Number Expression in Apurinã (Arawák)

Sidney Da Silva Facundes¹, Marília Fernanda Pereira de Freitas² and. Bruna Fernanda Soares de Lima-Padovani³

> ¹ Universidade Federal do Pará, Belém, Pará, Brazil sidi@ufpa.br

> ² Universidade Federal do Pará, Belém, Pará, Brazil mfpf@ufpa.br

³ Universidade do Estado do Pará, Belém, Pará, Brazil bflimapadovani@gmail.com

Abstract. Apurinã (Arawak), spoken along several tributaries of Purus River (Southwest of Amazonas State, Brazil), presents a plural morphological system that marks pronouns and nouns. The language has some free pronominal forms that distinguish singular from plural; additionally, it has bound pronominal forms, with singular/plural distinction made only in the first person for the enclitic forms. In the case of nouns, there are two suffixes that mark plural, -waku (that occurs only with [+human] nouns, as kyky-waku-ry (man-PL-M) 'men'), and -ny (that can occurs both, with [+human] nouns, as in pupyka-ry-ny-ry (indigenous person-M-PL-M) 'indigenous people'; or [-human] nouns, as in aikuny-ry (house-PL-M) 'houses'). The language also presents some quantifiers and numerals that encode number syntactically. The quantifiers are ithu, kaiãu and kuna-kamuny to encode the notion of 'much', puiãu, referring to 'some/few/little', and ykyny to mean 'all/every'. Additionally, there are the following numerals: (h)āty/(h)ātu) 'one' and epi 'two', which combine to derive higher numbers, and the word for 'hand', wakul piu, indicating the numeral five. Thus, the plural marking in the language can be marked in different ways, none of which is, however, required by the grammar. With that in mind, we discuss the extent to which plural marking is, to a great extent, constructed by the speakers in daily language use, according to whether it is contextually important to do so, and raise the question of the relevance of this problem to a computationally implementable grammar of the language.

Keywords: Number, Plural, Apurinã.

1 Initial Considerations

Native languages of Brazilian Amazon have received some special attention in the last three decades partly due to their contribution to language typology, also because most of them are in great state of disappearing once their elderly speakers pass away. Until recently, however, there had not been any work in the area of computational linguistics attempting to implement computational models of aspects of language grammars. This started to change when the authors of this paper came into contact with Jack Rueter and his work on minority languages Northern Europe. Although this paper is not in the field of computational linguistics, nor the authors are experts in this area, it shows an example of how questions raised by attempts to model grammatical properties of a language can contribute to improve language description and analysis. In other words, some of the questions addressed in this paper emerged from issues raised by Rueter in a visit to Brazil, and through continued interaction thereafter through exchange of e-mails.¹

The language whose grammatical properties will be discussed here is Apurinã, a minority language of the Arawak family, spoken by a population of around eight thousand people whose individuals are scattered in dozens of communities, almost all along the margins of several tributaries of the Purus River Basin in Northwestern Brazilian Amazon, as shown in the map in Figure 1:

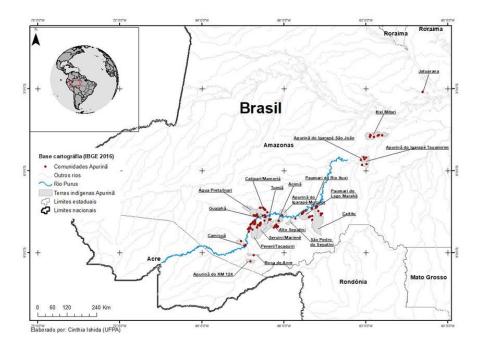


Figure 1: Approximate locations of the Apurina Villages (from the Apurina Language Archives of the Programa de Pós-Graduação em Letras)

There have been a number of studies about Apurinã languages, with some initial vocabulary and word lists being presented by Polak, 1894, Ehrenreich, 1897, Steere, 1901, and Koch-Grünberg, 1919. More recent works are Facundes, 2000, Lima-Padovani, 2016, and Freitas, 2017, among many others. None of these authors, however,

¹ Samples of Rueter's work in Brazil can be found at https://github.com/UniversalDependencies/UD_Apurina-UFPA, and https://github.com/giellalt/lang-apu.

