʐu] 'white-lipped peccary', which are also pronounced [ˈho.za] and [ˈhe.zo], respectively.

Table 20. The phoneme /ʐ/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	[ˈdʐa̰]	water	[ˈho.ʐa̰]	rain
i				
ɨ	[ˈdʐɨ.ɾi]	bird	[ˈdʐɨ.ʐɨ, ˈdʐɨ.dʐɨ]	language
o			[ha.ʐo.ˈma.ɾa]	aimara fish

The retroflex phoneme is quite rare in the Taruma data: there are six instances of the fricative and seven examples of the affricate. This suggests that this sound was not originally part of the Taruma phoneme inventory, as it would otherwise be found more frequently. Nevertheless, it is analysed as a phoneme here instead of as an allophone of another phoneme, as it does not appear synchronically in complementary distribution with any other phonemes. This leaves two possible reasons for the presence of the retroflex fricative in Taruma. Either the words containing these sounds are loanwords, or the phoneme developed diachronically. Crosslinguistically, /ʐ/ is a rare phoneme in South America: SAPhon, the South American Phonological Inventories database, lists only seven languages in whose inventories it is present (Michael et al. 2015). Among them is Wapishana, but none of the Taruma words containing the voiced retroflex fricative or affricate are of Wapishana origin. As far as I was able to establish, none of the other languages are possible donor languages either. Furthermore, water is such an essential resource and an omnipresent element of the landscape that it is (perhaps) a universal part of a language's core vocabulary and therefore it does not seem likely that this term was borrowed.

With regard to a potential diachronic development, the strongest correlation can be drawn with the alveolar tap. There are almost no instances of [ɾ] preceding the high central vowel, and [ɾ] is rarely found in word-initial onsets. In contrast, as pointed out above, the retroflex is most frequently found before the high central vowel, so that there is a "partial" complimentary distribution. However, there are some exceptions to this pattern. As shown above, both the retroflex fricative and affricate also appear before vowels other than /ɨ/, albeit less frequently. Furthermore, sequences of [ɾ] following [ɨ] appear in four words in the data. In one example,

---

[11] This exception is presumably a side effect of the sound symbolic nature of [ˈdʐɨ.dʐɨ], which is formed through the reduplication of the first syllable. As Hinton et al. (2010: 9) note, sound symbolic can allow for sounds or structures that are not permitted otherwise in the phonology of a given language.

rukonuko [ɾɨ.ku.'nɨ.ku] 'porridge', the syllable is even word-initial. A diachronic relationship between the two sounds is also not unheard of: according to Rodrigues and Cabral (2012: 509), the retroflex fricative in Awetí has been proposed as the reflex of a palatalised alveolar flap *ɾʲ in Proto-Tupían. In Taruma, the tap does not palatalise and there are many [ɾi] sequences in the data, so it does not appear that the retroflex fricative was initially an allophone of the alveolar tap when preceding /i/. Instead, however, it might have once been an allophone of [ɾ] before /ɨ/, which later became reanalysed as a separate phoneme. This would explain the patterning of the two phonemes /ɾ/ and /ʐ/.

3.1.1.3.2 Non-sibilant fricatives

The voiceless glottal fricative /h/ is realised as [h] in all environments (table 21). In the data, there are no examples of [h] preceding the high central vowel /ɨ/ word-medially and only one word-initial instance, *hukabato* [hɨ.ka.'ba.tu] 'true'. In high-frequency words such as *hama* /'ha.ma/ 'I' and *hĩdja* /hĩ'dʑa/ 'not', word-initial [h] is sometimes dropped.

Table 21. The phoneme /h/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	['ha.ma]	I	[do.'naha]	far
i	['hi]	dog	[gʷi.'hi.so]	roast
ɨ	[hɨka'batu]	true		
o	[ho.'dʑa.sɨ]	Taruma	[no.'ho.be]	teach

The voiceless labialised bilabial fricative /ɸʷ/ is realised variably as [ɸʷ] or as the voiceless labialised glottal fricative [hʷ] (table 22). There are only two words in which [hʷ] does not seem to alternate with [ɸʷ]: the homonyms *hwira* ['hʷi.ɾa, 'hʷe.ɾa] 'star' and *hwira* ['hʷi.ɾa] 'sloth'. In the data, there are no examples of /ɸʷ/ preceding /ɨ/ or /o/. Similarly to the origin of the other labialised consonants, the lack of [ɸʷo] is an indication that /ɸʷ/ developed from a combination of a consonant and /o/, before being reanalysed as a phoneme. The consonant in question is presumably /p/, which is often realised as [ɸ] or [h] before /o/. But since its distribution is more limited, the phonemic status of /ɸʷ/ is somewhat less certain compared to the other labialised consonants. However, the word *afwya* ['a.ɸʷja] 'path' is an indication that the labialised bilabial fricative is indeed a phoneme. If the labial element in *afwya* ['a.ɸʷja] 'path' was /o/ instead of a secondary articulatory feature of the consonant, the underlying form would be /'a.ɸoia/. Triphthongs are possible in Taruma, but the only case in the data is *dja'way* ['dʑa.ʔwaj] /'dia.ʔoai/ 'mother', in which the first and third vowel resemble approximants. In /'a.ɸoia/, in contrast, it would be the first and second vowels which are of an approximant-like quality. While it is not impossible that Taruma has both types of triphthongs, by parsimony, it seems more likely that there is only one kind. A final argument for the phonemic status of /ɸʷ/ is that while the allophone [ɸ] still alternates with [p] in many words, there are no instances of [pʷ] alternating with [ɸʷ]. There is only *pwuna* ['pʷɨ.na] 'full', which partly opposes the predictions made here for /ɸʷ/. On the one hand, it suggests that /p/ was subject to the same process that led to the development of /kʷ/ and /gʷ/, which conforms to the theory that /poV/

is the origin of /ɸʷV/. On the other hand, based on the analysis given here, the word would be expected to be realised only as ['ɸʷɨ.na], which is not the case. But ['pʷɨ.na] does not alternate with ['ɸʷɨ.na], so this word does not call the phonemic status of /ɸʷ/ into question.

Table 22. The phoneme /ɸʷ/ in different vocalic environments.

vowel	word-initial onset		word-medial onset	
	phonetic	meaning	phonetic	meaning
a	[ɸʷa, hʷa]	fire	['ba.ɸʷa, 'ba.hʷa]	scarlet macaw
i	[ɸʷi]	dead	['a.ɸʷi, 'a.hʷi]	sweet potato
ɨ				
o				

3.1.1.4 Liquids

Taruma has no laterals and only one rhotic consonant phoneme, the alveolar tap /ɾ/ (table 23). The tap is realised as such both in initial and intervocalic onsets and does not change in different environments. However, it is very rarely found in word-initial onsets. Furthermore, as mentioned above, there are only four instances of [ɾ] before [ɨ] in the data, which is a datapoint supporting the hypothesis that /ɾ/ used to have an allophone before [ɨ] which has by now become reanalysed as a separate phoneme, /ʐ/.

Table 23. The phoneme /ɾ/ in different vocalic environments.

	word-initial onset		word-medial onset	
vowel	phonetic	meaning	phonetic	meaning
a	['ɾa.ʔwa]	be born	[ko.'baɾa]	pineapple
i	[ɾi.'kʲi.nʲiu]	cricket	[ɕi.'ɾi.ka]	inga species
ɨ	[ɾɨ.ku.'nɨ.ku]	porridge	[kɨ.'ɾɨ.ɾɨ]	rub
o	[ro.ɾa.'bɛ.to]	family	[ko.'ro.to]	hole

3.1.2 Vowels

The following minimal pairs contrast the different vowel phonemes (Table 24).

Table 24. (Near-)minimal pairs showcasing the contrasts between Taruma vowel phonemes.

Contrast	Phonetic	Phonemic	Meaning
/i/ - /ĩ/	['bi.wa]	/'bi.oa/	moon
	['bĩ.wa]	/'bĩ.oa/	liana species
/a/ - /ã/	['ɸʷa]	/'ɸʷa/	fire
	['ɸʷã]	/'ɸʷã/	big
/i/ - /ɨ/ - /a/	['kʷi]	/kʷi/	poison
	['kʷɨ]	/kʷɨ/	there
	['kʷa]	/kʷa/	NEG
	['kʲi.na]	/'ki.na/	piranha

29

	['ku.na]	/'ko.na/	fruit
/ɨ/ - /o/	['ɨ:]	/'ɨ/	hard, firm, tight
	['o:]	/'o/	tree
/a/ - /o/	['ta.ɾa]	/'ta.ɾa/	bowl
	['to.ɾa]	/'to.ɾa/	chigger

3.1.2.1 Oral vowels

/a/

The low front unrounded vowel /a/ is realised as [a] except sometimes in diphthongs. In falling diphthongs when followed by /i/, [a] alternates with the low-mid front unrounded vowel [ɛ], as can be seen in *chayka* ['tɕa.j.ka, 'tɕɛj.ka] 'bow'.

/i/

The high front unrounded vowel /i/ is realised as [i], [j] or [e]. As the first element of rising diphthongs or as the last segment of falling diphthongs, /i/ is realised as [j], as in *horya* ['ho.ɾja] 'today' and *kway* [kʷaj] 'parch'. The allophones [i] and [e] seem to occur largely in free variation. That said, in unstressed syllables /i/ is more likely to be lowered: compare ['ta.ɾe] and [ta.'ɾi.hĩ], forms of *tari* /'ta.ɾi/ 'bake bread' with and without a suffix that results in stress reassignment from the first to the second syllable of the root. However, the alternation also appears in stressed syllables, such as in *hi* ['hi, 'he] 'dog'.

/ɨ/

The high central unrounded vowel /ɨ/ is realised as [ɨ] or [ɛ], for instance in *badu* ['ba.dɨ, 'ba.dɛ] 'axe' and *kasu* ['ka.sɨ, 'ka.sɛ] 'coati'. The allophone [ɛ] is often found in unstressed syllables, but not always, as the example of *aduko* [a'dɨku, a'dɛko] 'hair' shows. In this analysis, [ɛ] is considered not a phonemic vowel but an allophone of /a/ and /ai/, and mainly of /ɨ/, for the following reasons. For one, there are no minimal pairs contrasting [ɨ] and [ɛ]. Secondly, as an allophone of three different phonemes, it seems unlikely that this sound would also be a phoneme in its own right. Further, notably, [ɛ] and [ɨ] are both absent or mostly absent in certain environments, such as following /ɾ/, /ɕ/, and the labialised consonants, or in diphthongs. And while there is no clear complementary distribution, there are certain tendencies to their variation: in alveolar contexts, the allophones alternate in the same words, whereas after consonants, it is usually either one allophone or the other. For instance, the high central vowel [ɨ] is only found once after /h/ and twice after /p/, whereas there are several examples of [hɛ] and [pɛ]. In contrast, [ɛ] only appears once after /k/ and rarely after /s/, while [ɨ] is found after either onset in many examples. Finally, neither [ɨ] nor [ɛ] occur as frequently overall as /a/, /i/, and /o/, neither sound is found word-initially except in *u* [ɨ:] 'hard, firm', and there are also only few occurrences of the corresponding nasal vowels. Since the two sounds have so many distributional characteristics in common, this analysis postulates that they are realisations of

the same phoneme, /ɨ/, rather than two separate phonemes which occur infrequently and in similarly restricted contexts. The vowel /ɨ/ is considered to be the underlying phoneme due to the word *u*. Since this word is stressed, the vowel is unlikely to be reduced, and it consists of one vowel only, which minimizes the influence of other phonemes.

/o/

The close-mid back rounded vowel /o/ is realised as [u], [o], [ɔ], and [w]. This variation can partly be explained by a tendency towards assimilation based on height. Following a syllable that contains /i/, /o/ is often realised as [u], as in *akioko* [a.ki.'juku] 'brother-in-law'. After other vowels, /o/ is often realised as [o], as in *karo* ['ka.ɾo] 'make'. However, there is some variation between the two speakers: *dodo* /'do.do/ 'earth', a word in which there are no other vowels to harmonise with, is realised as ['du.du] by Mr Louis and as ['do.do] by Ms Suttie. Since Ms Suttie is the more proficient speaker, this suggests that /o/ is the underlying phoneme. Mr Louis' pronunciation might be influenced by Wapishana, in which [u] and [o] are realizations of /u/. In word-final unstressed syllables, /o/ is also sometimes lowered to [ɔ], for instance in *djakwado* [dʑa.'kʷa.dɔ] 'father'. In diphthongs, /o/ is often realised as [w], such as in *wara* ['wa.ɾa] /'oa.ɾa/ 'put' and *asuwba* [a.'sɨw.ba] 'lip' (see 3.1.2.3).

3.1.2.2 Nasal vowels

Taruma nasal vowels have phonemic status and are not nasalised due to nasal consonants in their environment. This can be seen in (1-4a), where nasal vowels appear without nasal consonants. However, before plosives, the nasal vowels are realised as a combination of a nasal vowel and a nasal stop which homorganically assimilates to the place of articulation of the oral stops (1-3b).

(1) /ĩ/   a. ['kʷĩ.ja]   /'kʷĩ.a/   'juice'
         b. [hĩn.'dʑa]   /hĩ.'dʑa/   'not'
(2) /ã/   a. [ɸʷã]   /ɸʷã/   'big'
         b. [hĩãm.bi'dʑi] /hĩã.bi.'di/   'hang'
(3) /ũ/   a. ['aɸõ]   /'aɸõ/   'hand'
         b. ['sũŋ.ku]   /'sõ.ko/   'navel'
(4) /ɨ̃/   a. [sɨ̃]   /sɨ̃/   'penis'

It is possible, however, that there is an optional phonological nasalisation process in addition to the phonemic nasal vowels in Taruma. This process does not seem to occur in polysyllabic words, but might apply in monosyllabic words beginning with a nasal: the word *mo* 'howler monkey' is realised as either ['mu] or ['mõ]. The monosyllable *nũ* [nɨ̃] 'sad' also consists of a nasal stop in the onset and a nasal or nasalized vowel, but there are no examples of this word being pronounced with an oral vowel instead. At this point, the data does not feature enough monosyllabic words containing a nasal stop to determine with certainty if Taruma has such a nasalisation process.

3.1.2.3 Diphthongs

In the analysis presented here, Taruma does not have phonemic approximants, even though phonetically, the inventory contains both a labiovelar approximant [w] and a palatal approximant [j]. The approximants are instead analysed here as allophones of /o/ and /i/, respectively, in diphthongs, or rather vowel sequences, due to their respective distribution before different vowels and the fact that they can also be syllabic.

Firstly, there are no minimal pairs that would confirm the labiovelar approximant as a phoneme separate from /b/. Secondly, in terms of distribution, the approximants are found before fewer vowels than what would be expected if they were phonemes (table 25). Specifically, the approximant [j] is not found before /ɨ/ or /i/ and the approximant [w] only appears once before /ɨ/, in *pwuna* ['pwɨ.na] 'full', and neither sound precedes / ɛ/.

Table 25. The distribution of [j] and [w] before different vowels.

vowel	[j]		[w]	
a	['do.ja]	house	['wa.ɾa]	to put
i			['wi.ˤkʲi]	opossum
ɨ			['pwɨ.na]	full
o	['tɕa.ju]	farm	['wo.ka, βo.ka]	blood

Moreover, similarly to [j], [w] does not appear before its corresponding vowel: as argued above, [wo] is underlyingly /bo/. This is because [wo] mainly occurs word-initially and alternate with [βo], similarly to word-medial [bo] alternating with [βo]; this realisation of /bo/ is analogous to /p/ undergoing lenition before /o/.[12] In one, there is a distributional gap: the approximants are not found before all oral vowels. A possible explanation of this distribution is that the approximants are underlyingly vowels themselves. Languages do not usually have all the diphthongs that are theoretically possible by combining all existing vowels in their inventory in every possible constellation. Therefore, if the approximants are vowels, it is to be expected that some vowel sequences, such as for instance /iɨ/, do not occur.

The second reason for analysing the approximants as vowels is their syllabicity. Compare the bare root *aria* ['a.ɾja] 'fish' (5a) and the inflected form *arika* [a.'ɾe.ka] (5b). The segment realised as [j] in the second syllable of the bare form could be interpreted as part of a complex onset. However, that same segment is realised as [e] and constitutes the nucleus when the root is followed by the non-subject marker *-ka*, which suggests that it is a vowel.

---

[12] There are two instances of [β] that do not fit this pattern, as they occur before a vowel other than /u/: ['tɕi.βi] 'ité palm' and [wi.'βi.ku] 'gourd'. In both cases, [β] is both preceded and followed by /i/; however, none of the other instances of [ibi] or [iwi] are pronounced with a fricative, so that the environment does not seem to be reason for this realisation. An explanation could be that ['tɕi.βi] is likely a loanword from Bahuana *itiwi* (Bredero 2021); [wi.'βi.ku] may well also be a borrowing.

(5) a. ['a.ɾ**j**a] b. [a.'ɾ**e**.ka]
       aria   *ari-ka*
       fish   fish-NS
       'fish'   'fish'

Similarly, the word 'mother' appears in four forms: ['dʑa.ʔo], ['dʑa.ʔ**w**a], ['dʑa.ʔ**w**aj], and ['dʑa.ʔ**w**ɛ]. It is not clear at this point what the difference between the forms is. Nevertheless, this example shows that the segment that is realized as [w] in ['dʑa.ʔ**w**a], ['dʑa.ʔ**w**aj], and ['dʑa.ʔ**w**ɛ] is the vowel [o] in ['dʑa.ʔ**o**]. This suggests that [w] in this word is underlyingly the vowel /o/ and that it constitutes the first element of the rising diphthong /oa/ instead of being part of the onset.

