

The rise of number agreement in Nheengatu A emergência da concordância de número em Nheengatu

Aline da Cruz

Universidade Federal de Goiás. Goiânia, Goiás, Brasil

Abstract: Number agreement systems often present traces of older elements common to different languages of the same family; thus, their emergence is difficult to reconstruct. One possible origin of such systems is the grammaticalization of plural words into bound morphemes, which, as a result of a long process, develop into agreement markers and may become obligatory. Various investigations have provided evidence for this hypothesis. However, the complete process of change from a system with no number as grammatical category into a number agreement system in a single language has not been documented. This paper analyses documents covering different stages of the development of the Nheengatu language from Tupinambá in order to observe how the number agreement system emerged in modern Nheengatu. By doing so, this paper supports the idea that grammaticalization may have occurred rapidly in intense contact situations.

Keywords: Number agreement. Grammaticalization. Língua geral (Nheengatu). Tupi-Guarani.

Resumo: Sistema de concordância de número frequentemente apresentam cognatos em diferentes línguas da mesma família, o que pode indicar que são muito antigos. Por isso, é muito difícil reconstruir a emergência desses sistemas. Uma possível origem para sistemas de concordância de número é a gramaticalização de palavras independentes com função de 'plural' em morfemas presos, que, como resultado de um longo processo, transformam-se em marcas de concordância, podendo inclusive se tornar obrigatórios. Várias pesquisas proveem evidências para essa hipótese. No entanto, o processo completo de transformação de uma língua sem a categoria gramatical de número em uma língua que apresenta concordância de número não foi ainda documentado. Este trabalho analisa documentos de diversas etapas de desenvolvimento do Tupinambá até o Nheengatu com o objetivo de observar como o sistema de concordância de número teria emergido em Nheengatu. Ao fazer isso, este trabalho dá suporte à ideia de que gramaticalização pode ocorrer mais rapidamente em situações de intenso contato linguístico.

Palavras-chave: Concordância de número. Gramaticalização. Língua geral (Nheengatu). Tupi-Guarani.

CRUZ, Aline da. The rise of number agreement in Nheengatu. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, v.10, n.2, p. 419-439 maio a ago, 2015. DOI: <http://dx.doi.org/10.1590/1981-81222015000200011>.

Autor para correspondência: Aline da Cruz. Universidade Federal de Goiás, Faculdade de Letras, Campus II – Samambaia, Vila Itatiaia. Goiânia, GO. CEP: 74690-900. E-mail: aline.da.cruz@live.com.

Recebido em 17/02/2014

Aprovado em 06/05/2015



INTRODUCTION

Studies on grammaticalization have predicted that the existence of plural words¹ in a language can lead into the development of bound morphology, and eventually into systems of great complexity involving agreement (Corbett, 2004, p. 266). However, to the best of our knowledge, there are few historical sources that can be used to describe the process from the moment a language has no number as a grammatical category, to the one when a number agreement system is in place². Documents registering the transformation of Tupinambá into Nheengatu can help fill this hiatus: sources show that there is no system expressing number as a grammatical category in Tupinambá, while Cruz (2011) shows that number agreement is present in Nheengatu. This is illustrated in examples (1) and (2) below. In example (1) from Tupinambá, the verb *juca* 'kill' is inflected by a third person marker that can be interpreted as singular or plural. Since there is only one mark for third person, the difference between singular or plural is context-dependent. In the pair of examples in (2) from Nheengatu, the cognate verb *juka* 'kill' occurs with different inflectional markers to distinguish between third person singular (2a) and third person plural (2b).

- (1) TUPINAMBÁ <ojucâ>³
 o-juca
 3.A-kill
 'He kills' / 'They kill' (Anchieta, 1990, p. 17v)
- (2) NHEENGATU (a) *u-juka* (b) *ta-juka*
 3SG.A-kill 3PL.A-kill
 'He kills' 'They kill'

The aim of this paper is to describe the grammatical changes that have contributed to the rise of the number agreement system in Nheengatu. The paper is organized as follows: the next section gives a brief introduction of the history of the social development of Tupinambá into Nheengatu. The following section indicates the sources used in this study. The fourth section of this article focuses on the beginning of the process of change, by observing the grammaticalization of a plural word in 18th century Brasília. The fifth section deals with the restructuring of the pronominal system in the 19th century Nheengatu. It shows the emergence of a third person plural pronoun, which provides the morphological material for the rise of number agreement. The sixth section shows the grammaticalization of this pronoun into prefixes of third person plural, used for expressing number agreement. Additionally, this section shows that the new-born third person plural prefixes are used as one of the criteria for distinguishing two varieties of currently spoken Nheengatu: the Negro-Içana variety spoken by Baré and Baniwa peoples, and the Xié variety spoken by Warekena people. The seventh section draws a hypothesis on how the contact with other languages can have triggered or at least accelerated the arising of a number agreement system in Nheengatu. The concluding section summarizes the process.