have presented an adequate description of number marking in the language. Such that when it was unclear how the computational grammar for the language could implement an analysis able to determine if a nominal form was singular or plural in a text. The absence of a comprehensive study of plural marking in Apurinã is due, in part, to the fact that such marking itself is not strictly required in the grammar of the language. That is, both of the following forms are perfectly grammatical in the language: epi sytuwaku-ru (two female-PL-F) and epi sytu (two women) mean 'two women'. This fact, combined with the importance of a thorough examination of the plural/singular uses in texts, the need of sociolinguistic stratification of the data, have inhibited further incursions into this issue in Apurinã. Therefore, the following issues are addressed in this paper:

- A) What formal resources are used in number marking (plural / singular) in Apurinã?
- B) What nouns can be marked in the plural in Apurinã, and what is the grammatical or discourse-pragmatic status of these markings, as well as possible semantic correlates?

In order to examine the Apurinã language, we made use of the Apurinã Language Digital Database (Fig. 2), which includes over three thousand lexical items, morphologically analyzed words, dialect variations, and texts transcriptions, which have been compiled by a group of researchers and students at the Universidade Federal do Pará, in Brazil, starting in 1990². Currently, this database makes use of the Fieldwork Language Explorer (FLEx) software, made by the SIL international (https://software.sil.org/fieldworks/), and freely available. See Butler and van Volkingurg (2007a, 2007b) and Rogers (2010), for reviews about FLEx.

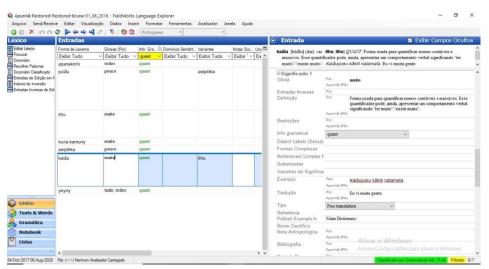


Figure 2: Sample of a register of the Apurinã Language Digital Database

² The terms under which the Apurinã Digital Database will be publically available are still being discussed with the Apurinã communities and the members of the research team.

2 The grammar of number in Apurinã

In this section, we describe how the notion of number is marked in the Apurinã language, both morphologically and syntactically, on the basis the both elicited and, primarily, text data. Most of the examples presented in the following sections have been attested in texts.

2.1 Number expression in pronominal forms

The plural versus singular distinction in Apurinã is encoded as part of the free and bound pronominal forms, and in the noun suffixes -waku and -ny. The same free pronouns can function as subject, object (of the verb or of the postposition) or possessor. As shown in Table 1, the number of distinction (singular / plural) is present in all grammatical persons for the free pronouns:

Table 1: Free pronouns as subject / object / possessor³

	Singular	Plural
1	nuta	atha / athe
2	pitha / pithe	(h)îthe / (h)îtha
3M	ywa	ynawa / ynuwa, nynawa / nynuwa
3F	uwa	

The pronominal bound forms differ between subject and object in the verb, marking the possessor noun, and present phonologically conditioned variants. As shown in Table 2, the subject / possessor markers also distinguish singular from plural in all persons. Leaving aside some cases of dialectal variation, the object markers distinguish between singular and plural only in the first and second persons (Table 3):

Table 2: Pronoun forms in subject/object/possessor function

	Singular	Plural
1	ny- / n- / nhi-	a- / ã-
2	p y - / p- / pi- / p i -	h y-/h-
3M	y- / Ø- / i-	yna / øna; ina
3F	u- / ũ-	u na

Table 3: Pronoun forms in object function

³ The pronouns that appear in more than one form in this table are cases of dialect variation.

	Singular	Plural
1	-nu	-wa
2	-i	-i / -na
3M	-ry	-ry
3F	- ru	-ru

2.2 The morphological expression of number

With regard to nouns in particular, there are two specialized suffixes in the plural marking, -waku and -ny. In contrast, singular forms do not receive any formal marking. In general, -waku is used in nouns related to humans, as in pupỹkary-waku-ry 'Apurinãs'; -ny is also used with nouns unrelated to humans, like aiku-ny-ry 'houses'. Note that both -waku and -ny are always accompanied by the gender marker, -ry 'MASCULINE' or -ru 'FEMININE'. (1) to (3) illustrate the use of the nominal plural markers:⁴

(1) <u>Sytu-waku-ru,</u> kiripa u-kama-ã-ne? woman - PL -F what 3PL.F- do- PROGR-3PL.F 'And the women, what do they do?'