In contrast to the first vocalic element in rising diphthongs, it is somewhat more difficult to prove that [j] and [w] as the second segment in falling diphthongs such as *aycho* ['aj.tɕu] 'face, eye' are vowels as well. It would be an indication that they are consonants if they were resyllabified when vowel-initial suffixes are added. However, in words such as *chayka* ['tɕaj.ka, 'tɕɛj.ka] 'bow' and the derived verb *chayakahĩ* [ˌtɕɛ.ja.ka.'hĩ] 'shoot', it is difficult to determine if the underlying form is /ˌtia.ja.ka.'hĩ/, with a phonemic approximant, or /ˌtiai.a.ka.'hĩ/, with [j] resulting from the transition from /i/ to /a/. However, as shown above, rising diphthongs formed of /iV/ or /oV/ exist in Taruma. It seems unlikely that they would contrast phonemically with syllables consisting of /jV/ and /wV/. Further, the sequence /io/ in *yoma* /'io.ma/ 'soft' can be realised as either a rising or a falling diphthong: this word appears in the data both as ['ew.ma] and as ['jo.ma], which is a further argument for considering both vocalic segments to be vowels. Therefore, [j] and [w] in syllable-final position are analysed as /i/ and /o/, similarly to their syllable-initial counterparts.

To sum up, [j] and [w] are not phonemes because their distribution is limited and because they alternate with /i/ and /o/. As a result, in words such as *adjoa* [a.'dʑo.wa] 'tail' and *biado* [bi.'ja.do] 'savannah', in which an approximant follows directly after its corresponding vowel, this is a result of the transition between the vowels; the underlying forms are /a.'dio.a/ and /bi.'a.do/. All other instances of [j] and [w], except in [wu] sequences, are analysed as being underlyingly /i/ and /o/, respectively (see also 3.1.2.3).

The following seven rising diphthongs are attested in the data, two of which feature a nasal vowel as the initial segment. Where possible, they are given in (6-12) without a consonantal onset word-initially (a) and word-medially (b) as well as with an onset (c):

(6) /ia/ a. [ja.'kʲi.kɨ]   /ia.'ki.kɨ/   'brown'
          b. ['do.ja]   /'do.ia/   'house'
          c. ['ta.nʲia]   /'ta.nia/   'forehead'
(7) /io/ a. ['jo.ɾo]   /'io.ɾo/   'slow'
          b. ['tɕa.ju]   /'tia.io/   'farm'
          c. [ɾi.'kʲi.nʲiu]   /ɾi.'ki.nio/   'cricket'
(8) /oa/ a. ['wa.sɨ]   /'oa.sɨ/   'open'
          b. ['bi.wa]   /'bi.oa/   'moon'
          c. ['ɾa.ʔwa]   /'ɾa.ʔoa/   'be born
(9) /oi/ a. ['a.tɕi 'we]   /'a.ti 'oi/   'eyebrow'
          b. ['sɨ.wi]   /'sɨ.oi/   'throw (away)'

33

(10) /oɨ/ a. ['pwɨ.na]   /'poɨ.na/   'full'
(11) /õa/ a. ['õã]   /'õã/   'sun, day'
(12) /ĩa/ a. ['hĩã]   /'hĩã/   'good'

There are also two falling diphthongs:

(13) /ai/ a. ['aj.tɕʊ̈]   /'ai.tio/   'bone'
           b. [kʷaj]   /kʷai/   'parch'
(14) /ɨo/ a. [a.'sɨw.ba]   /a.'sɨo.ba/   'lip'

The following observations can be made with regard to the diphthongs. Firstly, in the data, the diphthongs involving nasal vowels or /ɨ/ are rare. Secondly, the vowel sequence /ao/ is also attested only once, in *kwawpidap* [kʷaw.pʷi.'dʲap] 'fishing rod'.[13] This word might be an ad-hoc coinage, as 'fishing rod' is otherwise given as *kwakipa* [kʷa.'kʲi.pa, kʷa.'kʲe.pa] in the data. Finally, /ai/ sequences are occasionally realised as [ɛ], as in (15). As a consequence of analysing all vocalic elements within a syllable as vowels, examples such as in (15a) are considered to be triphthongs.

(15) a. ['dʑa.ʔwaj, 'dʑa.ʔwɛ]   /'dia.ʔoai/   'mother'
     b. [gʷaj, gʷɛ]   /gʷai/   'fetch'

3.2 Phonological processes

This section describes the phonological processes found in Taruma. First, the processes are shown which are considered to be Taruma in origin, namely palatalisation, lenition and pre-aspiration (3.2.1). This is followed by a discussion of the processes that are ascribed to language contact, namely the speakers' bilingualism in Taruma and Wapishana. These processes are labialisation after /o/ and final devoicing (3.2.2).

3.2.1 Language-internal processes: palatalisation, lenition, and pre-aspiration

Firstly, as has been shown in §3.1.1, most Taruma stops palatalise when preceding /i/:

(16) a. /ni/   ['ta.nʲia]   /ta.nia/   'forehead'
     b. /pi/   ['dʑi.pʲi]   /'di.pi/   'name'
     c. /ti/   ['a.tɕi]   /'a.ti/   'face, eye'
     d. /di/   ['dʑa.bɨ]   /'dia.bɨ/   'older sister'
     e. /ki/   ['ma.ko.kʲi]   /ma.ko.'ki/   'go'
     f. /gi/   ['gʲe.ma]   /gi.ma/   'I say'

This process applies to all plosives except the voiced bilabial stop /b/, the glottal stop and the labialised velar stops. However, while the alveolar stops always palatalise, /p/ and /k/ do not

---

[13] This word most likely ends in a devoiced vowel, e.g. /ḁ/.

palatalise in unstressed initial syllables, and the two words in which [gʲ] is found are also sometimes realised without palatalisation.

Secondly, the bilabial oral plosives usually undergo lenition before /o/. For the voiceless bilabial plosive, this takes the form of spirantization or debuccalisation, as it is realised as either [ɸ] or [h] (17). For the voiced bilabial plosive, lenition takes the form of spirantisation to [β] in word-medial onsets and of approximation to [w] in word-initial onsets (18).

(17) /po/ a. ['ɸo]   /'po/   'saliva'
          b. [ka.dja.'ɸo]   /ka.dia.'po/   'people'
(18) /bo/ a. ['wo.da]   /'bo.da/   'rope'
          b. [wa.βu]   /oa.bo/   'açai'

A third phonological process that takes place in Taruma is the pre-aspiration of voiceless consonants, mainly /k/. The realisation of the pre-aspiration is conditioned by the preceding vowel, similarly to what Borgstrøm (1940, cited in Silverman 2002: 579) describes for some Scottish Gaelic dialects. Thus, /ʰk/ is variously realised as [ˁk] or [ˣk]:

(19) a. [i.'ˁkʲia]   /i.kia/   '3'
      b. [a.'sɨ.ˣku]   /a.'sɨ.ko/   'bark'
      c. ['a.ˣka]   /'a.ka/   'macaw'
      d. [ma.ˣku.'ˣkʲi]   /ma.ko.'ki/   'go'

After the high front vowel /i/, the pre-aspiration is palatal, while it is velar following vowels that are either not high or not front. While pre-aspiration most commonly occurs with /k/, it is also found with [tɕ], for instance in *awicha* [a.'wi.ˁtɕa] 'belly', and with /kʷ/, such as in *akwacho* [a.'ˣkʷa.tɕu] 'creek'. As the following examples show, pre-aspiration not only occurs within morphemes, but also at the morpheme boundary, for instance between words and the non-subject-marking suffix *-ka* (20.a) and across word boundaries (20.b):

(20) a. [mu.ˣku.ɾu.'ˣka]   /mo.ko.ɾo.'ka/   'mukru plant'
      b. ['ha.ma ˣkʷa.'kʲe.pa] /'ha.ma kʷa.'ki.pa/ 'my fishhook'

As is typical for pre-aspiration cross-linguistically, this phenomenon is not phonemic (Ladefoged and Maddieson 1996 73); in Taruma, it might serve to help disambiguate the plain unvoiced consonants from their voiced counterparts. Furthermore, it does not occur every time a given word is pronounced.

3.2.2 Transfer from Wapishana: occasional labialisation after /u/ and final devoicing

Two phonological processes in Taruma can be traced back to the speakers' proficiency in and daily use of Wapishana. Both processes are attributed to Wapishana influence because they do not appear consistently throughout the data and often only occur in the wordlists, in the first instance but not subsequent repetitions of a given word.

Firstly, velar stops in Wapishana are labialised after /u/ and /ũ/:

(21) /ũgaɾɨ/   [ũ'gʷaɾɨ]   'I/me/mine'   (dos Santos 2006: 42)

(22) /uka:wan/ [uk^wa:wan] 'she arrived'

Similarly, labialization of velars after /o/ can occasionally be found in Taruma utterances, as in *wokawa* [wo.'k^wa.wa] 'raw' or *wa'ew karudja* [wa.'ʔɛw k^wa.'rɨ.dʑa] 'make food'. However, in neither example, labialisation occurs consistently. The labialised velar stop in *wokawa* only appears once in the elicitation session, with the word subsequently being pronounced as [wo.'ka.wa]. Furthermore, the phrase *wa'ew karudja* appears frequently in the data and is pronounced without labialization in all other instances, as [wa.'ʔɛw ka.'rɨ.dʑa]. Finally, as the following examples show, labialization is not usually at work in Taruma:

(23) a. [a.ko.'ka] 'you'
     b. [pi.'ru.ka] 'banana'
     c. ['ɸo.ga]   'bacaba palm'

Here, the velar stops in (23) would be expected to labialize if labialization was a productive process in Taruma. Thus, both of these examples illustrate that the labialization of velars after /o/ in Taruma is irregular and presumably caused by the speakers' bilingualism.

    The second phonological process in Wapishana is the apocope or devoicing of final vowels and devoicing of the preceding consonant:

(24) a. ['watо̥]   'crow'   (Gomes 2022: 152)
     b. [iɾip]   /iɾib/ 'much, many'   (dos Santos 2006: 64)[14]

This devoicing also sometimes appears in the Taruma data. (25a) is an example of the final devoicing of the vowel, whereas (25b) shows the devoicing of both the final vowel and consonant:

(25) a. ['ma.k^wḁ]   /'ma.k^wa/ 'tinamou'
     b. ['wo.d̥o̥]   /'bo.do/   'urine'

The word *wodo* is alternatively pronounced ['wo.do], with voiced segments in the final syllable. Furthermore, instances of devoicing are not frequent overall, especially within sentences. Thus, this process does not regularly form part of Taruma phonology but sometimes carries over from Wapishana.

3.3 Syllable structure

This subsection describes Taruma syllable structure, vowel length, syllable weight, and the phonotactic constraints that operate on the level of words.

---

[14] In the Wapishana dictionary, this word is written as <iriba>, which suggests that the underlying form is /iɾiba/ instead of /iɾib/ (WLP and WWA 2000: 66). However, since dos Santos' (2006) description is of Wapishana as spoken in Brazil, whereas the dictionary was compiled by Wapishana speakers living in Guyana, this might be due to differences between the two varieties of the language.

The syllables found in the data are of the type V, VV, Vn, CV, CVV, CVn and CVVV (Table 26). Syllables with the structure V, VV, CV and CVV are canonical, while Vn, CVn and CVVV are each represented by one example only.

Table 26. Types of Taruma syllables.

Syllable structure	Example (phonemic)	Meaning
V	/**i**.ˈkia/	3
VV	/**io**.ˈma.kɨ/	soft
Vn	/bo.bo.ˈ**an**/	waterfall
CV	/ˈ**ka**.ɾo/	make
CVV	/ˈta.**nia**/	forehead
CVn	/pɛ.ɾi.ˈ**ton**/	plant species
CVVV	/dia.ˈ**ʔoai**/	mother

Other than the few instances of /n/ in the coda, syllables are open. Onsets are optional and may consist of no more than one consonant. Nuclei, in their turn, may contain between one and three vowels. The minimal nucleus, which consists of a single vowel as in *ikya* /i.ˈkia/ '3', is also the minimal syllable. However, when a monosyllabic word, which is always stressed, is composed only of a vowel, the vowel is lengthened, as in *o* [o:] 'tree' and *u* [ɨ:] 'hard, firm'. The length of these vowels does not appear to be lexical; when the non-subject suffix *-ka* is attached to *o* to form *oka* [ˈo.ka], /o/ is similarly in length to other stressed vowels and significantly shorter than in the bare morpheme. Unlike *o*, *u* has not been attested yet with additional morphemes, so that it cannot be determined with certainty, but at this point it seems that [o:] and [ɨ:] are lengthened are lengthened as a means of emphasis, to make the stress apparent even in a monosyllable.

Vowel length is not contrastive in Taruma, as the example of [o:] and [ˈo.ka] shows. Furthermore, there are no minimal pairs distinguished by length. Therefore, vowel length is not lexical, maybe with the exception of *djiidji* [ˈdʑi:dʑi] 'grandfather', *sooso* [ˈsu:.su] 'flower' and *djiicho* [ˈdʑi:tɕu] 'camp', in which the vowel in the first syllable is always long. In addition to these words and the monosyllabic words without onset, word-final vowels that are not long in other instances are sometimes lengthened for pragmatic reasons. In these cases, the speaker usually hesitates briefly after the long vowel while thinking about how to continue the utterance. Prolonging the vowel before a brief pause presumably serves to signal that the utterance is not yet completed, which suggests that the long vowels in these examples do not result from the underlying phonological structure.

In terms of their weight, Taruma syllables can be monomoraic (26a), bimoraic (26b-d) or trimoraic (26e-f):

(26) a. CV   b. CVV   c. CṼ   d. CVn

σ   σ   σ   σ
| / \   / \   / \
μ   μ   μ   μ   μ   μ   μ
/\   /\ | /\ | /\ |
m a   k i a   h o +nasal t o n

[ma]   [kʲia]   [hõ]   [ton]

e. CVVV   f. CṼV

σ   σ
/ | \   / | \
μ μ μ   μ μ μ
/\ | | /\ | |
ʔ o a i   h i a +nasal
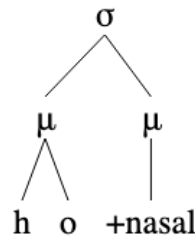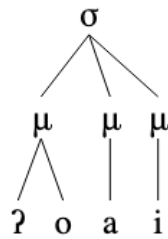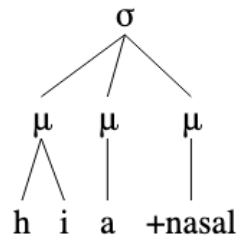
[ʔwai]   [hĩã]

Monomoraic syllables of the type (C)V are light. Bimoraic and therefore heavy types of syllable structure include (C)VV, (C)Ṽ, and (C)Vn. (C)Ṽ syllables are analysed as bimoraic since they behave like (C)VV syllables in terms of stress or stress reassignment (see 3.4).[15] Trimoraic, super-heavy syllables contain three vocalic segments or a sequence of nasal vowels.

Syllables are combined to form words with two restrictions. Firstly, syllables ending in /n/ are always word-final. Secondly, syllables consisting exclusively of vowels cannot be combined infinitely. The longest words without consonants have a V.VV structure, such as *iwa* /'i.oa/ 'smoke from fire'. The highest number of adjacent vowels attested is four, in a CVV.VV structure such as in *chayo* /'tia.io/ 'farm'. There are no particular types of syllable structure that serve to distinguish different word classes.

The canonical structure of Taruma words is di- and trisyllabic; monosyllabic and monomorphemic words with more than three syllables are less frequently found. Words with more than three syllables tend to be complex forms, often as a result of reduplication in the case of nouns, as in in *shirishiri* [ɕi.ɾi.'ɕi.ɾi] 'intestines' and several animal names (see §3.5).

---

[15] This is similar to phonemically nasal vowels in Romance languages such as French and Portuguese, which are longer than the corresponding oral vowels (Delattre and Monnot, 1981; Beddor, 1993; cited in Rothe-Neves and Valentim 2012: 110). It has been suggested that the properties of Portuguese nasal vowels are a result of vowels becoming associated with the time unit and the [+nasal] feature left by deleted nasals (Moraes and Wetzel 1992). Therefore, it is possible that the nasal vowels in Taruma also arose from Vn sequences. In the case of nasal diphthongs such as [hĩã], the [+nasal] trace presumably spread to both vowels.

Verbs with more than three syllables usually contain derivational suffixes, such as *mizokahĩ* [me.z̢o.ka.'hĩ] 'play', which consists of the root *mizo* 'toy', non-subject marker *-ka* and the suffix *-hĩ*. The longest words that appear to be monomorphemic consist of five syllables: the number four, *dja'obarada* [dʑa.ʔo.'ba.ɾa.da], and *kiaramapa* [ki.a.ɾa.'ma.pa] 'Waiwai'. In Farabee's (1918) wordlist, several numbers have up to ten syllables. However, these are also complex forms, as they can be parsed to some extent. For instance, <ociwai> 'one', <jowa> 'two', and <mĩkyahahi> 'three' are combined with <-yakana'ahami> to form <ociwaiyakana'ahami> 'six', <jowakyakana'ahami> 'seven', and <mĩkyahahiyakana'ahami 'eight'>.

3.4 Prosody

This subsection describes the suprasegmental features of stress and intonation in Taruma. Phonetically, stress is often realised as higher pitch and increased volume relative to the unstressed syllables. Furthermore, the vowels in unstressed syllables can be reduced in length. Stress and intonation interact with each other to some extent, since pitch is a component of both prosodic features.

3.4.1 Stress

Stress in Taruma is not lexical, but usually fixed on the penultimate syllable (27a-d). Words ending in *-n* have final stress (27e-f). Monosyllabic words are always stressed (27g).