(2) Kitxakapirîka <u>atukyry-</u>w<u>aku-ri-nhi</u> in the old days 2sg-grandfather.in-PL-M-AFET

sãpira - ta-nu kuna watxa atuku i-txa.
count - VBLZ-1SG.O not today as 3SG.M -be
'In the old days your grandparents told me it was not like today.'

(3) ... *kutxi karywa xipuka -ry <u>pupỹkary-waku-ry.</u>*because non.Indian acabar- 3SG.M.O Indian-PL-M
'(Preceding context: Today there are few Indians) because the White destroyed the Indians.'

There are at least two cases that can be considered extensions of these uses: the first includes the nouns for animals, which are pluralized by adding -waku, as in (4). This use of plural with nouns for animals is one of the ways to derive names for clans and other subgroups Apurinã the society:

(4) Kamỹyry-**waku** -ry kuwary - ny -ry- taka umanata-wa apaka.

⁴ SG = natural; PL = plural; F = female; M = male; O = object; PROGR = progressive; AFET = affected; VBLZ = verbalizer; POSSD = owned; N.POSSD = not possessed; PRED = predicate; DISTR = distributive; GER = gerund; NC = classificatory name; IPFTV = imperfective; RESTR = restrictive; LOC = locative; ATRIB. INTENS = enhancing attribute; FOC = focus; AUX = auxiliary; REFL = reflective; PRIV = private; Frust = frustrativo; ENF = emphatically.

```
macaw-PL-M kuwary - PL -M - part fazer.maldade- 1PL.O also 'The people of the macaw clan did not like us.'
```

The second case includes at least some nouns for farm animals, as in the example below:

```
(5) Ny-pyra-waku-ry pathery-waku-ry.

3SG-animal-PL.M chicken-PL-M

'I have many chickens' Lit: 'My chicken (domestic) animals'
```

Therefore, the use of -waku distinguishes between nouns for human beings and related concepts, and other nouns in Apurinã.

The plural suffix -ny occurs with some nouns denoting humans, as shown in (6) and (7), and, at least in elicited data, it also occurs with non-human nouns, as in (8). However, some observations about -ny are in order: after examining a sample of 27 texts, Freitas (2017) noticed that there were 33 ocurrences of -waku and 20 of -ny, where - waku appears in nouns for people, domestic animal, game meat or names of clans, whereas -ny appeared only in verbal form (pa-puxuku-ny-ry 'they are happy'), in pupykary-ny, and some nouns where -ny semi-grammaticalized, such as amary-ny 'children', and kuwary-ny-ry 'part of; name of one of the Apurina moiety'. With that in mind, we can note that in (7) pupykary-ny is used to refer to the Apurina people as a nation, that is, to highlight a collective with its traditional customs, in (a) to contrast it with other nations, such a Paumari or Non-Indigenous people, and, in (b), to contrast with the nation of other animals, in cultural coherence with the Apurina cosmology⁵. The fact that examples such as (8), where -ny is used with a noun for an object are attested only in elicited data suggests that such a noun semantic is not generally morphologically marked for plural in spontaneous speech.

(6) Wera-ĩkara-ry kitxity-ny-ry txa-ka-ta-ry Awãĩ. there-again-M patauá.fruit-PL-M be-indeed-VBLZ-3M.O Awãĩ 'That one is in fact the Patauá person, Awãi.'

```
(7) a. <u>Pupỹkary-ny</u> awa y-kyynyry-te-na .
Indian-PL have 3PL.M-xingané-POSSD-3PL.M
y-serēka-na mitxi-ry
3PL.M-dance-3PL.M be.first-3SG.MO
'The Apurinã have their party; first they dance (traditional dances).'
```

b. *ymaruta txa-ry ynawa, pup<u>ỹ</u>kary-ny-ry, asike-tikinhi-t-inhi-ry* know AUX-3M.O they Apurinã-PL-3M track-behind-VBLZ-GER-3M.O *î-iãkynỹ-ã.*3M-footprint-LOC

⁵ In this sentence the narrator is talking from the perspective of a tapir who had kidnapped a woman from the village, made her his spouse and now is being hunted by woman's relatives.