(27) a. ['**sɨ**.wi]   'throw'
     b. [ma.'**ɾa**.gʷi]   'stingray, firefan'
     c. [ka.ɾi.'**ka**.ɾi]   'parrot species'
     d. [ki.a.ɾa.'ma.pa]   'Waiwai'
     e. [pi.'**ɕan**]   'cat'
     f. [wu.wu.'**wan**]   'waterfall'
     g. ['**gi**]   'say'

Exceptions to this pattern include words with ante-penultimate stress, namely *kiribi* ['**kʲe**.ɾe.be] 'mix' and *dja'obarada* [dʑa.ʔo.'**ba**.ɾa.da] 'four', as well as words with word-final stress, namely *akoka* [a.ko.'ka] 'you', *ikya* [i.'**kʲia**] '3', *makoki* [ma.ko.'**kʲi**] 'go' and any verb ending the suffix *-hĩ*, such as *chokohĩ* [tɕo.ko.'**hĩ**] 'tell'.[16] It is not apparent at this point what causes these exceptions. Syllable weight, which plays a role in stress reassignment with the suffix *-ma* (see below), does not account for syllable-final stress, since the stressed syllables in [ma.ko.'kʲi], ['**kʲe**.ɾe.be] and [dʑa.ʔo.'**ba**.ɾa.da] are not heavier than the penultimate syllables which normally bear stress. While [hĩ] and the final syllable in [i.'kʲia] are indeed heavy, this does not provide an explanation since there are counter-examples of words in which the final

---

[16] *-hĩ* is a verbal suffix whose function is not yet clear.

syllable is the heaviest which follow the canonical penultimate stress pattern, such as *ahõ* ['**a**.hõ] 'hand'.

When suffixes attach to nouns, stress is often reassigned so that it remains penultimate. This is exemplified by the suffixes *-kwa* 'locative' and *-ki* 'instrumental':

(28) *-kwa*   a. ['**tɕa**.jo]   b. [tɕa.'**jo**.kʷa]
                   chayo   *chayo-kwa*
                   farm   farm-LOC
                   'farm'   'at/to/from the farm'

              c. [no.'**ko**.da]   d. [no.ko.'**da**.kʷa]
                   nokoda   *nokoda-kwa*
                   forest   forest-LOC
                   'forest'   'at/to/from the forest / from the forest'

(29) *-ki*   a. ['**ma**.rja]   b. [ma.'**ri**.kʲe]
                   maria   *mari-ki*
                   knife   knife-INS
                   'knife'   'with a knife'

With *-kwa*, stress is reassigned and remains on the penultimate syllable in disyllabic and trisyllabic roots. With *-ki*, there are no examples of trisyllabic roots, but the suffix similarly effects stress reassignment in disyllabic roots. However, with other suffixes, there is some variation depending on the number of syllables in the root and the weight of the final syllable of the root. Compare the stress reassignment (or lack thereof) in *djipi* ['**dʑi**.pʲi] 'name' and *djakwado* [dʑa.'**kwa**.do] 'father' with the suffixes *-ni* '2SG.POSS' and *-ma* '1SG.POSS':

(30) *-ni*   a. [dʑi.'**pʲi**.nʲi]   b. [dʑa.'**kʷa**.do.nʲi]
                   djipi-ni   *djakwado-ni*
                   name-2SG.POSS   father-2SG.POSS
                   'your name'   'your father'

(31) *-ma*   a. ['**dʑi**.pʲi.ma]   b. [dʑa.'**kʷa**.do.ma]
                   djipi-ma   *djakwadoma*
                   name-1SG.POSS   father-1SG.POSS
                   'my name'   'my father'

While the stress is reassigned within the disyllabic root in [dʑi.'**pʲi**.nʲi], this is not the case with the trisyllabic root in [dʑa.'**kwa**.do.nʲi]. When followed by *-ma*, the stress remains unchanged in both roots. The only examples of stress reassignment in the presence of *-ma* are disyllabic roots in which the second syllable is heavier than the first.

(32) a. [a.'**hõ**.ma]   /a.hõ.ma/   'my hand'
      b. ['**a**.da.ma]   /a.da.ma/   'my head'

In *ahõ* ['a.hõ] 'hand' (33a), the nucleus of the second syllable consists of the nasal vowel /õ/, which is bimoraic (see 3.3) and therefore heavier than the oral vowel in the light first syllable /a/. In contrast, in disyllabic roots consisting of two equally light syllables, such as *ada* ['**a**.da]

'head', stress is not reassigned but remains on the same syllable in *adama* [**'a**.da.ma] 'my head' (32b). If both syllables are heavy, stress does not shift either:

(33) a. [**'hã**.dʑa.ma]   /'hã.dia.ma/   'my armpit'
      b. [**'aj**.tɕu.ma]   /'ai.tiu.ma/   'my face'
      c. [**'dʑa**.ʔwa.ma]   /'dia.ʔoa.ma/   'my mother'
      d. [dʑa.**'ʔwɛ**.ma]   /dia.'ʔoai.ma/   'my mother'

In both *hãdjama* [**'hã**.dʑa.ma] 'my armpit' (33a) and *aychoma* [**'aj**.tɕu.ma] 'my face' (33b), each of the syllables is bimoraic, as they consist of either a diphthong or a nasal vowel, and the stress does not shift from the penultimate syllable. Compare also *dja'wama* [**'dʑa**.ʔwa.ma] (33c) and *dja'wayma* [dʑa.**'ʔwɛ**.ma] (33d) 'my mother'. In the former, both syllables are bimoraic and the stress remains on the first syllable. In the latter, the first syllable is bimoraic and the second syllable is trimoraic; as a result, the stress shifts to the second syllable, as it is heavier.

With the non-subject suffix *-ka*, finally, the most important factor determining stress reassignment is the number of syllables in the root. In roots consisting of four syllables, such as *karikarika* [ka.ɾi.**'ka**.ɾi.ka] 'parrot species', stress does not shift when the suffix is added, and is therefore antepenultimate. In contrast, in trisyllabic roots, stress does get reassigned, as can be seen in *biadoka* [bi.a.**'do**.ka] 'savannah'. In disyllabic roots, there is some variation; in *gikaka* [gi.**'ka**.ka] 'man', stress assignment takes place, but in *hichika* [**'hi**.tɕi.ka] 'deer' it does not. In contrast to disyllabic roots inflected with *-ma*, syllable weight does not provide an explanation here, as all syllables in both examples are light.

In verbs, stress also fulfils a grammatical function: word-finally, it signals the imperative, both singular (34a) and plural (34b). In the indicative mood, these verbs follow the regular pattern of penultimate stress, as in *bu'a* [**'bɨ**.ʔa] 'run' and *apara* [a.**'pa**.ɾa] 'bury', respectively. and the second person singular agreement suffix *-'a* is unstressed, as in *nakani'a* [na.ka.**'**nʲi.ʔa] 'you know'. In the imperative, however, the stress shifts to the ultimate syllable.

(34) a. [bɨ.ʔa.**'ʔa**]
          bu'a–'a
          run–2SG
          '(You) Run!' (2nd SG)

      b. [a.pa.**'ra**   ɾo.ɾa.**'**be.to   **'**ɸʷi.ʔa]
          apara   rorabeto   fwi–'a
          bury   family   dead–2PL.POSS?
          Bury your family when they die.' (2nd PL)

Therefore, the stressed final syllable distinguishes the imperative mood from the indicative mood. This applies both when the stressed syllable is the person agreement suffix, as in [bɨ.ʔa.**'ʔa**], and when it is the last syllable of the root, as in [a.pa.**'ra**], since the second person plural does not have a verbal agreement suffix.

To sum up, stress is usually penultimate in uninflected roots and is often reassigned to penultimate syllable when nouns are inflected with *-ki*, *-kwa*, and *-ka*. With *-ni*, stress is only

reassigned in disyllabic words, and with *-ma* it is only reassigned in disyllabic words where the second syllable is heavier than the first. In verbs, word-final stress signals the imperative. It should be noted, however, that the suffixes *-ma* and *-ka* appear more frequently in the data than the others, so that the stress patterns described here can be subject to change as more data emerges.

3.4.2 Pitch and intonation

Two intonational patterns are discernible. Compare the pitch contours of an affirmative (35a) and an interrogative (35b) utterance drawn in Praat (Figures 5 and 6).

(35) a. /ˈki.ɾi.bi i.ˈkia.ka/　　　　　b. /ha.bi.a.ˈʔa/
　　　　kiribi　*ikya–ka*　　　　　　*habia–'a*
　　　　mix　　3–NS　　　　　　　　EXIST–2SG
　　' [I] mix it'　　　　　　　　　'Are you there?'
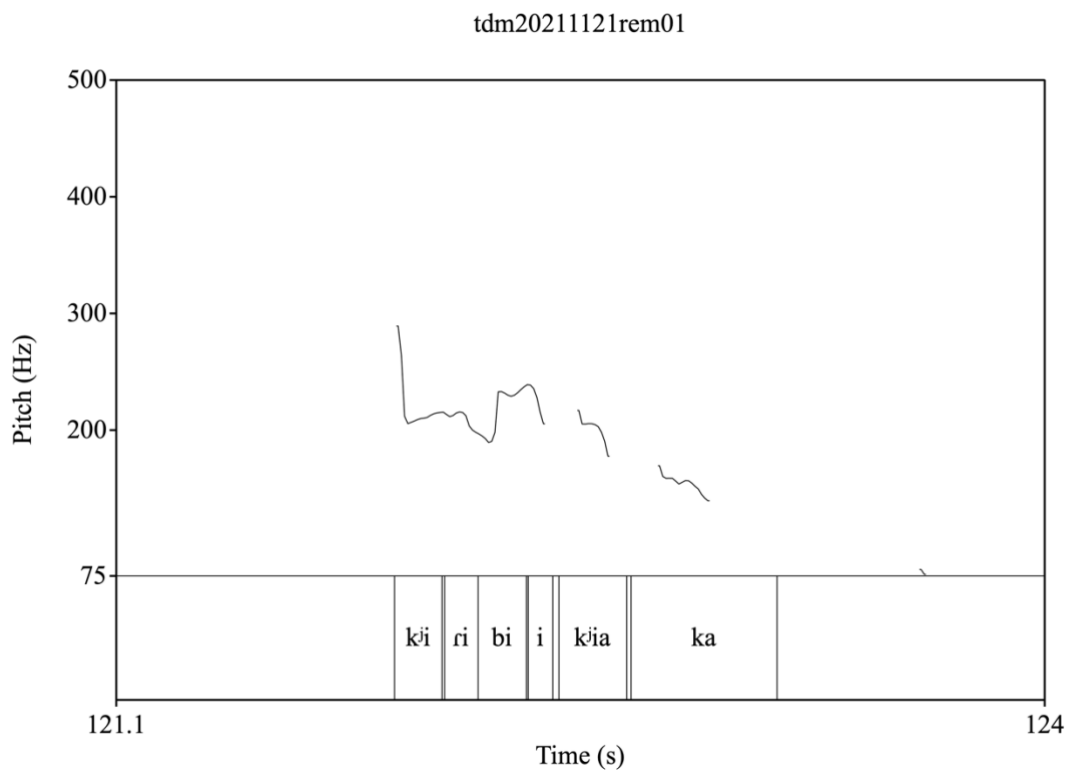
Figure 5. Utterance with affirmative intonation.

42
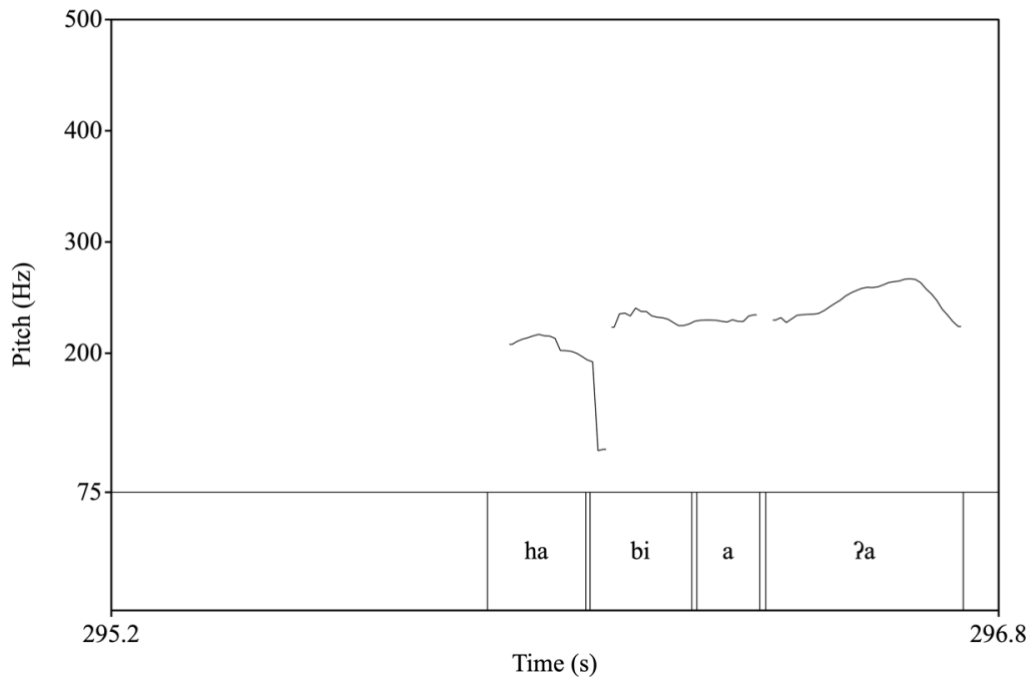
Figure 6. Utterance with interrogative intonation.

In affirmative utterances, the last syllable is characterised by a falling pitch, as in Figure 5. In contrast, in interrogative utterances, the pitch of the final syllable rises (Figure 6). Since higher pitch is one of the phonetic properties of stressed syllables, the interrogative intonation can appear similar to the intonation in imperative forms, which end in a stressed syllable. However, the difference between the two is that the raised pitch of the stressed syllable in imperatives is word-final but not necessarily utterance-final. An imperative is characterized by the same intonational pattern as interrogative utterances only when the command does not contain an object following the verb, so that the last syllable of the imperative is coincidentally also utterance-final.

3.5 Sound symbolism

This subsection summarises the typology of sound symbolism given by Hilton et al. (2010) in order to discuss some sound-symbolic Taruma words and strategies through which they are formed. Sound symbolism is defined as the "direct linkage between sound and meaning" and can be divided into four categories on a gradable scale of gradually increasing arbitrariness (Hinton et al. 2010: 1-2). The first category is corporeal sound symbolism, which serves to externalise a speaker's internal "emotional or physical" state, for instance, by means of interjections or involuntary noises (Hinton et al. 2010: 2). The second category, imitative symbolism, comprises words and utterances representing "environmental sounds", such as *knock*, but also "bird and animal sounds" by means of onomatopoeia (Hinton et al. 2010: 3). Synaesthetic sound symbolism, the third category, refers to the use of certain segments to

43

express non-acoustic characteristics; for instance, high vowels are often found in diminutives or words denoting objects of a small size (Hinton et al. 2010: 5). The fourth type, conventional sound symbolism, contains segments whose associations with certain meanings have become conventionalised (Hinton et al. 2010: 5). In contrast to synaesthetic sound symbolism, in which some cross-linguistic tendencies can be observed, conventional sound symbolism is often language-specific; however, there can be a degree of overlap between conventional sound symbolism and the previous two categories (Hinton et al. 2010: 5).

Hinton et al. (2010: 9) name three strategies that are used to form sound-symbolic words. Firstly, a common strategy is reduplication, especially when other lexical items of a given language also feature reduplication (Hinton et al. 2010: 9). Secondly, the segments used in sound symbolism are usually a restricted subset of the phoneme inventory, or may include sounds that are not part of the phoneme inventory otherwise (Hinton et al. 2010: 9). Thirdly, certain types of segments are often used to represent similar characteristics cross-linguistically: as Hinton et al. (2010: 9) write, "stops are used for abrupt sounds and acts, and continuants for continuing sounds and acts". This strategy is most often found in imitative and synaesthetic sound symbolism.

Most of the sound-symbolic Taruma words described here fit best into the category of imitative sound symbolism. They can be loosely divided into four semantic categories: bird sound, the names of birds and invertebrates, and words representing environmental sounds. With regard to the first category, the expression *brro iki* [ʙoː i.ˈkʲe] is used to portray the sounds of birds such as chickens and parrots in an oral tradition. Here, the voiced bilabial trill /ʙ/ appears, which is not part of the regular Taruma phoneme inventory. The sound-symbolic bird names in the second semantic domain are *aka* [ˈa.ˣka] 'macaw', *fofoki* [ɸo.ˈɸo.kʲi, ho.ˈɸo.kʲi] 'curassow' and *karikari* [ka.ɾi.ˈka.ɾi] 'parrot species'. These names are presumably imitations of the sounds of these birds, with the plosive /k/ imitating the abrupt nature of their cawing. In contrast to the bird sound discussed above, the names are conventionalised as lexical items and do not feature extra-systemic sounds. Furthermore, they are characterized at least partly by reduplication: in the word [ka.ɾi.ˈka.ɾi], the last two syllables are identical to the first two, and the first two syllables of the word [ɸo.ˈɸo.kʲi] are also identical. Strictly speaking, the word [ˈa.ka] does not consist of reduplicated syllables but the vowel /a/ is repeated in both syllables.

Similarly to the bird names, some names of invertebrates (36) imitate the sounds produced by their referents: the bilabial fricative [ɸ] in *fona* [ˈɸo.na] 'wasp' is reminiscent of the buzzing of wasps, the reduplication in *kamakama* [ka.ma.ˈka.ma] 'butterfly' represents the repetitive flapping of a butterfly's wings, and the nasal vowel /ã/ followed by the [jo] sequence in [ko.ˈʔã.jo] 'mosquito' represents the whining sound of mosquitos.

(36) a. [ˈɸo.na]   'wasp'
      b. [ko.ˈʔã.jo]   'mosquito'
      c. [ɾi.mi.ˈɾi.mi]   'horsefly'
      d. [ɕu.ˈɕu.da]   'cockroach'
      e. [ka.ma.ˈka.ma]   'butterfly'
      f. [tɕi.tɕi.ˈha.da]   'grasshopper'
      g. [ɾi.ˈkʲi.nʲiu]   'cricket'
      h. [pi.ˈdʑi.dʑi̥]   'tick'
      g. [ˈβo.βo]   'termite'

The other names are less transparent in terms of mimicking a specific sound, which is perhaps not suprising considering that some of these animals, such as ticks, do not produce sounds audible to humans. However, it is noticeable that reduplication is common in this semantic domain, as most of the words in (36c-g) have at least one reduplicated syllable. There are also several names of invertebrates in the data where this is not the case, but many of these lexical items are characterized by assonance, as they feature identical vowels in adjacent syllables, such as in *pararo* [pa.'ɾa.ɾo] 'earthworm', *korori* [ko.'ɾo.ɾi] 'spider' or *binini* [bi.'nʲi.nʲi] and *dara* ['da.ɾa], two crab species.

The final semantic domain concerns words representing environmental sounds. For instance, in *wowoan* [wu.wu.'wan] 'waterfall', the reduplication of [wu] expresses the steady roaring of a waterfall. Other examples include fwifwi ['ɸʷi.ɸʷi] 'whistling', *fwa* [ɸʷa] 'fire', and *pofo* [po.'ɸo] 'fan'. Similarly to the words discussed above, ['ɸʷi.ɸʷi] and [po.'ɸo] are formed by reduplication, in this case of the first syllable; in [po.'ɸo], the reduplication also represents the repetitive movement of fanning. In addition, as is common crosslinguistically, these words feature continuant phonemes, namely bilabial fricatives, to express continuing actions and sounds, such as the hissing noise of fire, whistled notes or the airstream produced by fanning. But despite its prevalence in these lexical items, [ɸ(ʷ)] does not seem to be an example of conventional sound symbolism in Taruma. It is not associated with a single semantic domain that could be defined easily; while both whistling and fanning are connected to the manipulation of air, this category does not easily fit the word *fire*. Furthermore, [ɸ(ʷ)] does not appear in every word connected to air: the word for *wind*, which could be expected to contain this phoneme, is *kobato* [ko.'ba.to]. Conversely, there are several semantically unrelated words featuring this sound, such as *fwi* [ɸʷi] 'dead' or *afwya* [a.'ɸʷia] 'road'.

To sum up, the sound symbolism contained in the data is of imitative nature; it includes bird sound, the names of birds and invertebrates, both of which often feature reduplication, words with labial phonemes representing environmental sounds. Many of these sound-symbolic words adhere to the crosslinguistic tendency of representing abrupt actions with plosives and continuant actions with continuants. The list of words discussed here is by no means exhaustive. The analysis of a larger set of lexical data and non-lexical utterances may well highlight more examples and patterns of sound symbolism in Taruma, especially sound symbolism of the corporeal and synaesthetic type.

3.6 Loanword phonology

The Taruma vocabulary contains several lexical items that are shared with other Indigenous languages, as well as words borrowed from Indo-European languages (Table 27). The Indo-European loanwords often denote objects or animals that were imported to South America by Europeans, while many of the words shared with other Indigenous languages refer to plants. It should be noted that it is difficult to establish the direction of the borrowings with regard to many of the shared words, especially in the absence of historical reconstruction for Taruma. For instance, similar names for the saki monkey are found in several languages spoken in and around the Guianas: Lokono *bisa*, Wapishana *wisa*, Bahuana *wica* and Manao *huitcha* (Rybka

2020: 249). One of these forms could have been borrowed into Taruma as *hisai*, but it is also possible that the name originated in Taruma and was borrowed by the Arawakan languages (Rybka 2020: 249). However, determining the direction of borrowing goes beyond the scope of this thesis, which focusses on the phonological correspondences between such words.

Table 27. Some shared vocabulary and loanwords in Taruma (plant vocabulary based on Bredero, 2021).

Contact language	Shared word (orthography)	Taruma form (phonetic)	Meaning
Bahuana	*itiwi*	['tɕi.βi]	ité palm
	box	['bak.sa]	box
	light	['laj.te]	light
English	*to put*	['po.ta]	to put
	rice	[ɹaɪs]	rice
	sorrel (via Wp. *sorrel?*)	[sa'ɾiɾi]	sorrel
	spoon (via Wp. *chupoon*)	[tɕu.'pun, tɕɨ.'pon, tɕi.'pon, su.'pun, sɨ.'pun, spun]	spoon
Mawayana	*kalua*	[ka.'ɾu.wa]	bell yam
Portuguese	*arcabuz*	a.ɾa.ka.'ba.sa]	shotgun
	camisa	[ka.'mi.sa]	clothing
	carta (via Wp. *karita?*)	[ka.'ɾi.ta]	paper
	cochino (via Wp. *kooshi?*)	['ku.ɕ(i)]	pig
	cavalo (via Wp. *kawaro?*)	[ka.'wa.ɾo]	horse
	prata 'silver' (via Wp. *puraata*?)	[pɨ.'ɾa.ta]	money
	saco, saca	['sa.kʲi]	bag
	vaca	['pa.ka]	cow
Wapishana	*bai*	[ɓai, bai]	Muscovy duck
	kanawa	[ka.'na.wa]	canoe
	marii	['ma.ɾja]	knife
	o'i	[o.'ʔi]	farine
	pishan	[pi.'ɕan]	cat
	soom (also Ma. *tuma*)	['su.ma]	tobacco
	wabo	['wa.bu]	manicole

Many of the shared words have been adapted based on the phonological constraints of the target language, which often happens when words are borrowed from one language into another (Kang 2011: 2258). These adaptation processes concern 'all facets of phonological structure, reflecting the segmental, phonotactic, suprasegmental, and morphophonological restrictions of the borrowing language' (Kang 2011: 2258). This subsection examines the adaptation processes at play in loanwords in Taruma, discussing segmental and phonotactic adaptation as well as differential importation. Finally, an analysis of the epenthetic vowels is given based on the data that is currently available.

3.6.1 Segmental adaptation

According to Kang (2011: 2265), when a loanword contains a non-native phoneme, this segment is replaced with the closest approximation in the borrowing language.[17] Taruma shares many phonemes with the Arawakan languages it has been in contact with, so that many words, such as *kanawa* [ka.ˈna.wa] 'canoe', can be borrowed without phonological adaptations. Sounds that do undergo adaptation include the implosives found in Wapishana, which are replaced with the corresponding plain voiced stops at least in some of the repetitions of a given word. Wapishana *bai* /ɓai/ 'Muscovy duck', for instance, is is pronounced by Mr Louis as [ɓai] or as [bai].[18] Further, the lateral *l* in Mawayana *kalua* 'bell yam' corresponds to the only Taruma liquid, the alveolar tap /ɾ/, in Ta. *karoa* [ka.ˈɾu.wa]. This mapping of /ɾ/ onto /l/ also applies in loanwords of Indo-European origin, such as in Portuguese *cavalo* 'horse' and English *sorrel*, which are borrowed as *kawaro* [ka.ˈwa.ɾo] and *sariri* [sa.ˈɾiɾi], respectively. However, these words might have been borrowed via Wapishana. In this case, the replacement of the lateral with a rhotic occurred during the borrowing into Wapishana, whose phoneme inventory does not include a lateral either. When borrowed into Taruma, the Wapishana rhotic /ɽ/ was then replaced with /ɾ/.

Other sounds that are replaced in Indo-European loanwords include the voiced labiodental fricative /v/ and the near-open central vowel [ɐ], neither of which are found in Taruma. This can be seen, for instance, in Portuguese *vaca* [ˈva.kɐ] 'cow', which is borrowed as *paka* [ˈpa.ka] (Ashby et al. 2012c). Here, instead of the near-open [ɐ], the word is borrowed with the open front unrounded vowel /a/, and /v/ is replaced with the voiceless bilabial stop /p/. Thus, the labial place of articulation and the obstruent nature of the borrowed segment /v/ were preserved, while the manner of articulation changed from a fricative to a plosive and the sound became voiceless. The reason why /v/ in Ptg. *cavalo* 'horse', Taruma *kawaro* [ka.ˈwa.ɾo] 'horse' is not analogously replaced with /p/ is that Taruma presumably borrowed it via Wapishana when it had already undergone adaptation to Wapishana phonology in the form of replacing /v/ with /w/ and /l/ with /ɽ/. But the form [ˈpa.ka] could also have been borrowed via another Indigenous language, so that more Portuguese loans in Taruma would be needed to ascertain if Taruma /p/ is consistently mapped onto Ptg. /v/. Finally, the voiceless alveolar fricative /s/ in En. *spoon* sometimes surfaces as either [s] or [tɕ]. This loanword does not seem to be a standardised form in Taruma: in the data, it appears as [tɕu.ˈpun], [tɕi.ˈpon], [tɕi.ˈpon] [su.ˈpun], [sɨ.ˈpun] and [spun]. Here, the variation between [s] and [tɕ] depends on whether the word is a borrowing of the English loanword *chupoon* in Wapishana, or is taken directly from English. The forms [su.ˈpun], [sɨ.ˈpun] and [spun] are likely borrowed directly from English, as the initial [s] is preserved, while [tɕu.ˈpun] and [tɕi.ˈpon] are probably adaptations of the Wapishana loanword [tɕi.ˈpon] to Taruma phonology, which does not allow [tɕi].

---

[17] There are several approaches to the factors that determine the similarity of sounds in the language of origin and the borrowing language. For instance, Steriade (2009, quoted in Kang 2011: 2266) suggests that certain features, such as differences in the type of stricture, are more salient than voicing and place of articulation with regard to the perceptual similarity of sounds. See Kang (2011) for a detailed discussion.

[18] The Taruma word for this animal given by Ms Suttie is *pere* [ˈpɛ.ɾɛ], which makes it likely that [ɓai] is an ad-hoc borrowing. This explains the implosive, which is otherwise absent in Taruma.

A potential case of a different type of segmental adaptation can be seen in *arakabasa* [a.ɾa.ka.ˈba.sa] 'shotgun', which is a borrowing of Portuguese *arcabuz* [aɾ.ka.ˈbus] (Ashby et al. 2012a). The [u] changing to [a] in the loanword is an example of what Peperkamp (2005, cited in Kang 2011: 2260) terms *unnecessary repair*. In this process, an element is adapted even though it is not in violation of the native phonology; [u] exists in Taruma as well, as an allophone of /o/, so that the repair to /a/ is not necessary to conform with Taruma phonology. It should be noted though that, similarly to the other loanwords of Indo-European origin, it is not entirely certain that Taruma borrowed this word from Portuguese directly; it might have been borrowed in its already adapted form via a different Indigenous language.

3.6.2 Phonotactic adaptation

On the level of phonotactic structure, there are two constraints that are often violated in loanwords: Taruma does not allow consonant clusters, nor consonants other than /n/ in the coda (and other nasals in surface forms of combinations of nasal vowels followed by a stop). These illicit structures are usually repaired by means of vowel epenthesis. According to Kang (2011: 2270), this is a common adaptation strategy cross-linguistically, especially in word-initial consonant clusters, and often preferred to deletion. This tendency can be explained based on Paradis and LaCharité's (1970) Preservation Principle, according to which loanwords are "preserved as much as possible, unless the cost of preservation is too extreme". An example of vowel epenthesis in initial consonant clusters is *spoon*. In most renditions of this word, a vowel is inserted between the [s] and [p]. As a result, the two consonants both form a simple onset in CV syllables instead of a CC cluster. The repair of coda consonants can be seen, for instance, in [a.ɾa.ka.ˈba.sa], where an epenthetic [a] is inserted after the [ɾ] and [s] of *arcabuz.* Consequently, coda [ɾ] and [s] are reanalysed as onsets of the resulting CV syllables, so that Taruma syllable structure is preserved. The same process can be observed in *karita* [ka.ˈɾi.ta] 'paper', originating from Portuguese *carta* [kˈaɾ.tə], as well as in the final vowel of [sa'ɾiɾi] (Ashby et al. 2012b).[19] Most cases of phonotactic adaptations are found in loanwords of Indo-European origin (likely because many languages of the region, have a CV syllable structure). There is only one example of vowel epenthesis in a loanword from an Indigenous language, namely the insertion of [a] to derive *soma* [ˈsu.ma] 'tobacco' from Wapishana *soom*. This can be expained by the fact that the Taruma borrowed this Arawakan root from Mawayana *tuma* first, as per Farabee (1918: 282), and only recently adapted its first vowel to /s/ in line with its form in Wapichana, a language they speak today.

3.6.3 Differential importation

While most loanwords undergo phonological adaptation as discussed above, this is not always the case. In some cases, part of the structure of a loanword is borrowed into the native language despite being illicit in its phonology, which is referred to as differential importation (Kang 2011: 2260). In the data, there are some examples of both phonemes and phonotactic structures

---

[19] *Karita* may well be a borrowing via Wapishana, but the same adaptation process applies in Wapishana, which does not permit consonant clusters consisting of a rhotic and [t].

that are not found in words of Taruma origin. Examples of imported phonemes can be found in the borrowing of En. *light* and *rice*, which appear in the data as ['laj.te] and [ɹaɪs], with the voiced alveolar lateral approximant [l] and voiced alveolar approximant [ɹ] instead of the expected Taruma liquid /ɾ/. Illicit syllable structures include [spun], one iteration of the word *spoon* in which there is no epenthetic vowel repairing the initial consonant cluster. Further examples include ['bak.sa] 'box', where an epenthetic [a] is inserted to repair the coda consonant but not the consonant cluster formed of [k] and [s], and ['ɹaɪs.ka] 'rice', in which no vowel is inserted before the non-subject marker *-ka* to repair the syllable with [s] in the coda. It should be noted, however, that these English loanwords are infrequent in the data and might simply be cases of ad-hoc coinages which are not properly integrated into Taruma vocabulary and phonology.

3.6.4 Epenthetic vowels

Several theories have been put forward to account for the cross-linguistic differences between the vowels inserted epenthetically into loanwords (see Kang 2011: 2273 for a broad overview). The most commonly accepted theories include vowel harmony, consonantal assimilation, and the insertion of a default vowel, which Uffmann (2006) argues interact with each other. The first two are assimilatory processes, in which the epenthetic vowel copies the features of a neighboring vowel or consonant. The third process stipulates that a given language has a vowel that is simply the default and is inserted independently of other phonological factors (Uffmann 2006: 1080-1081).

At this point, it is not possible to determine with much certainty which principles govern the quality of the epenthetic vowels in Taruma. With regard to many of the loanwords, especially those originating from European languages, it is difficult to ascertain if they were borrowed directly or via other Indigenous languages that Taruma was in contact with in the past. Based on the data available, there are several possible principles that could dictate which vowel is inserted. Firstly, it seems that /a/ could be the default vowel, as it appears in [a.ɾa.ka.ˈba.sa], ['su.ma], ['bak.sa] and ['po.ta] 'put'. It is not only inserted within morphemes but also at morpheme boundaries, as in *spoonama* ['spu.na.ma] 'my spoon', where it intervenes between *spoon* and *-ma*, the first-person possessive suffix. However, [a] is not always the epenthetic vowel; [i] is inserted in [sa.ˈɾi.ɾi] and [ka.ˈɾi.ta], and /i/, /ɨ/ or /u/ are inserted in *spoon*. Neither can this be explained by vowel harmony: while the epenthetic /i/ in /sa.ˈɾi.ɾi/ and /ˈlai.ti/ might be a copy of the preceding vowel, this process does not account for [ka.ˈɾi.ta] nor for the different forms of *spoon*. Consonantal assimilation does not provide an explanation either. In [sa.ˈɾi.ɾi] and [ka.ˈɾi.ta], the high vowel /i/ might be linked to the coronal feature of the preceding rhotic, but there are several examples of coronal consonants which are followed by epenthetic [a], as in ['po.ta] and [a.ɾa.ka.ˈba.sa]. It might be the case that /i/ is only inserted in environments with both coronal consonants and high vowels, or that the exceptions can be explained due to being borrowed via other languages, with already existing phonological adaptations. Neither of these analyses is entirely satisfactory for the variation in the word *spoon*, however. To sum up, Taruma might have an epenthetic default vowel /a/ but in some words, /i/ is inserted instead, potentially due to processes such as assimilation to a neighboring vowel or consonant.

3.7 Morphophonological processes

This subsection describes the two morpho-phonological processes that are apparent in the data at this point: the insertion of epenthetic vowels and the truncation of stem-final diphthongs in some words. However, the functions of some suffixes in the data are unclear, and there are likely more suffixes that have not yet appeared in the data. Therefore, more data is necessary to explain why the truncation process occurs in some cases but not others.

Firstly, epenthetic vowels appear in consonant clusters in loanwords, but are also inserted when the first-person possessive suffix *-ma* attaches to roots ending in *-n*:

(37)　*spoonama*
　　　spoon–a–ma
　　　spoon–EP–1SG.POSS
　　　'my spoon'

Here, an epenthetic [a] is inserted between the root itself and the first-person possessive suffix *-ma*, so that /n/ in the coda and /m/ in the following onset are not adjacent. It is not certain at this point to which suffixes this process applies. The only other example of an inflected word with stem-final /n/ is *peritonka*, which lacks an epenthetic vowel between the root *periton* 'wild banana' and *-ka*. It might be the case that vowel insertion after stem-final /n/ only takes place if the onset of the suffix is a nasal.

The second process concers the diphthongs /ia/ and /oa/ in stem–final syllables, which are sometimes truncated when the words are inflected:

(38)　a.　['a.ɾ**ja**]　　b.　[a.'ɾ**i**.ka]
　　　　　aria　　　　　　*ari–ka*
　　　　　fish　　　　　　fish–NS
　　　　　'fish'　　　　　　'fish'

(39)　a.　['tɕe.**wa**]　　b.　['tɕi.**u**.ka.hĩ]
　　　　　chiwa　　　　　*chio–ka–hĩ*
　　　　　song　　　　　　song–NS–TAM
　　　　　'song'　　　　　　'sing'

In these examples, the second vocalic segment of the diphthongs is elided when followed by non-subject suffix *-ka*. As a result, only the first vowel remains in the nucleus, so that the diphthongs /ia/ and /oa/ are reduced to /i/ and /o/, respectively. Further examples of truncation include *arino* and *mariki*, in which collective number marker *-no* and instrumental *-ki* are attached to aria 'fish' and maria 'knife', respectively.