In Hämäläinen, M., Partanen, N., Alnajjar, K. (eds.) Multilingual Facilitation (2021), pages 31–42. — CC BY 4.0

(8) <u>Aiku-ny-ry</u> home-PL-M 'houses'

2.3 The syntax of number expression

In addition to the plural marks, there is in Apurina another strategy to encode the existence of a certain number of elements: the use of certain quantifiers, *i.e.* numerals, and indeterminate quantifiers. Syntactically, they function as nominal modifiers.

The numerical Apurinã system contains two primary forms (h)āty 'one' and epi (or ipi) 'two', and the remaining numbers are derived from the combination of these two primary forms: hāty epi 'three', epi epi 'four epi epi hāty 'five'. Sometimes the pakyny form, meaning 'plus', comes with the numerals, as in epi hāty pakyny, but its use is not mandatory. Although the system permits continued counting, for as long the memory allows computing the number of repetitions of, as in epi epi epi epi hāty 'eleven', in over 30 years of contact with Apurinã, we found only a young Apurinã that seemed comfortable enough to count without hesitation over five using the language system. Normally, they count up to three. Even counting up to four involve some hesitation. Also, five can also be computed as hāty waku / piu 'a hand'.

As for the phrase structure that contains a numeral, the numeral always precedes the modified noun, and the modified noun does not usually receive the morphological plural mark in texts, although this may occur in elicited data, and illustrated in (9):

In semantic terms, there are some quantifiers in the language to mean 'a lot': *ithu*, *kaiãu* and kuna-kamuny. As shows (10), *ithu* can be used with nouns referring to individually distinguishable elements (*i.e.* each 'parent'), thus as counting noun, as well as with nouns whose referents are not made up of individually distinguishable elements:

(10) Ithu a. Nuta awa ithu ny-nyrymane 1sg have much 1sg-relatives.of 'I have many relatives'

b. Nuta awa -ry ithu ikira 1sg have-3SG.M.O much salt

Thave a lot of salt.

c. Sytu ãta-ru ithu ipara woman drink-3SG.F.O much water

'The woman drank a lot of water'

The data (11-12) show a similar behavior for the quantifiers *kaiãu* and *Kuna-kamury*. The difference between some of these quantifiers is that they are associated with distinct language varieties, as described later:

(11) *Kaiãu*

a. *Kaiãu-puku a-parika awa*very - DISTR 1PL- trabalho.de exist
'It's our job'

b. *Kaiãu-puku* kãkyty n-atama-ta very - DISTR people 1SG- see -VBLZ

'I saw a lot of people'

c. Ximaky kaiãu-puku ny-mãỹka fish much-distr 1sg- catch

'I got a lot of fish'

(12) Kuna-kamuny

a. Nuta itimata-ru kanawa kuna-kamuny 1sg see- 3SG. F.O canoe much

'I saw a lot of canoe'

b. *Nuta itimata* **kuna-kamuny**-ry kaniti
1sg see much - M people
'I saw a lot of people'

c. *Kuna-kamuny - panhi -ry* kãkyty nuta akirita-ka-inhi very-IPFTV -M person 1SG call- PRED - GER

'I invited a lot of people'

In addition, there is also the quantifier $pui\tilde{a}u$ that corresponds to 'little' as in the following examples (13). In the data below, we call special attention for the use of the classificatory noun -pe in katarukyry-pe (flour-NC) in (13d), for this morpheme allows to derive the measure term for concepts not made up of perceptually important individual parts. Thus, -pe is a way that some mass nouns can be quantified. As in the other quantifiers described so far, the modified noun does not receive the plural mark.

(13) *Puiãu*

a. *Nuta* **puiãũ**-puku nhipuku-ry py-syka 1SG little-DISTR food-N.POSSD 2SG-give

'Give me some food'

b. *Watxa puiaũ -nuka pupỹkary awa-ry kutxi* today little- RESTR Apurinã exist - 3sg.mo because

karywa xipuka pupỹkary-waku -ry. not.indian finish-3SG.M.O Apurinã- PL -M

'Today there are (only) some Indian because the Branco people destroyed the Indians.'

c. *Nuta* awa -ry puiãu -ka iukira
1SG have -3SG. M.O little -pred salt
'I have little salt'

d. *Atha* awa -ry puiãu -ka katarukyry-pe
1PL have-3SG.M little-PRED flour-NC

We have little flour '