However, this process does not apply to all cases of stem-final diphthongs. The diphthong /ia/ in *ikya* /i.kia/ '3' is not reduced when followed by the plural or non-subject suffixes in *ikyano* and *ikyaka*. Moreover, there are no examples of this process applying to diphthongs other than /ia/ and /oa/. However, currently, the number of words with stem-final diphthongs which appear with and without additional morphemes is too limited to establish if factors such as word class of the root, the type of suffix (inflectional or derivational) or further phonological characteristics of either the root or affix play a role in this pattern.

50

3.8 Suggested orthography

This subsection gives an overview of the linguistic and non-linguistic aspects that need to be considered when developing an orthography for an unwritten language (3.9.1). Based on these considerations, an orthography for Taruma is suggested (3.9.2).

3.8.1 Orthography development

The two elements defining orthographies are "a set of graphemes, such as an alphabet, and a set of accompanying rules regulating their use", both of which have to be standardised (Seifart 2006: 277). There are many factors influencing the choice of both the graphemes and the corresponding orthographic conventions to be used; as Seifart (2006: 275) writes:

> orthography development is […] a highly complex issue, which involves not only phonological, prosodic, grammatical, and semantic aspects of the language to be written, but also a wide variety of non-linguistic issues, among them pedagogical and psycholinguistic aspects of reading and writing and the sociolinguistic situation.

Before these issues are discussed in more detail, it is important to note that orthography development is a collaborative process, during which the speakers have the final say about the orthographic conventions (Lüpke 2011: 335). They are, after all, the principal users of the orthography and have a right to express their language in a way that suits them best. Cahill and Karan (2008: 12) also note that speakers are more likely to use the orthography if they have taken part in its development.

To address the linguistic considerations of devising an orthography, the structure of the language in question needs to be determined, especially with regard to its phonology and morphosyntax (Seifart 2006: 276). Once the analysis is completed, the first choice that has to be made pertains to the type of writing system. Two principal types are distinguished based on the linguistic unit which their graphemes represent (Seifart 2006: 277). In morphographic writing systems, such as Chinese characters, morphemes constitute the unit that is represented. In contrast, the graphemes in phonographic writing correspond to elements of the phonology of the language, usually either syllables, such as in the Japanese syllabaries, or phonemes, as in the case of the Latin alphabet (Seifart 2006: 277).[20] After choosing a writing system, one needs to decide which graphemes to employ and which level of the (morpho)phonological structure they represent; this is also referred to as the "depth" of the orthography (Seifart 2006: 279). In shallow orthographies, graphemes correspond to phonetic realisations rather than to underlying phonemes, whereas deep orthographies tend to employ the same graphemes for the different surface realisations of phonemes (Seifart 2006: 279). As a result, languages with few (morpho)phonological processes can be represented well by means of a shallow orthography, while languages characterised by more complex morphophonology tend to employ a deeper orthography (Seifart 2006: 279).

---

[20] It should be noted that in practice, morphographic and phonographic writing are not polar opposites, as many languages combine aspects of the two (Seifart 2006: 278).

In addition to the level of linguistic structure, it also needs to be decided which phonological contrasts to represent in the orthography. Based on Pike's (1947) phonemic principle, the ideal orthography was long considered to be one in which each phoneme corresponds to one grapheme and vice versa (Lüpke 2011: 313); however, as has already become clear with regard to the orthographical depth, this is not always the case. A related issue to consider is underrepresentation: in order to avoid an overly complicated orthography, some phonological features may not be represented if they have a low functional load (Seifart 2008: 280). This refers to their frequency of occurrence, especially in phonologically contrastive contexts; thus, if a given feature differentiates only a small number of words, it is said to have a low functional load and need not be represented in the orthography (Seifart 2006: 280). The underrepresentation of such features results in homographs, such as English *protest*, where stress is the differentiating feature. This ambiguity usually does not cause readers difficulty, as they are able to decode the correct meaning due to the syntactic and semantic cues, such as the different word classes of the above example (Seifart 2006: 280). Finally, after careful morphosyntactic analysis, the word boundaries need to be established and it can be decided how to mark them orthographically. It is advised to distinguish between affixes, independent morphemes and clitics (Kutsch Lojenga 2014: 82, cited in Jones and Mooney 2017: 20-21): affixes are often written conjunctively with the morphemes they are bound to, whereas free morphemes tend to be written separately from each other. Clitics, which are phonologically bound but syntactically independent, can be represented through either of the two options.

When it comes to choosing between monographic and phonographic scripts and the depth of the orthography, sociolinguistic and pedagogical factors also play an important role, especially in the context of endangered languages. As they do not represent phonological information, morphographic writing systems have the advantages that they can be pronounced differently according to different dialects and that they remain stable over time; in contrast, phonographic systems usually require reforms to adapt to sound changes (Bradley 2005: 7, Jones and Mooney 2017: 14). However, the speaker demographic also needs to be taken into account. As Seifart (2006: 283) points out, in the context of endangered languages, it is often the case that the principal age group that will make use of the orthography is the younger generation. The heritage speakers of this age group often have not acquired the language from childhood and are therefore comparable to L2-learners in some respects. Therefore, phonographic scripts are often preferable due to the finite number of graphemes, which makes them easier to learn compared to morphographic writing systems (Jones and Mooney 2017: 14). Furthermore, unlike in morphographic scripts, the sounds in phonographic writing systems are identifiable from the graphemes, which also facilitates the learning process for heritage speakers and L2-learners (Grenoble and Whaley 2006: 158, cited in Jones and Mooney 2017: 14). For the same reason, morpho-phonological complexity is not the only deciding factor for the depth of an orthography. Choosing a shallower orthography can be advisable, as the transparency allows speakers to pronounce words accurately even without L1-speaker intuition of their morphophonology (Seifart 2006: 283). A final practical consideration is that the orthography should not constitute an obstacle with regard to technology and digitisation and should therefore be reproducible on a keyboard. Hence, the use of already existing scripts is usually preferable to the invention of a completely new script (Seifart 2006: 286).

A further factor that is shaped by both sociolinguistics and pedagogical considerations is the relation with existing orthographies. Endangered languages are often spoken in multilingual environments, so that speakers are usually already familiar with the orthography of at least one dominant language (Lüpke 2011: 313). In these multilingual settings the question of the ideological distance to already existing orthographies is especially important, as languages and their scripts and orthographies are often intrinsically linked to the speakers' sense of identity (Lüpke 2011: 330). Depending on how important it is for the users to emphasise their identity as speakers of a minority language, a different script or orthography may be used to mark the endangered language as different from the majority language(s) (Lüpke 2011: 330). But the speakers may also prefer their orthographic conventions to resemble those of the dominant language as much as possible for pedagogical reasons, namely in order to facilitate transfer of literacy: if they are already literate in the dominant language(s), then similar conventions make it easier to learn to read and write (Jones and Mooney 2017: 26). However, the orthography should still be adapted to the specific structural characteristics of the language, as it is not necessarily possible to accurately represent a language through the orthographic conventions of a different language (Seifart 2006: 282).

3.8.2 Suggested Taruma orthography

The guiding principle for the development of this suggested orthography is to maximise ease of transfer since the principal users of the orthography will likely be the heritage speakers, who write in Wapishana and English. Therefore, the proposed standard resembles Wapishana orthography as much as possible (Table 28). The Taruma speakers are well-integrated into the Wapishana community where they live, so that it is not necessary to distance the Taruma orthography from the Wapishana writing for sociopolitical reasons.

Table 28. Present orthography of Guayanese Wapishana (based on Gomes 2022: 146-148).

Phoneme	Orthographic representation	Orthographic example	IPA example	Meaning
/a/	<a>	<awaru>	[aˈwaɾɨ]	wind
/aː/	<aa>	<paapai>	[paˈpai]	father's brother
/ã/	<ã>	<ã'ãi>	[ãˈãi]	okay, all right
/ɓ/	<b>	<badi>	[ˈɓadɨ̥]	cassava bread
/tʃ/	<ch>	<chaakoi>	[tʃaːˈkoj]	toucan
/d/	<d>	<dawatan>	[dawaˈtan]	search
/ɗ/		<daari>	[ˈɗɑːɾɨ̥]	black mud
/ɗj/	<dy>	<dyuwuza>	[ɗjiˈwɨzḁ]	a type of palm
/ɡ/	<g>	<agao>	[aˈɡao]	ouch
/h/	<h>	<hai>	[ˈhai]	yes
/i/	<i>	<iripi>	[iɾipi̥]	a type of fruit
/iː/	<ii>	<piimuda>	[piːˈmɨɗḁ]	hummingbird
/ĩ/	<ĩ>	<ĩnao>	[ĩˈnjao]	they, them
/k/	<k>	<kuruku>	[kɨˈɾɨkɨ̥]	fowl

/m/	\<m\>	\<maamai\>	[maˈmai]	mother's sister
/n/	\<n\>	\<ninoba\>	[niˈnoɓa̰]	tongue
/ɲ/	\<ny\>	\<nyukunuu\>	[ɲɨkɨˈnɨː]	heart
/u/	\<o\>	\<oridi\>	[uˈɽid̥ḭ]	dove
/uː/	\<oo\>	\<pootan\>	[puˈtan]	blow
/ũ/	**\<õ\>**	**\<õgaru\>**	[ũˈɡaɽɨ̰]	I, me
/p/	\<p\>	\<pinidi\>	[piˈnid̥i̥]	grass
/ɽ/	\<r\>	\<ramita\>	[ɽaˈmitjɛ]	break
/s/	\<s\>	\<soowa\>	[suːwa̰]	spider
[ʃ]	\<sh\>	\<shii\>	[ʃiː]	swelling
/ʂ/		\<shawarau\>	[ʂawaˈɽaɨ]	a type of palm
/t/	\<t\>	\<tikazi\>	[ˈtiˈkɛzi̥]	fire
/ɨ/	\<u\>	\<uruu\>	[ɨˈɽɨː]	he, him
/ɨ̃/	\<ũ\>	\<zũũ\>	[zɨ̃ː]	sound of motor vehicle
/ɨː/	\<uu\>	\<tuukii\>	[tɨːˈkiː]	very
/w/	\<w\>	\<wapichan\>	[wapĩˈt͡ʃan]	Wapishana
/j/	\<y\>	\<yawari\>	[jaˈwaɽḭ]	mouse opossum
/ʐ/	\<z\>	\<zakapu\>	[ʐaˈkapḭ]	farm
/ʔ/	\<'\>	\<imi'i\>	[imiʔi̥]	sand, soil

The Wapishana writing system is phonographic, with the graphemes representing individual phonemes; the same applies to the Taruma orthography suggested here, which is given here again for ease of reference (Table 29). An additional advantage of largely basing the Taruma orthography on the written conventions of Wapishana is that no problems arise with regard to reproducing it on computers.

Table 29. Suggested Taruma orthography.

Phoneme	Allophone	Grapheme	Example (phonetic)	Example (orthographic)	Meaning
/a/	[a, ɛ]	\<a\>	[ˈa.tɕi]	\<achi\>	face, eye
/ã/	[ã, ãN]	\<ã\>	[ka.ˈɾã̃ŋ.gʷa]	\<karãgwa\>	otter
/b/	[b]	\<b\>	[aˈba.ɾa]	\<abara\>	back
	[w, β]	\<w\>	[ˈwo.βo]	\<wowo\>	termite
/d/	[d]	\<d\>	[ˈda.nɨ]	\<danu\>	jaguar
	[dʑ]	\<dj\>	[dʑa.ˈʔa.ɾa]	\<dja'ara\>	pet
/ɸʷ/	[ɸʷ]	\<fw\>	[ɸʷa]	\<fwa\>	fire
/g/	[g, gʲ]	\<g\>	[go.ˈmi.tɕi]	\<gomichi\>	woman
/gʷ/	[gʷ]	\<gw\>	[ˈgʷi]	\<gwi\>	dig
/h/	[h]	\<h\>	[ˈha.ma]	\<hama\>	I, my
/i/	[i, e]	\<i\>	[bi.ˈkʲe.ɾu]	\<bikiro\>	balata tree
	[j]	\<y\>	[ˈjo.ɾo]	\<yoro\>	slow
/ĩ/	[ĩ, ĩN]	\<ĩ\>	[ˈhĩn.dʑa]	\<hĩdja\>	no
/k/	[k, kʰ, ʰk, kʲ]	\<k\>	[ˈko.ba]	\<koba\>	arrow

/kʷ/	[kʷ]	<kw>	['kʷa.sɨ]	<kwasu>	person
/m/	[m]	<m>	[ma.'ɾa.gʷi]	<maragwi>	stingray, firefan
/n/	[n, nʲ]	<n>	[na.'ka.nʲi]	<nakani>	know
/o/	[o, u]	<o>	[o.ɾo.'ko.da]	<orokoda>	knee
	[w]	<w>	['wa.ɾa]	<wara>	put
/õ/	[õ, õN]	<õ>	['a.ɸõ]	<afõ>	hand
/p/	[p, pʰ, pʲ]	<p>	['pa.na]	<pana>	dawn
	[ɸ]	<f>	['ɸo.na]	<fona>	wasp
/ɾ/	[ɾ]	<r>	[ɾi.'kʲi.nʲiu]	<rikinyo>	cricket
/s/	[s]	<s>	['sɨ.wi]	<suwi>	throw
/ɕ/	[ɕ]	<sh>	[ɕo.'no.ɾa]	<shonora>	crane species
/t/	[t, tʰ]	<t>	['ta.kʲi]	<taki>	ask
	[tɕ, ʰtɕ]	<ch>	[tɕi.'na.ɾi]	<chinari>	shin
/ɨ/	[ɨ]	<u>	[a'dɨku]	<aduko>	hair
	[ɛ]	<e>	['hɛ.ɾa]	<hera>	arrive
/ɨ̃/	[ɨ̃]	<ũ>	[nɨ̃]	<nũ>	bad, ugly
	[ɛ̃]	<ẽ>	['ɛ̃.hɛ̃]	<ẽhẽ>	flesh
/z̺/	[z̺]	<z>	['ho.z̺a]	<hoza>	rain
	[dz̺]	<dz>	['dz̺ɨ.dz̺ɨ]	<dzudzu>	language
/ʔ/	[ʔ]	<'>	['ɾa.ʔwa]	<ra'wa>	give birth

Wapishana and Taruma share a large part of their phoneme inventory, which is represented by the same set of graphemes. This concerns <a, ã, d, g, h, i, ĩ, k, m, n, o, õ, p, s, t, u, ũ, z, '>.[21] In addition, there are several Wapishana and Taruma phonemes which are similar to each other and are therefore also represented by the same graphemes. The Taruma rhotic /ɾ/, the alveolo-palatal fricative /ɕ/ and the voiced bilabial plosive /b/ are expressed by means of <r, sh, b>, respectively, similarly to Wapishana /ɽ, ʂ, ɓ/. The Taruma phonemes /kʷ/, /gʷ/ and /ɸʷ/ do not have an equivalent, as labialisation is a phonological process in Wapishana. But since <w> represents /w/ in Wapishana, this grapheme was chosen as the second element of the digraphs <kw, gw, fw>. Wapishana does not have phonemic front mid vowels, and even though [ɛ] is also an allophone in this analysis of Taruma, it is nevertheless written as <e> in this thesis in the contexts where it is not clearly alternating with /a/. In other contexts, it is not entirely predictable which underlying phoneme [ɛ] is a realisation of, but representing it by means of <e> makes the pronunciation apparent.

There are relatively few morphophonological processes in Taruma, which is why the orthography suggested here is quite shallow in so far that several allophones resulting from phonological processes are represented by separate graphemes. Neither pre-aspiration nor aspiration of the voiceless plosives are represented, as they are not contrastive and occur in free variation. The palatalisation of /n/, /p, /k/ and /g/ before /i/ is not marked either but instead defined as a pronunciation rule. However, the palatalised allophones of /t/ and /d/, [tɕ] and [dʑ],

---

[21] As can be seen in table 28, the grapheme <d> in Wapishana denotes both the voiced alveolar plosive /d/ and the alveolar implosive /ɗ/. Taruma does not have the implosive, so that <d> marks the alveolar plosive only.

are represented by <ch> and <dj>, respectively. Making the pronunciation of these allophones transparent in the orthography is beneficial for the heritage speakers, especially since there is also phonological palatalisation in Wapishana, which is progressive rather than regressive, as in Taruma (dos Santos 2006: 46-47). Therefore, representing the palatalised allophones separately from their non-palatalised counterparts lessens the chance of interference from Wapishana. The grapheme <ch> was chosen to represent [tɕ] because it is used for the voiceless postalveolar affricate /tʃ/ in Wapishana. The grapheme <dj> expresses [dʑ] because the heritage speakers are familiar with using <j> for English /dʒ/. Due to the comparative similarity of /dʒ/ and /dʑ/, <j> was chosen as the second half of the Taruma digraph, while <d> was added in order to make it more apparent that the represented sound is an affricate. For the same reasons, [dʐ] is expressed by means of <dz>; using this digraph makes it clear that the sound in question is an affricate corresponding to the fricative /ʐ/.

The vowels /i/ and /o/ and their allophones [j] and [w] are written as distinct graphemes as well, namely <i, o> and <y, w>, respectively. Representing these vowels by means of graphemes that often express consonants may not be ideal; it obfuscates the relation between the allophones, while suggesting a connection between /o/ and /b/, which can both be realized as [w] in certain circumstances. However, this orthographic choice facilitates the parsing of syllables and syllable boundaries significantly, which is beneficial for the heritage speakers. If the vowels in words such as *koya* ['ku.ja] 'calabash' and *rikinyo* [ɾi.'kʲi.nʲiu] 'cricket' were written exclusively with vowel graphemes, as *koia* and *rikinio*, then the spelling would be ambiguous, as the words could also be pronounced as ['kwi.a] and [ɾi.kʲi.'nʲiw]. Furthermore, using <y> and <w> disambiguates words that would otherwise be homographs. Both *kway* /kʷai/ 'parch' and kwai /'kʷa.i/ 'spill' would be spelt *kwai*, while *koi* /'ko.i/ 'take off, take out' could be read as ['kʷi] and be confused with *kwi* /'kʷi/ 'poison'. Both word pairs encode transitive verbs, so that there are no syntactic cues disambiguating them. As a result, they are not easily distinguishable from each other based on context only, which is why orthographic differentiation is useful in this case. The nasal diphthongs are written according to the same principals as the oral ones insofar that the approximant-like segment is represented by a consonant grapheme; for instance, /õã/ 'sun' is written as <wã>. Long vowels as in <djiidji> 'grandfather', which are not lengthened due to phonological processes, are represented with doubled vowel graphemes, like in Wapishana.