Finally, there is the universal quantifier ykyny- 'Everything, all'. As shown in (14), it can receive -puku, the distributed plural suffix modifying the nou. As in other cases, the noun modified by the universal quantifier does not receive the plural mark:

(14) ykyny

a. Ywa awa -ry ykyny -puku aapuku-taxi 3SG.M have-3SG.M.O all-dISTR house.of-N.POSSD -L.OC

'He has everything in his house'

- c. *Kiripa py-kama ykyny uty?*what 2SG-make all day
 'What do you do in the day-to-day?'
- d. *Ywasaaky* **ykyny** -mane sãkira-wa-ta pupỹkary sãkire.
 At.that.time all-body.of speak-REFL-VBLZ Indian language 'At that time, everyone spoke the Apurinã language.'

The notion 'no one' is expressed as follows:

(15) *M-inha-ka-ti-ma* kuna aiata watxa
PRIV-be- PRED-3SG.MO-FRUST not hunt today
'Nobody is going to hunt today.'

In addition to these quantifiers, the notion 'a little bit, a little' can be expressed by the descriptive verb *axi* 'being little, which in certain contexts can acquire a sense of 'have / be a little bit, a little', as in the following example:

(16) Axi-ka-piti awa -ry kyky
be.little-PRED-INTS live- 3SG.M.O man
'The man lived very little' Lit :: 'It was little (that) the man lived'

Thus, the preliminary analysis of the quantifiers in Apurinã, on the basis of the examples presented so far, suggest that they may occur when nouns are perceived as made up of individual parts (*ithu nyrymane* 'many relatives'; *ximaky kaia-puku* 'very many fish', *nuka puiaũ pupykary* 'only little Indian'), and also when noun referents are not perceived as made up of individual parts (*ithu iparãa*" 'plenty of water'; *puiãu-ka iukyra* little salt'). There are also cases of certain nouns referring to entities without individual parts, where a classificatory noun is used to refer to a container, as seen in the example (12d).

2.4 Quantifiers and dialectal differences

As described above, in Apurinã, there are five different forms associated with the notion of quantification, *ithu / ithu hãitery, kaiãu, kuna-kamuny*, encoding the notion of 'very' and *puiãu* and *axipitika* encoding the notion of 'little'. Such forms are dialect variants which seem not to co-occur in the same variety of Apurinã language, i.e. when in a given community uses one of these forms, other forms are not employed. Therefore, the diversity of forms that denote the notion of quantification is related to the wide geographic dispersion of the communities, with large distances from each other. This situation favors the lexical variation in the language as a whole, making certain variants be used in some communities, but not in others (see LIMA-PADOVANI, 2016).

There is, however, one exception to the variation pattern just mentioned: in one community, in addition to *ithu*, it was also identified the form kuna kamuny to denote the notion of 'a lot'. The latter form was, nonetheless, provided by older speakers of the community. Once this data was checked with speakers of other communities, we noticed that the latter recognized the second form, but said they did not use it. In conclusion, *kuna-kamuny*, at the current stage of Apurinã is associated not only to geographic variation (as occurs only in a community), but also to generational variation, since it was attested only in the speech of older individuals. According to Freitas (2017), given the low frequency of use *kuna-kamuny* we could perhaps conclude that this form is falling out of use in the Apurinã language. Note the geographic distribution of these quantifiers in Table 4 below:

Table 4: Geographic Distribution of Quantiers Apurinã

	K	Km	K	Tu-	Aci-	Sepa	Sã	Japi	Terri	Ter	Vila
Quant	m		m	miã	mã	-tini	o	-im	-nha	-ra	Nova
			45								

	12	137					Jo-			No-	
	4						sé			va	
ithu	X	X	X	X	X	X					
kaiãu							X	X	X	X	X
puiã				X	X	X	X		X	X	X
Kuna-					X						
kamuny											
axipi- tika								X			
tika											

The first, 'ithu', above has greater distribution, being identified in six different communities: Km 124, Km 137, Km 45, located in the Middle Upper Purus region; and in the Tumiã above and Sepatini, located in the Middle Purus region. In addition, members of other communities recognize this form, but do not use it.

Item three, *kaiãu*, follows a similar distribution, occurring in four different places, as in Japiim and Terrinha communities in the Middle Purus and the newly founded communities, and in Vila Nova in the Lower Purus.

Item three, $pui\tilde{a}$, encodes the notion of 'little'; it also has a wide distribution, occurring in eight different communities, which are distributed in the regions of Middle and Lower Purus.