In terms of suprasegmentals, stress is unmarked, as it is largely predictable and not contrastive for the most part. Imperatives, which are distinguished from most other verb forms by means of stress, can be disambiguated based on contextual cues. So far, it was not possible to discuss this orthography with the Taruma speakers. During the fieldwork, the phonological analysis had not yet been completed, and afterwards, the technological problems made working on the orthography together remotely impossible. Therefore, the orthographic conventions presented here should be understood merely as a suggestion and as a guide to the transcriptions. During the next visit to Maruranau, the matter will be discussed with the speakers and the changes they suggest will be implemented.

3.9 Comparison with historical wordlists

In this subsection, the contemporary Taruma data are compared to the historical wordlists compiled by Schomburgk (1848) and Farabee (1918), to yield additional information about the history of the language. This analysis comes with the caveat that the international phonetic alphabet did not yet exist at the time of their travels and that neither of them was a trained linguist. Therefore, their transcriptions are not fully reliable; consequently, the results of this comparison have to be considered tentative.

A first observation is that some words in Farabee's wordlist, such as <dun> 'jaguar' and <kwakĕp> 'fish hook', appear without the final vowel which is found in our data: *danu* ['da.nɨ] and *kwakipa* [kʷa.'kʲe.pa]. This suggests that the deletion of final vowels might not simply be transfer from Wapishana resulting from Ms Suttie's and Mr Louis' personal linguistic background, but that it might have already been a phonological process in the speech of the Taruma that Farabee encountered at the beginning of the twentieth century. However, at the time of Farabee's visit, the Taruma were already in close contact with the Wapishana (Carlin and Mans 2015). Therefore, this phonological process could have still been caused by contact with Wapishana, where it occurs more consistently than in Taruma.

Secondly, Schomburgk's (1845: 45) comment that Taruma has "the sound of the English th and the Greek ph" has become clearer. The Greek letter <φ> is usually transliterated as <ph> in the Roman alphabet and represents the voiceless labiodental fricative /f/ (Arvaniti 1999: 168). Therefore, "the Greek ph" likely refers to the bilabial fricative [ɸ] in Taruma. However, it is unclear which sound Schomburgk perceived as "the English th", or if his description refers to a single sound similar to both "the English th and the Greek ph". In any case, voiceless dental fricatives are not found in our data, and nor does the digraph <th> appear in his transcriptions, so that no inferences can be drawn from the contemporary pronunciation. Another question remains with regard to Schomburgk's record: his list does not distinguish between what are voiced and voiceless plosives in the current pronunciation. The voiceless plosive in *api* [apʲe] 'foot' is transcribed as a voiceless stop in <appa>, but so are the sounds that are voiced plosives in our data. For instance, *ada* ['a.da] 'head' and *biwa* ['bi.wa] 'moon' are transcribed as <atta> and <piwa>, respectively. This is puzzling, since Schomburgk, as a speaker of German and English, would certainly have been aware of the contrast between voiced and voiceless plosives. If he did not simply mishear the voiced stops, then a potential explanation could be that the variety of Taruma that he transcribed had a contrast between aspirated and unaspirated voiceless stops instead of between voiceless and voiced stops. Since aspirated and unaspirated voiceless plosives are allophones of the same phonemes in German and English, Schomburgk might not have been able to distinguish between these sounds (Antonsen 2007: 72; Jensen 1993: 2).

Thirdly, the wordlists are in accordance with the analysis of the allophones of /p/, /t/ and /d/. There are no <po> sequences in either list, which conforms to the analysis that /p/ is subject to allophonic variation when preceding /o/. Similarly, neither Schomburgk nor Farabee lists the sequence <ti>, and while the sequence <di> appears in Farabee's list, the contemporary pronunciation of the words which also appear in our data features the high central vowel /ɨ/ instead of /i/. For instance, Farabee's <diwi> 'near' and <badibapi> 'axe' are pronounced ['dɨ] and ['ba.dɨ] in our data. This fits the analysis of [tɕ] and [dʑ] as allophones of /t/ and /d/ before

/i/. What is more, Schomburgk's transcription of the number two, <tyuwa> (['dʑu.wa] in our data), provides additional support for this analysis. The digraph <ty> could correspond to a palatalized alveolar stop [tʲ], which is the expected intermediate step in the palatalization of the alveolar stops: [ti] and [di] sequences would likely have palatalized to [tʲ] and [dʲ] before being realized in their current form as [tɕ] and [dʑ].

The final observation concerns the nasals and nasal vowels. Neither Schomburgk nor Farabee transcribe nasal vowels; instead, they note sequences of vowels and nasals. Farabee only transcribes these instances of nasals before consonants, such as in <inja> 'not' and <karangwi> 'otter', which resemble the contemporary pronunciation ['hĩ.dʑa] and [ka.'rã.gʷa]. Possibly, Farabee was not familiar with nasal vowels and only perceived nasality before consonants due to the homorganic nasal allophones. Schomburgk transcribes what he presumably perceived as a velar nasal [ŋ] in <ouang> 'sun' but does not make note of nasality in <ahu> 'hand'. These words are pronounced ['õã] and ['a.hõ] in our data, with nasal vowels in the respective final syllable. The transcriptions suggest that if the nasal vowels do indeed originate from Vn sequences, this development would have been completed at the time when Schomburgk collected his wordlist. Otherwise, <ahu> would have been transcribed with a nasal as well, but Schomburgk presumably did not perceive the nasality in this word, as the nasal vowel appears in an unstressed syllable. Therefore, these transcriptions do not provide evidence for a diachronic development of the nasal vowels, but they do not contradict this assumption either. However, they might shed light on a different word in which there is a mismatch between the historical and contemporary data with regard to nasality. Both Schomburgk and Farabee transcribe the word 'star' with a nasal, as <wingra> and <uinra> respectively, while it is pronounced without nasality on our recordings. It is possible that this word used to contain a nasal vowel which is now lost; this would explain why it is a homophone of *hwira* ['hʷi.ɾa] 'sloth' today.

4 SUMMARY AND CONCLUDING REMARKS

The Taruma phoneme inventory consists of 17 consonants as well as four vowels and their nasal counterparts. The phonological processes identified include pre-aspiration, the palatalisation of /p/, /k/ and the alveolar nasal and oral stops before /i/, and the lenition of the bilabial stops before /o/. In addition, one morpho-phonological process apparent at this point consists of the truncation of diphthongs in some words when a suffix is attached. Syllables are primarily open, with /n/ being the only consonant permitted in the coda, and onsets consist of maximally one consonant. The violation of these phonotactic constraints in loanwords is often repaired through the insertion of epenthetic vowels. The minimal syllable consists of a single vowel, and syllables can be light, heavy, or superheavy. Stress is usually fixed on the penultimate syllable in uninflected roots and reassigned to remain penultimate in inflected words, but there is some variation with the non-subject suffix *-ka* and the first-person singular possessive suffix *-ma*. The sound symbolism in the data is of the imitative type, is found primarily in animal names and terms for environmental sounds and is often achieved through reduplication. Both in the Taruma phonology and its vocabulary, signs of language contact are apparent. In terms of the lexicon, there are several words shared with other Indigenous

languages. With regard to the phonology, the Taruma speakers occasionally transfer phonological processes from Wapishana to Taruma, namely the labialisation of velars after /o/ (/u/ in Wapishana) and the deletion or devoicing of final vowels.

Compared to the typology of phonological systems of South American lowland languages compiled by Storto and Demolin (2012), the Taruma system has several similar characteristics. The consonant inventory consists of an average number of segments, with 17 as opposed to an average of 17.5 consonants per language in the 27 language isolates and language families under investigation. It conforms to the observation that almost all languages have at least the voiceless plosives *p*, *t* and *k*, the nasal stops *m* and *n*, and "at least one fricative, *s* or *h*" (Storto and Demolin 2012: 333). Vowel inventories containing four vowels are not uncommon, and the Taruma vowel system resembles the pattern of *i*, *a*, *u*, *ɨ* described for Wapishana (Storto and Demolin 2012: 335, dos Santos 2006). However, there are also several aspects which set Taruma apart. Phonemic nasal vowels, while not uncommon in South American languages generally (Campbell 2012b: 267), are not found in the other languages which are spoken in Guyana and listed in SAPhon, with the exception of Mawayana (Michael et al. 2015). Additionally, in contrast to most other South American languages, the Taruma inventory contains the voiced retroflex fricative /ʐ/. Retroflex consonants are not very widespread outside of the Highland Andean region but are found in several Arawakan languages (Campbell 2012b: 271). Among these is Wapishana, the only other language listed in SAPhon which is spoken in Guyana and shares /ʐ/ with Taruma, and one of only seven other languages on the continent to have this sound. A further unusual characteristic, at least compared to the Cariban, Arawakan, Tupían, and Panoan lowland language families, concerns Taruma syllable structure. While the phonotactic restriction of the onset and coda are found in several of these families as well, none of them are described as permitting triphthongs, as their nuclei allow for maximally two vowels (Storto and Demolin 2012: 340).

Several avenues for further research remain. Firstly, there are several parts of this phonology sketch that would benefit from a larger amount of data. In terms of the history of the language, this concerns the direction of borrowings shared with other Indigenous languages, and the hypothesised diachronic development of phonemes such as /ʐ/, /ɕ/, and the nasal vowels. Secondly, while the influence of Wapishana on Taruma has been laid out here, it has not been examined if there are also transfer effects from Taruma in the Taruma speakers's use of Wapishana. Thirdly, it is likely that more types of sound symbolism exist in Taruma, such as corporeal and synaesthetic symbolism. Forthly, more occurrences of suffixes other than *-ma* would contribute to filling the current gaps in our knowledge, for instance with regard to stress reassignment. The morpho-phonological process of vowel deletion in diphthongs in the presence of suffixes also requires more data to establish a pattern that explains why it occurs in some words but not in others. It is also possible that there are more morpho-phonological processes in Taruma which would become apparent with more data. In addition, the function of some suffixes remains unclear, and it is possible that there are others which have not yet appeared in the data. This can be addressed in a grammar sketch of Taruma, along with other aspects of the morphology and syntax of the language. The list of intonational patterns presented here can also be significantly extended, especially based on recordings of conversations and genres of speech other than wordlist elicitation and narratives. Such data would also allow for an analysis of other aspects of the Taruma language, such as its pragmatic

characteristics. The gaps laid out here can be filled through the continued documentation and analysis of Taruma. This would not only aid linguistic analysis but also continue to create a lasting record of Taruma culture and materials which the heritage speakers can use to improve their proficiency. For this endeavor to be successful, it is essential that the Taruma continue to be actively involved in and shape the documentation process.

REFERENCES

Antonsen, Elmer H. 2007. *Elements of German Phonology and Morphology*. Tuscaloosa: University of Alabama Press.

Arvaniti, Amalia. 1999. "Standard Modern Greek." *Journal of the International Phonetic Association* 29, no. 2: 167–172.

Ashby, Simone, Sílvia Barbosa, Silvia Brandão, José Pedro Ferreira, Maarten Janssen, Catarina Silva, and Mário Eduardo Viaro. 2012a. "arcabuz." In *A rule based pronunciation generator and regional accent databank for Portuguese. Proceedings of Interspeech 2012*. http://www.portaldalinguaportuguesa.org/%3E?action=fonetica&region=lbx&act=details&id=2565 Accessed 3rd June 2022.

Ashby, Simone, Sílvia Barbosa, Silvia Brandão, José Pedro Ferreira, Maarten Janssen, Catarina Silva, and Mário Eduardo Viaro. 2012b. "carta." In *A rule based pronunciation generator and regional accent databank for Portuguese. Proceedings of Interspeech 2012*. http://www.portaldalinguaportuguesa.org/%3E?action=fonetica&region=lbx&act=details&id=6384 Accessed 3rd June 2022.

Ashby, Simone, Sílvia Barbosa, Silvia Brandão, José Pedro Ferreira, Maarten Janssen, Catarina Silva, and Mário Eduardo Viaro. 2012c. "vaca." In *A rule based pronunciation generator and regional accent databank for Portuguese. Proceedings of Interspeech 2012*. http://www.portaldalinguaportuguesa.org/%3E?action=fonetica&region=lbx&act=details&id=30838 Accessed 3rd June 2022.

Beddor, Patrice S. 1993. "The perception of nasal vowels." In *Nasals, nasalization, and the velum*, edited by Marie K. Huffman, and Rena A. Krakow, 171-196. Academic Press.

Boomert, Arie. 1977. "De Taruma: Een verdwenen Indianenstam uit Suriname." *Informatie-tijdschrift over ontwikkelingsvraagstukken van Suriname* 2, no. 2: 13-17.

Borgstrøm, Carl H. J. 1940. *The dialects of the Outer Hebrides*. Vol.1 of *A linguistic survey of the dialects of Scotland* (*Norsk Tidsskrift for Sprogvidenskap*, Supp. Bind 1). Oslo: Norwegian Universities Press.

Bowerman, Melissa, Marianne Gullberg, Asifa Majid, and Bhuvana Narasimhan. 2004. "Put project: The cross-linguistic encoding of placement events." In *Field Manual Volume 9*, edited by Asifa Majid, 10-24. Nijmegen: Max Planck Institute for Psycholinguistics.

Bradley, David. 2005. "Issues in orthography development and reform." In *Heritage maintenance for endangered languages in Yunnan, China*, edited by David Bradley, 1-10. Melbourne: La Trobe.

Bredero, Robin D. 2021. "Retracing the footsteps of the Taruma people." MSc thesis, Wageningen University and Research.

Bureau of Statistics, Guyana. May 2021. "Year-end and Mid-year Population Estimates by Sex, Guyana: 1992 to 2019." Retrieved 17th March 2022. https://statisticsguyana.gov.gy/subjects/demography-vital-and-social-statistics/year-end-and-mid-year-population-estimates-by-sex-guyana-1992-to-2019/

Butt Colson, Audrey, and John Morton. 1982. "Early Missionary Work among the Taruma and Waiwai of Southern Guiana: The Visits of Fr. Cuthbert Cary-Elwes, S.J. in 1919, 1922 and 1923." *Folk* 24: 203–261.

Cahill, Michael, and Elke Karan. 2008. "Factors in designing effective orthographies for unwritten languages." *SIL International*.

Campbell, Lyle. 2012a. "Classification of the indigenous languages of South America." In *The Indigenous Languages of South America: A Comprehensive Guide*, edited by Lyle Campbell and Verónica Grondona, 59-166. Mouton de Gruyter.

Campbell, Lyle. 2012b. "Typological characteristics of South American indigenous languages". In *The Indigenous Languages of South America: A Comprehensive Guide*, edited by Lyle Campbell and Verónica Grondona, 259-330. Mouton de Gruyter.

Carlin, Eithne B. 2011. "Nested identities in the southern Guyana-Suriname corner." In *Ethnicity in Ancient Amazonia: Reconstructing past identities from archaeology, linguistics, and ethnohistory*, edited by Alf Hornborg and Jonathan D. Hill, 225-236. Boulder: University Press of Colorado.

Carlin, Eithne B., and Jimmy Mans. 2015. "Movement through Time in the Southern Guianas: Deconstructing the Amerindian Kaleidoscope." In *In and Out of Suriname: Language, Mobility and Identity*, edited by Eithne B. Carlin, Isabelle Léglise, Bettina Migge and Paul B. Tjon Sie Fat, 76-100. Leiden, Boston: Brill.

Delattre, Pierre, and M. Monnot. 1981. "The role of duration in the identification of French nasal vowels. In *Studies in Comparative Phonetics*, edited by Pierre Delattre, 17-38. Heidelberg: J. Groos.

Eberhard, David M., Gary F. Simons, and Charles D. Fennig (eds.). 2022a. "Guyana." *Ethnologue: Languages of the World.* Twenty-fifth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Eberhard, David M., Gary F. Simons, and Charles D. Fennig (eds.). 2022b. "Taruma." *Ethnologue: Languages of the World.* Twenty-fifth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Endangered Languages Archive. 2022. "Using ELAR." Accessed 19th May 2022. https://www.elararchive.org/how-to-use/

Evans, Clifford, and Betty Jane Meggers. 1960. *Archeological Investigations in British Guiana*. Vol. 177. Washington: US Government Printing Office.

Farabee, William C. 1917. "A Pioneer in Amazonia: The Narrative of a Journey from Manaos to Georgetown." *The Bulletin of the Geographical Society of Philadelphia* 15, no. 2: 57-103.

Farabee, William C. 1918. *The Central Arawaks.* Philadephia: University Museum.

Farabee, William C. 1924. *The Central Caribs*. Philadelphia: University Museum.

Gomes, Adrian. S. 2022. "Introduction of a Wapishana-English bilingual education programme: an evaluation of the early stages." PhD diss., Leiden University. doi:10.48273/LOT0615

Grenoble, Lenore A., and Lindsay J. Whaley. 2006. *Saving languages: An introduction to language revitalization*. Cambridge: Cambridge University Press

Hinton, Leanne, Johanna Nichols, and John J. Ohala. 2010. "Introduction: sound-symbolic processes." In *Sound symbolism*, edited by Leanne Hinton, Johanna Nichols, and John J. Ohala, 1-12. Cambridge University Press.

Holmes, Janet. 2013. *An Introduction to Sociolinguistics* (4th edition). London: Routledge.

Holt, Anne Marie. 2021. "Developing contactless methodology for ethnobotanical research: preliminary documentation of Taruma and Wapichan plant knowledge." MSc thesis, Wageningen University and Research.

Jensen, John T. 1993. *English Phonology*. Amsterdam ; Philadelphia: J. Benjamins Pub. Co.

Jones, Mari C., and Damien Mooney. 2017. "Creating Orthographies for Endangered Languages." In *Creating Orthographies for Endangered Languages*, edited by Mari C. Jones and Damien Mooney, 1-35. Cambridge: Cambridge University Press.

Kang, Yoonjung. 2011. "Loanword phonology." In *The Blackwell companion to phonology, Volume IV: Phonological* interfaces, edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hunne and Keren Rice, 2258-2282. Wiley.

Kaufman, Terrence, with help from Brent Berlin. 2007. "South America." In *Atlas of the World's Languages* (2nd edition), edited by R. E. Asher and Christopher Moseley, 59–94. London: Routledge.

Kutsch Lojenga, Constance. 2014. "Basic principles for establishing word boundaries." In *Developing orthographies for unwritten languages*, edited by Michael Cahill and Keren Rice, 73–106. Dallas TX: SIL International.

Latham, Robert G. 1862. *Elements of comparative philology*. London: Walton and Maberly.

Ladefoged, Peter, and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford, Cambridge (MA): Blackwell Publishing.

Lewis, M. Paul, and Gary F. Simons. 2010. "Assessing Endangerment: Expanding Fishman's GIDS." *Revue roumaine de linguistique* 55, no.2: 103-120.

Loukotka, Čestmír. 1949. "La Langue Taruma." *Journal de la Société des Américanistes* 38: 53-65.

Lüpke, Friederike. 2011. "Orthography development." In *The Cambridge Handbook of Endangered Languages*, edited by Peter K. Austin and Julia Sallabank, 312-336. Cambridge: Cambridge University Press.

von Martius, Carl F. P. 1867. *Wörtersammlung Brasilianischer Sprachen.* Vol. 2 of *Beiträge zur Ethnographie und Sprachenkunde Amerikas zumal Brasiliens*. Leipzig: Friedrich Fleischer.

Michael, Lev, Tammy Stark, Emily Clem, and Will Chang (compilers). 2015. *South American Phonological Inventory Database v2.1.0. Survey of California and Other Indian Languages Digital Resource.* Berkeley: University of California.

Ministry of Amerindian Affairs. 2022a. "About the Ministry of Amerindian Affairs." Accessed 23rd March 2022. https://moaa.gov.gy/about-us/

Ministry of Amerindian Affairs. 2022b. "Amerindian Heritage." Accessed 22nd March 2022. https://moaa.gov.gy/amerindian-heritage/

Ministry of Amerindian Affairs. 2022c. "Indigenous Nations." Accessed 22nd March 2022. https://moaa.gov.gy/amerindian-nations/

Ministry of Amerindian Affairs. 2022d. "Maruranau." Accessed 27th March 2022. https://moaa.gov.gy/indigenous-villages/maruranau/

Ministry of Finance Guyana. 2022a. *Cooperative Republic of Guyana estimates of the public sector: Current and capital revenue and expenditure for the year 2022 as presented to the the National Assembly. Volume 1*. Accessed 23rd March 2022. https://finance.gov.gy/wp-content/uploads/2022/01/Budget_Estimates_Volume_1_2022_webVer_reduced.pdf

Ministry of Finance Guyana. 2022b. *Cooperative Republic of Guyana estimates of the public sector: Current and capital revenue and expenditure for the year 2022 as presented to the the National Assembly Volume 2*. Accessed 23rd March 2022. https://finance.gov.gy/wp-content/uploads/2022/01/Budget_Estimates_Volume_2_2022_webVer_reduced.pdf

Ministry of Finance Guyana. 2022c. *Cooperative Republic of Guyana central government development programme: Capital projects, profiles for the year 2022 as presented to the National Assembly. Volume 3*. Accessed 23rd March 2022. https://finance.gov.gy/wp-content/uploads/2022/01/Budget_Estimates_Volume_3_2022_webVer_reduced.pdf

Moraes, João Antônio, and W. Leo Wetzels. 1992. "Sobre a duração dos segmentos vocálicos nasais e nasalizados em Português. Um exercício de Fonologia Experimental." *Cadernos de estudos linguísticos* 23, 153-166.

Paradis, Carole, and Darlene LaCharité. 1997. "Preservation and minimality in loanword adaptation." *Journal of linguistics* 33, no. 2: 379-430.

Peperkamp, Sharon. 2005. "A psycholinguistic theory of loanword adaptation." *Proceedings of the Annual Meeting, Berkeley Linguistics Society* 30, 341-352.

Pike, Kenneth Lee. 1947. *Phonemics: A Technique for Reducing Languages to Writing*. Ann Arbor: University of Michigan Press

Rivet, Paul. 1924. "Langues de l'Amérique du sud et des Antilles." *Les langues du Monde par un groupe de linguistes*. Paris: 639-707.

Rivière, Peter. G. 1966. "Some ethnographic problems of Southern Guyana." *Folk* 8, no. 9: 301-312.

Rybka, Konrad. 2020. "Linguistic, Cultural, and Environmental Aspects of Ethnoprimatological Knowledge Among the Lokono, Kari'na, and Warao of the Moruca River (Guyana)." In *Neotropical Ethnoprimatology*, edited by Bernardo Urbani and Manuel Lizarralde, 239-262. Ethnobiology. Springer, Cham

Queixalós, Francisco, and Odile Renault-Lescure, eds. 2000. *As Línguas Amazônicas Hoje*. 1a ed. São Paulo [Paris, France]: Instituto Socioambiental, Museu Paraense Emílio Goeldi; Institut de recherche pour le développement.

Rodrigues, Aryon Dall'Igna, and Ana Suelly Arruda Câmara Cabral. 2012. "Tupían." In *The Indigenous Languages of South America: A Comprehensive Guide*, edited by Lyle Campbell and Verónica Grondona, 495-574. Berlin, Boston: De Gruyter Mouton.

Rothe-Neves, Rui, and Hellen Valentim. 2012. "On the duration of nasal vowels in Brazilian Portuguese." *Revista Diadorim* 12, 108-128.

dos Santos, Manoel Gomes. 2006. "Uma Gramática do Wapixana (Aruák): aspectos da fonologia, da morfologia e da sintaxe." Instituto de Estudos da Linguagem.

Schomburgk, Robert H. 1845. "Journal of an Expedition from Pirara to the Upper Corentyne, and from Thence to Demerara." *The Journal of the Royal Geographical Society of London* 15: 1–104. https://doi.org/10.2307/1797903

Schomburgk, Robert H. 1848. "Remarks to accompany a comparative vocabulary of eighteen languages and dialects of Indian Tribes inhabiting Guiana." *Simmond's Colonial Magazine* 15: 46-64.

Seifart, Frank. 2006. "Orthography development." In *Essentials of language documentation*, edited by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 275-299. Berlin: Mouton de Gruyter.

Silverman, Daniel. 2003. "On the Rarity of Pre-Aspirated Stops." *Journal of Linguistics* 39, no. 3: 575–98.

Smithsonian Institution. 2021. "Collections search: Taruma." Accessed 27th March 2022. https://americanindian.si.edu/collections-search/search?edan_q=Taruma

Steriade, Donca. 2009. "The phonology of perceptibility effects: The P-map and its consequences for constraint organization." In *The nature of the word: essays in honor of Paul Kiparsky*, edited by Kristin Hanson and Sharon Inkelas, 151-180. Cambridge, MA: MIT Press.

Storto, Luciana, and Didier Demolin. 2012. "The phonetics and phonology of South American languages." In *The Indigenous Languages of South America*, edited by Lyle Campbell and Verónica Grondona, 331-390. Mouton de Gruyter.

Uffmann, Christian. 2006. "Epenthetic Vowel Quality in Loanwords: Empirical and Formal Issues." *Lingua* 116, no. 7: 1079–1111.

Wapishana Language Project, and Wapichan Wadauniinao Ati'o. 2000. *Scholar's Dictionary and Grammar of the Wapishana Language. Tominpainao Ati'o Wapichan Paradan Parada-karu na'iki Paradauzo-karu kaduzu*. Lethem, Port Velho: Wapishana Language Project.

APPENDIX A. WORDLISTS

A.1 Historical wordlists

A.1.1 Schomburgk (1848: 59-60)

English	Taruma	English	Taruma
sun	ouang	mouth (my)	merukukanna
moon	piwa	hand (my)	ahu
stars	wingra	foot (my)	appa
earth	toto	bow	tzeika
fire	hua	arrow	kupa
water	tza	dog	hi
head (my)	atta	one	oshe
eyes (my)	atzi	two	tyuwa
nose (my)	assa	three	ungkeha

A.1.2 Farabee (1918: 277-283)

Key to the transcriptions (Farabee 1918: 185)

a	as in	father	ai	as in	aisle
ă	"	hat	au	"	how
e	"	fete	oi	"	oil
ĕ	"	met	c	"	ship
i	"	pique	tc	"	chain
ĭ	"	pin	hw	"	when
o	"	note	kw	"	quake
ŏ	"	not	ñ	"	cañon
u	"	rule	a'a, i'i	as broken vowels	
ŭ	"	but	a·i, a·u, o·i	as individual sounds	

English	Taruma		
		are you there?	habia
		armadillo	marura
agouti	tuh	armadillo, small	kabaiyo
agouti, small	huki	arrow	kuba
anaconda	helitcitci	aunt	maini
anato paint	kalo·il	axe	badibapi
aniqua bird	haragua		
another	utcaka	balata tree	bikeru
anything fat	hatcăkwa	barbecue (v.)	jauyi

beads	kasur	earth, black	jongwi
belt	ĭnjabo	east	atcakwa
bird	zili	eat, let us	gwaíniwí
blue	daitwik	essiquibo	zi'ikidju
boil	jiku	eye	atci
bottle	pateli		
bow	tceka	far	dunahá
bowl	tala	fat	tcăkwa
bread	tciwi	father	paiyo
bring	awadasi	field, new	tceyo
brother	wi	field, old	gúmeja
brother-in-law	akiwa	finished	aima
		fire	fwa
calabash	gŏlie	fire, to make	kwami
capabara	guara	fish	ale
carrying basket	ásigi	fish hook	kwakĕp
casarip	batcuku	fish line	fitema
chicken	akala	foot	apa
chop	bitchi	forest	nukuda
clay	kuline		
cold	siwa	give water	daziza'
color	samaku	go into hole	hwakikuraba
comb	paleli	go to rear	gwanana
corral	bawasi	good	hia
cotton	mulu	grandfather	abana
creek	kidju	grandmother	agigi
crown	tcĭnika	grater	tcĭmari
curassow	ohona	gun	arkebusa
curassow, white-head	kuyui	gun cap	duladula
cut (*v.*)	tatuka	gunpowder	gulaparu
cutlass	katcupara		
		hammock	sai'ĭ
day after tomorrow	saliki	hand	a·u
day after day after	slikabada	hard	u'u
tomorrow		haul	jujukwa
day before yesterday	noki	having	ha or habia
day before day before	nukibada	head	adam
yesterday		heavy	dupá
deep	gwoa	here	mahiki
deer, bush	ki'ilya	herron, blue	junari
deer, red	kauyi	hot	biau
deer, savannah	hitci	house	duiya
dog	hi	how many?	kwatciwigueraba
duck, cuervus	zakoka	hungry	hukapadika
duck, diver	cuniwad		
duck, mareka	kaibadje	incense wax	zi
duck, moscova	peli	it is so	haka
eagle	mokaua	jaguar	dun
earth	dudo		

kill	băhé	parrot, noisy	kiukiu
kingfisher	bitcitcu	peccary	baiyi
kingfisher, red-brested	sĕleli	peccary (large variety)	hisu
knife	malia	pink	jŏli
know, I do not	ndau	plantain	bĭli
		powis	wokuki
large	wakana	puma	kwailipaua
leaf	jukár		
leak	tcuraka	quick	gwásiki
lean	wĭruwĭru	rain	fuza
lie (v.)	hadé	rain is coming	madakĭnafuza
light, weight	kabú	razor	kídimic
little	gaciwi	red	iciku
lizard	uwana	river	kidju
lizard	asijĕri	rock	api
made	mukimuki	saet	ukwai
make	kuru	scissors	cirata
make fire (v.)	kurukwami	scrub forest	kasikasagwa
male	maito	seat	moeli
man	gika	shallow	jama
many	tcegitcegi	short	ciaiwi
many fish	tcegiale	shoulder basket	pakar
medicine man	kiliku	sick	baukánama
monkey, black	kwagi	sister	atci
monkey, black howler	rumi	slow	hiabakwai
monkey, friarly	koku	soft	yuma
monkey, hooded	kwisa	sore	kwara
monkey, marmoset	hisai	spirit	kantu
monkey, red howler	mu	stars	uinra
monkey, striped	sikwa	stone	api
moon	biwa	strangers	wanitcú
mother	maiyu	strangers are coming	madkĭna–wanitcu
mouth	kukana	strong	dumi
mountain	fwai	sun	hwa
my	hamu	suncreek	wakikidju
near	diwi	take hold	tcuni
nephew	jayiwi	tapir	baki
no	ĭnja	there	duia
nose	aza	thing	kokuna
		thirsty	wanapadika
oil	tcĭmi	throw away	siwi
otter	karangwi	tobacco	tuma
otter pup	kwahau	today	fole
		tomorrow	sali
paca	kaikwi	tortoise	gini
parrot	kurigua	toucan, red-breasted	liyu
parrot, green	karikari		

toucan, white-breasted	ziku	yesterday	nok
trumpet bird	naki	you	a'a
turn	bainikwa	you have come	kwegiag
turtle	yariko	younger brother	u·ui
		your	ani
very little	gaciwigaciwi		
wash	gwaciki	1	ociwai
watch (*v.*)	gugua	2	jowa
water	za	3	mĭkyahahi
waves	uriparu	4	jobarada
weak	ulu	5	dobangurubăb
weary	ulukantua	6	ociwaiyakana'ahami
west	zuniakwakiwaka	7	jowakyakana'ahami
what	gaga	8	mĭkyahahiyakana'ahami
when	hama	9	jobaradakyakana'ahami
where	ndatci, ungwá	10	dobong
white	wokik	11	ociwaiyakanahapabi
without	yulu	12	jowayakanahapabi
woman	gumitca	13	mĭkyahahiyakanahapabi
wood	jukánahu	14	jobaradakyakanahapabi
wrap	duma	15	urawapaba

A.2 Contemporary wordlist

Taruma	English	Notes
A		
abana	grandmother	
abara	back	
abashoro	then	
achi	eye, face	
achi wi	eyebrow	
ada	head	
ada aycho	skull	lit. *bone of the head*
aduko	hair	
adjino	backyard	
adjoa	tail	
afwi, ahwi	sweet potato	
afwya	path	
ahõ	hand, brain	
aka	macaw	
akara	chicken	
akara djai	chicken egg	
akiko	liver	
akioko	brother-in-law	
akosa	needle	
akwa	arm	
akwacha	hip, thigh	
akwacho	creek	
ana	you (plural)	
ane	necklace	
apa	foot	
apara	to bury	
apere	testicles	
api	stone	
arakabasa	shotgun	
aria	fish	
asa, asu	nose	
asa koroto	nostril	lit. *nose hole*
asa ashishi	snot	lit. *nose dirt*
asadji	basket type	
asaka	heart	
asape	piaba fish	
ashishi	dirt	
ashishi, ashishiku	lip	
asidjari	lizard species	
asukidjo ashishi	asɨ'kidʑo a'ɕiɕi	lit. *ear dirt*
asukidjo	ear	
asuko	bark, skin	

70

asuo, asuoba	lip
atakwa	a'tak^{w}a
atɛ̃ku	long
ateri, aderi	wrist
awicha	belly
aycho	bone
a'oyhĩ	cry

Ã

ãdjako	sister in law

B

badu	ax
badehĩ	laugh
badjara	heel
badjihĩ	search
badu	axe
badu chinari	axe handle
bafwa	scarlet macaw
bãgwada	gecko
bahĩ	club ; hit, kill
bahĩ fwaka	extinguish the fire
bahõ	snake
baki	tapir
bakĩhĩ	carry
bape	lake
baraka	manioc pan
baridjana	basket
barima	wound, sore
baro	forest frog
bashada	drink type
ba'i	collared peccary
ba'o	warm
ba'okahĩ	be sick
bikidjo	pay, payment
bikiro	balata tree
binini	crab species
biridjanako	caimbe tree
bisãnda	hummingbird
biwa	moon
bĩwa	liana species
biado	savannah
bu'a	run

CH

chabi	wait
chada	kokorite palm
chagi	a lot
chamani	bacaba (small)
charaba	water turtle
charato	hungry
charoko	fish trap
chawi, chayakahĩ	shoot
chayka	bow
chayo	field
chiwa	song
chichihada	grasshopper
chichihado	centipede
chinari	shin
chinika	hat
chiokahĩ	sing
chikwi	fill
chiwi	ite palm
chochoari	catfish
choka	corn
chokohĩ	tell
chomari	manioc grater
choni	catch, grab, hold
chupon, spoon	spoon

D

dabaro	shadow
dakika	mukru plant
damẽ'ẽ	answer
dani	sharpen
danu	jaguar
dara	crab species
da'akwi	come
djabu	older sister
djãda	comb
djadehĩ	grate
djagwuna	visit
djai	egg; child; pregnant
djaika	husband
djaka	dry
djakwado	father
djanese	girl
djasu	wife
djawã	child
dja'adji	father in law, maternal uncle
dja'ara	pet, domestic animal
dja'o, dja'wa	mother

dja'obarada	four
dja'okwahĩ, dja'okwihĩ	help
djiba	afternoon
djiicho	shelter, camp
djiidjima	grandfather
djikʷo, djiko	cooked
djikwina	waist
djimiku	tasty
djini	land turtle
djini toro	land turtle's ladder
djipi	name
djiriri	toad, frog
djisu	buttocks
djitoka	sugar cane
djoko	leaf
djomi	hide
djoa	two
doba	all
dodo	earth,
doko	chest do'koma
domi, domiko	strong
donaha	far
dopa	heavy
doriki	close
dorikina	door, entrance
doya	house
doya ada	roof lit. *head of the house*
du	here
dukato	fart
dukwu	black
dza	water
dzakuri	thirsty
dzudzu	language
dzuko	toucan
dzumi	disappear
dzumusu	forget
dzupi	babysling
dzuri	bird
dzuri asu	bird's beak