In contrast, the last two items of the table were each attested in only one community. Item four was attested only in the Acimã village, being used by older speakers. Item five, *axipitika*, also encodes the notion of 'little' and occurs only in the Japiim community, in the Middle Purus.

Although these forms are actively used in some communities and not in others, following a geographic distribution, most of the time, some speakers recognize and accept more than one of such forms to designate the same quantifier. This is due to the constant migration among communities, and, more recently, intense contact among its members. The geographical space shows the particularity of each community, showing the variety that the language assumes from one region to another, as a way of characterizing cultural diversity.

3 Concluding Remarks

We have shown that, in Apurinã, nouns can be marked in the plural by one of the two plural marker suffixes: -waku, occurs only with terms related or associated with human beings, while the other, -ny, occur with less frequency with human related nouns given it some more general/abstract or spiritual meaning and, at least in elicited data, with objects. Both plural suffixes, however, are not grammatically required. In addition, although the syntactic modifiers can be used to mark the plurality of a modified now, they also are not grammatically required for a noun to be interpreted as plural. We could then say that the distinction between singular and plural is not strictly bound to the grammar of the language, and more to pragmatic forces, such that it is the pragmatic context (rather than the grammar) that will determine whether to use it or not plural

marking. Which pragmatic forces are these, however, require investigation. Thus, the plural marking in Apurinã is likely to pose the same sort of problem for a computational implementation of a grammar, which is how to build contexts into it.

Finally, in a typological note, when we take into consideration the typological survey presented in Corbett (2001), we can now examine the place of the Apurina number expression system in contrast to other languages of the world. In terms of the number marking, despite an apparent opposition between plural and singular, which would manifest itself in the presence or absence of plural suffixes, in fact, it would be more appropriate to describe the opposition as being between plural and generic. This is justified because the plural nouns admit interpretation in the plural only, while nouns not marked for the plural may be interpreted both as singular or plural. That is, while *nyry-mane-waku-ry* entails interpretation 'more than one relative' *nyrymane* entails 'one or more of a relative'.

References

- Butler, Lynnika, and Heather Van Volkinburg. "Fieldworks Language Explorer (FLEx)." Technology Review 1.1 (2007): 1.
- Butler, Lynnika, and Heather van Volkinburg. "Review of Fieldworks Language Explorer (FLEx)." Language Documentation & Conservation 1.1 (2007): 100-106.
- 3. Corbett, Greville g. Number. New York: Cambridge University Press, 2004
- 4. Ehrenreich, P. Beiträge zur Völkerkunde Brasiliens. Veröffentlichungen aus dem Königlichen Museum für Völkerkunde II. Bd., 1/2. Heft. II. Über einige Völker am Rio Purus (Amazonas). SS. 48-72. Berlin, 1891.
- Facundes, Sidney da Silva. The Language Of The People Of Apurina Brazil (Maipure / Arawak). New York, Buffalo: Faculty of the Graduate School of State University of New York at Buffalo (Doctoral Thesis), 2000.
- FREITAS, Marília Fernanda Pereira de. A Posse em Apurinã: descrição de construções atributivas e predicativas em comparação com outras línguas Aruák. Tese de Doutorado. Universidade Federal do Pará. Belém, 2017.
- 7. KÖCH-GRÜNBERG, von Theodor. Ein B eitrag zur Sprache der Ipuriná- Indianer, Rio Purus, Brasilien. Journal de la Societé des Americanistes, NS Paris, 1919.
- LIMA-PADOVANI, Bruna Fernanda. Levantamento Sociolinguístico do Léxico da Língua Apurinã e sua contribuição para o conhecimento da cultura e história Apurinã (Aruák). Dissertação de Mestrado. Universidade Federal do Pará, Belém (2016).
- 9. POLAK, Jacob. The ER Grammar and the Vocabulary of the Ipuriná Language. Vocabulary Publication Fund, No. 1. London: Kegan Paul, Trench, Trubner, and CO., Charing Cross Road, 1894.
- 10. Rogers, Chris. "Review of fieldworks language explorer (flex) 3.0." Language Documentation & Conservation 4 (2010): 78-84.
- 11. STEERE, JB Narrative of a Visit to Indian Tribes of the Purus River, Brazil. Report of the United States National Museum; 1903 pp.378-380.