Ẽ

ẽ'ẽ	game animal
ẽhẽ	flesh

F

fidjihĩ	want

fo	saliva
fona	wasp
fota	seat
fofoki	curassow
foga	bacaba (large)
fwa	fire
fwã	big
fwi	dead; plant
fwifwi	whistling

G

gana	belongings
gashiwi	small
gatuka	always
gwa'ena	body part
gi	say
gibi, gibiri, gibiro	hit
gida	louse
gigina	wash, clean
gika	man
go	go
gomichi	woman
goro	enter
gwa	ant
gwasuki	right away
gwa'ena	neck
gway	fetch
gwi	dig
gwihiso, wikiso	roast
gwini	bathe
gwiri	fingernail
gwuna	swallow

H

hacha	fat
hãdja	armpit
hagori	corpse
hakari	wake up
hama	I
hami	lizard species
hana	fly
harakwa	maroodi bird
hayari	plant species
hazomara	aimara
hẽda	keep
hera	arrive

here	small alligator	
herichichi	anaconda	
hewe	magic spell	
hi	dog	
hĩ	see	
hyã, hyãku	good, clean	
hyãbidji	hang, place	
hĩdja gana	nothing	lit. *no thing*
hĩdja hama'o	never	
hĩdja hyãkwa	not good	
hĩdja kokina	nowhere	lit. *no place*
hyãdokahĩ	to fear	
hibadjo	pot	
hibuku	salty	
hiso	white-lipped peccary	
hichi	gray brocket deer	
hichi asukidjo	gray brocket deer's ear	
hishiku	red	
hiso	stomach	
hisu	saki monkey	
hodjasu	Taruma	
hodjasu dzudzu	Taruma language	
hoza da'akwi	the rain is coming	
hogiku	white	
hoki	agouti	
horya	now, today; red paint	
horido	pigeon	
horiku	wet	
hoza	rain	
hukabato	true	
hwira	sloth; star	

I

ikwa	to cut	
iroka	thin	
iwa	breast; smoke from fire	
iwa ada	nipple	lit. *head of the breast*
iwana	iguana	

K

kabayo	armadillo
kabi	dance
kamakama	butterfly
kamisa	clothes
kamoti	clay pot
kano	healthy

kanawa	canoe
kaneko	basin
kaniku	bitter
kano	alive
karãgwa	large caiman; river otter
karikari	parrot species
karita	paper
karo	make
karoa	bell yam
kashopara	cutlass
kasu	coati
kato	soul, spirit
kawaro	horse
keki	happy
ki	hear
kida	burn
kio	anteater
Kiaramapa	Waiwai
kibiku	glue
kina	piranha
kirafo	thorn
kiribi	mix
ko	bite; eat (not fruit)
koaba	flea
ko'ãyo	mosquito
koba	arrow, arrow cane
koba sachi	fishing arrow
kobara	pineapple
kobato	wind
kobedo	cassava stick
koi	take out, take off
koki	fox
komata	manioc sieve
kona	fruit
koria	waro parrot
koroba	jump
korona	eddoe
korori	spider
koroto	hole
koshi	pig
koya	calabash
koyaku	blue, green
kururu	rub
kwai	vomit
kwaihĩ	vomit , spill
kwaykwi	labba
kwahĩ	grow
kwakipa	fish hook
kwakidjapa	cassava squeezer
kwakwani	paddle

kwãdeko	older brother
kwãdji	capuchin monkey
kwara	hurt
kwarabato	snot, cold (the illness)
kwaraku	painful, spicy
kwasu	person
kwatoro	elderly person
kway	parch
kwi	poison
kwira	bush deer
kwiriku	medicine man
kwisa	monkey species
kwĩa	juice
kwĩaku	sweet
kwu	there

M

mabo	return
makika	cashew
makashizu	sweet cassava
makoki	go
makwa	tinamou
makwadji	sleep
makwihĩ	stop
makwe	wild cashew
mapaya	papaya
mapipa	fish bait
Mapumapu	Wapishana
maragwi	firefan; stingray
maria	knife
marora	giant armadillo
mayoku	slow
meme	stepmother, maternal aunt
mezekahĩ	love
migwi	put
mĩhĩ	brazil nut
mizo	rattle, toy
mizokahĩ	play
mia	sleepy
mo, mõ	red howler monkey
mo adjoa	red howler monkey's tail
muhwa	a while ago

N

nachi	locust tree
nado	pepper

nado naykonomo	pepper pot
nakani	know, understand
naki	trumpeter bird
nakipa	plant species
nokoda	forest
namaro	sweat
nanuto	weeds
nito	cassava
nohobi	learn, teach
noa	electric eel
no'okwakoro	story
no'wa	long ago
nĩ	sad, bad of character
nukihĩ	fight

O

o	tree	
o acha	tree root	
o asukidjo	mushroom	lit. *tree ear*
o djoko	tree leaf	
o kona	tree fruit	
orokoda	knee	
oshĩwi	one	
o'i	farine	

Õ

õ'õko	younger brother

P

pachori	elbow
paka	cow
pakidako	clay
pakwata	medicine, medicate
pana	daybreak
papa	paternal uncle
pararo	earthworm; vein
parawa	shoulder
pateda	plate
pere	duck
periton	plant species
pidjidji	tick
pikiako	broom
pirisha	long yam
piroka	banana

pishan	cat	
pofo	fan	
porori, horori	millipede	
pwuna	full	
purata	money	

R

ra'wa	be born, give birth	
rikinyo	cricket	
rimirimi	horsefly	
romi	spider monkey	
rora	many	
rorabeto	family	
rukonuko	porridge	

S

sachi	spear	
saki	bag	
sari	tomorrow	
sariri	sorrel	
sa'i	hammock	
sa'ino	village, country	
sa'ino ada	chief, toshao	lit. *head of the village*
sa'o	beetle	
shabana	grandchild	
shaboro	type of drink	
shidjira	common eel	
shipito	mud, swamp	
shipikidjo	mirror	
shirika	inga palm	
shirishiri	intestines	
shonora	heron	
shoshoda	cockroach	
sidja	cotton	
siriri	bat	
sodo	strap	
sõko	navel	
soma	tobacco	
sooso	flower	
sõwa	sõwa	
sũ	penis	
suba	vagina	
sukacho	night	
suko	caterpillar	
sukwi	bee, honey	
suwa	cold	

suwi	throw

T

ta	tucumã palm
taka	awara palm
taki	ask
tania	forehead
tara	bowl, big gourd
tari	bake bread
tatama	spine
tẽba	shoe
tobito	mushroom
tomiku	sour
tora	chigger
toro	bridge, ladder
tu	agouti
tunawo	belt, bracelet

U

u	tight, firm, hard
ushukuwi	touch

W

wã	sun
waba	guava
wabo	manicole palm
wara	put; capybara
wasu	open
wa'adji	talk
wa'e	food
wa'ẽhĩ	eat (fruit)
wa'iri	gourd
wiki	opossum
wikyahã	three
wiwiko	gourd
woda	rope
wodo	urine
woka	blood
wokia, woka	salt
wokawa	raw
wowo	termite
wowoan	waterfall

Y

yakiku	brown
yama	pumpkin
yoma	soft
yoro	slow

Z

zu'i	plant species

APPENDIX B. TRANSCRIBED TEXT

The following is an oral tradition about an abandoned child that is raised by a giant armadillo. The narrative was told by Ms Irene Suttie and recorded by Dr Sérgio Meira on the 31st January 2015. It was translated and transcribed in December 2021. The recording and the corresponding ELAN file will be deposited at ELAR.

(1) *Bakoko   aaa....   djadehĩ*
bakoko   aaa....   djade–hĩ
bird.species   HESI   grate–TAM
'Bakoko was grating.'

(2) *Djaimaki   a'oyhĩ.*
djai –maki   a'oy–hĩ
child–3.POSS   cry–TAM
'His child was crying.'

(3) *Hĩdja   aaa....   fidjihĩ   kwa   djaimakika   a'oyhĩ.*
hĩdja   aaa....   fidji–hĩ   kwa   djai–maki–ka   a'oy–hĩ
NEG   HESI   like–TAM   NEG   child–3.POSS–NS   cry–TAM
'[He] didn't like his child crying.'

(4) *Suwi   meka.*
suwi   meka
throw.away   3.OBJ
'[He] threw it away.'

(5) *Suwi   meka.*
suwi   meka
throw.away   3.OBJ
'[He] threw it away.'

(6) *Habia   marora.*
habiya   marora
EXIST   giant.armadillo
'There was a giant armadillo.'

(7) *Doha   meka   djawãka   djaimakika.*
doha   meka   djawã–ka   djai–maki–ka
bring   3.OBJ   child–NS   child–3.POSS–NS
'[He] took it, the child, his child.'

(8) *Donaha goro   bakĩhĩ   korotoka   djaimakika.*
donaha goro   bakĩ–hĩ   koroto–ka   djai–maki–ka
far   enter   carry–TAM   hole–NS   child–3.POSS–NS
'Far [he] carried his child inside his hole.'

(9) *Hikwa   kwahĩ.*
hikwa   kwa–hĩ
ANPH.LOC   grow–TAM
'There [he] grew.'

(10) *Mmm...   korotoka   kwahĩ   djaimaki   marora*
mmm...   koroto–ka   kwa–hĩ   djai–maki   marora
HESI   hole–NS   grow–TAM   child–3.POSS   giant.armadillo

    maroraka   korotoka   doyaka
    marora–ka   koroto–ka   doya–ka
    giant.armadillo–NS   hole–NS   house–NS

'In the hole grew his child, in the hole of the armadillo, in his house.'

(11) *Hikwa   kwahĩ.*
hikwa   kwa–hĩ
ANPH.LOC   grow–TAM
'[He] grew there.'

(12) *Habia   maraho.*
habia   maraho
EXIST   tree.species
'There was a maraho tree.'

(13) *Ikya   hama   ikya   aaa....   doha.*
ikya   hama   ikya   aaa....   doha
3   1SG   3   HESI   bring
'He, I, he carried it.'

(14) *Habia   ikya   badjihĩ   wa'ekaromakika   djawã.*
habia   ikya   badji–hĩ   wa'e-karo-maki–ka   djawã
EXIST   3   search–TAM   food-make-3.POSS–NS   child
'The child looked for his food.'

(15) *"Mioko   suwidja   hakwa."*
mioko   suwi–dja   hakwa
mother   throw.away–TAM   1SG.OBJ
'"My mother threw me away."'

(16) *"Mioko   suwi   hakwa."*
mioko   suwi   hakwa
mother   throw.away   1SG.OBJ
'"My mother threw me away."'

(17) *"Mioko   mioko   suwi   hakwa."*
mioko   mioko   suwi   hakwa
mother   mother   throw.away   1SG.OBJ
'"My mother, my mother threw me away."'

(18) *"Haka aba, mmm," maroro damẽ'ẽ meka.*
haka aba mmm maroro damẽ'ẽ meka
?? grandma mmm giant.armdadillo answer 3.OBJ
'"Granny, hhh," the armadillo answered him.'

(19) *Marora damẽ'ẽ meka.*
marora damẽ'ẽ meka
giant.armdadillo answer 3.OBJ
'The armadillo answered him.'

(20) *"Mioko mioko suwi hakwa."*
mioko mioko suwi hakwa
mother mother throw.away 1SG.OBJ
'"Mother, mother threw me away."'

(21) *"Mioko mioko suwidja djiko hakwa."*
mioko mioko suwi–dja djiko hakwa
mother mother throw.away–TAM ?? 1SG.OBJ
'"Mother, mother threw me away."'

(22) *"Haka aba, mmm," mmm... marora damẽ'ẽ meka.*
haka aba mmm mmm... marora damẽ'ẽ meka
?? grandma mmm HESI giant.armadillo answer 3.OBJ
'"Granny, hhh," the armadillo answered him.'

(23) *Habia ikya... ikyano kwasu afwikwa.*
habia ikya... ikya–no kwasu afwikwa
habia 3 3–COL people walk
'There were people walking.'

(24) *Badjihĩ ẽ'ẽ karoma kaon.*
badji– hĩ ẽ'ẽ karo–ma kaon
badji–TAM game.animal make–1SG ??
'[He is] looking for his game.'

(25) *Afwikwa.*
afwikwa
walk
'[He is] walking.'

(26) *Aaa... ikya ki chiomakika.*
aaa... ikya ki chio–maki–ka
HESI 3 hear song–3.POSS–NS
'He heard his song.'

(27) *Ikya ki chiomakika.*
ikya ki chio–maki–ka
3 hear song–3.POSS–NS
He heard his song.

(28) *"Mioko mioko   suwidja   djiko   hakwa."*
  mioko   mioko   suwi–dja   djiko   hakwa
  mother mother throw.away–TAM ??   1SG.OBJ
  '"Mother, mother threw me away."'

(29) *"Mioko mioko   suwidja   djiko   hakwa."*
  mioko   mioko   suwi–dja   djiko   hakwa
  mother mother throw.away–dja ??   1SG.OBJ
  '"Mother, mother threw me away."'

(30) *"Haka   aba,   hmmm,"   marora   damẽ'ẽ   meka.*
  haka   aba   hmmm   marora   damẽ'ẽ   meka
  ??   grandma   hmmm   giant.armadillo answer   3.OBJ
  '"Granny, hhh," the giant armadillo answered him.'

(31) *Ga   hama   ki   aaa.... djawãka   chioka.*
  ga   hama   ki   aaa.... djawã–ka   chio–ka
  ??   1SG   hear HESI   child–NS   song–NS
  'I heard the child sing.'

(32) *Gi   ikyano,*
  gi   ikya–no
  say   3–COL
  'They said [to him],'

(33) *"Habia   djaini   kwu."*
  habia   djai–ni   kwu
  EXIST   child–2SG.POSS   there
  '"There is your child."'

(34) *"Habia   djaini   kwu."*
  habia   djai–ni   kwu
  EXIST   child–2SG.POSS   there
  '"There is your child."'

(35) *"Hama   hĩ,"   gi   ikya."*
  hama   hĩ   gi   ikya
  1SG   see   say   3
  '"I find [him]," he said.'

(36) *"Chokohĩ   dja'wamakika."*
  choko–hĩ   dja'wa–maki–ka
  tell–TAM   mother–3.POSS–NS
  '"[I] tell his mother."'

(37) *"Choni choni   hia,"   gi.*
  choni   choni   hia   gi
  take   take   ??   say
  '"[I] take [him] back," [he] said.'

85

(38) *"Choni   hia"*
    choni   hia
    take   ??
    '"I take [him] back."'

(39) *Makoki   ikyaka   ichoni.*
    makoki   ikya–ka   i–choni
    go   3–NS   PRP?–take
    'So he went to take him.'

(40) *?? badjihĩ   ikya   djaimakika.*
    ??   badji–hĩ   ikya   djai–maki–ka
    ??   search–TAM   3   child–3.POSS–NS
    'He was looking for him.'

(41) *Aaa....   habia   mmm...   ikya   chabi   korotamakika.*
    aaa....   habia   mmm...   ikya   chabi   korota–maki–ka
    HESI   EXIST   HESI   3   wait   hole–3.POSS–NS
    'There... they are waiting [for him] at his hole.'

(42) *"Dja'we   bu'a'a."*
    dja'we   bu'a–'a
    mother   run–2SG.IMP
    '"Run to your mother."'

(43) *Nnn... aaa....   abanamakika,   aaa...   choni   meka.*
    nnn... aaa....   abana–maki–ka   aaa...   choni   meka
    HESI   HESI   grandma–3.POSS–NS   HESI   take   3.OBJ
    'To his grandma, [they] take him.'

(44) *Dja'wemaki   choni.*
    dja'we–maki   choni
    mother–3.POSS   take
    'The mother holds [him].'

(45) *A'oyhĩ   chagi*
    a'oy–hĩ   chagi
    cry–TAM   much
    '[He] is crying a lot.'

(46) *Hĩdja   fidjihĩ   kwa   dja'okamaka.*[22]
    hĩdja   fidji–hĩ   kwa   dja'o–kamaka
    NEG   like–TAM NEG   mother–1SG.POSS
    'Because [he] doesn't like my mother.'

---

[22] The speaker accidentally used first-person marking here and repaired it in the following sentence.

(47) *Hĩdja fidjihĩ   kwa   dja'wamakika.*
hĩdja   fidji–hĩ   kwa   dja'wa–maki–ka
NEG   like–TAM   NEG   mother–3.POSS–NS
'Because [he] doesn't like his mother.'

(48) *Hĩdja   mezekahĩ   kwa   ikya.*
hĩdja   mezeka–hĩ   kwa   ikya
NEG   love–TAM   NEG   3
'[He] doesn't love her.'

(49) *Eee...   a'oyhĩ   hĩdja   makwadji   kwa.*
eee...   a'oy–hĩ   hĩdja   makwadji   kwa
HESI   cry–TAM   NEG   sleep   NEG
'[He] is crying, [he] cannot sleep.'

(50) *Hĩdja   hidjihĩ   kwa   sa'imakika .*
hĩdja   hidji–hĩ   kwa   sa'i–maki–ka
NEG   want–TAM   NEG   hammock–3.POSS–NS
'[He] doesn't want his hammock.'

(51) *Gwaka   aaa...   dja'wamaki   hẽda   aaa…,*
gwa–ka   aaa...   dja'wa–maki   hẽda   aaa..
ant–NS   HESI   mother–3.POSS   keep   HESI

    sa'imakika   ikya   djina   makwadji.
    sa'i–maki–ka   ikya   djina   makwadji
    hammock–3.POSS–NS   3   ??   sleep

'His mother put him with the ants, inside his hammock so he can't sleep well.'

(52) *Makwihĩ   a'oyhĩ.*
makwi–hĩ   a'oy–hĩ
stop–TAM   cry–TAM
'[He] stopped crying now.'