¹ The term 'plural word' is used in this work as defined by Dryer (1989).

² Examples to the development of number as an instance of grammaticalization is discussed by Lehmann (2002, p. 50-53).

³ In this work examples from old documents are presented as follows: in the first line the data is provided in its original spelling, indicated by angular brackets; in the second line the data is provided with adaptation of spelling and indication of morphemes; in the third line the gloss is provided; and in the final line, the original translation is provided.

OVERVIEW OF THE SOCIAL HISTORY OF NHEENGATU⁴

In the 16th century, Portuguese colonizers landed on the northern coast of Brazil, where they established contact with Tupinambá people. Some colonizers fathered children with local Tupinambá women. These children, their Portuguese fathers, and their indigenous families formed a new Tupinambá-speaking society: the *caboclo*⁵ society.

In 1616, the first Portuguese settlement in Amazonia was created with the foundation of Belém do Pará. It was from there that *caboclo* people, Portuguese colonizers and Jesuit missionaries started to occupy larger portions of the Amazon. When the *caboclo* people moved deeper into the Amazon region, they came into contact with speakers of many other indigenous languages. Due to this migration, Tupinambá became increasingly used for communication in the Amazon: the *caboclo* people used it to communicate among themselves, the colonizers used it to speak with indigenous groups, and indigenous groups used it to communicate with each other if they had no language in common (Rodrigues, 1996a; Bessa Freire, 2004). As Tupinambá was spread deeper into the Amazon region, it underwent changes and became known as *língua geral Brasileira* (Oliveira, 2008; Lee, 2005).

In the 18th century, *Brasílica* expanded along different rivers of the Amazon and evolved into different varieties (Bessa Freire, 2004). In the 19th century *Brasílica* was called Nheengatu⁶, a name still used by the remaining speakers. Based on the analysis of 19th century documents, Felix (2002) recognized eleven varieties of the language: Solimões, Juruá, Upper Rio Negro, Low Rio Negro, Tapajós, Low Tocantins, Monte Alegre, Ereré, Santarém, Mudurukú and Mawé. Despite the widespread use of Nheengatu's varieties, they had been mostly replaced by Portuguese by the early 19th century.

Even though Nheengatu was disappearing in most parts of Amazonia in the 19th century, the language continued expanding in the Upper Rio Negro. There, Nheengatu speakers came into contact with other languages, such as Baré, Baniwa, and Warekena from the Arawakan family; Tukano from the Tukanoan family; Dâw from the Nadahup family (Cabalzar; Ricardo, 2006).

The outcome of this linguistic contact was varied: while the Tukano, the Dâw and the Baniwa from the Upper Içana have kept their original languages, the Baniwa from the Lower Rio Içana, the Warekena from the Rio Xié and the Baré from Rio Negro have switched from their original languages to Nheengatu. The beginning of the language shift process was registered in the early 20th century by Koch Grünberg:

Diese Aufnahme stellt sich insofern schwieriger, als die Leute fast kein Spanisch, sondern nur Uerekéna, Baniwa und Lingoa geral verstehen. Da helfe ich mit meinem Baniwa-Vokabular, oder der Patrão verdolmetscht es ihnen aus dem Spanischen in die Lingoa geral.

[This undertaking [construction of Uerekéna vocabulary] is more difficult [compared to the construction of Baniwa and Baré vocabularies], since these people [the Uerekéna-Indians Koch is traveling with] almost do not understand any Spanish, but only Uerekéna, Baniwa and Língua geral. Because of that, I help myself with my Baniwa vocabulary, or the Patrão translates for them from Spanish to Língua geral] (letter from July 8, 1903 from the Upper Rio Negro, in Von den Steinen 1904, p. 294; translation by Johan Vincent de Vries).

The bilingualism between Nheengatu and Warekena, and the bilingualism between Nheengatu and Baniwa – described by Koch Grünberg in 1903 –, did not remain. During the 20th century, the Warekena from the Rio Xié

⁴ A reduced version of this overview of the social history of Nheengatu was published in Cruz (2014).

⁵ Since there is no proper English translation, the Brazilian Portuguese term *caboclo* is used to refer to this society.

⁶ According to Rodrigues (1996a), the term Nheengatu was created by Couto de Magalhães (1876). It is a compound formed by the words *nheen* 'speak' and *katu* 'be good', which means the 'good language'.

switched from their original language to Nheengatu. The same occurred with the Baniwa from the Lower Içana and the Baré from Rio Negro, who also switched to Nheengatu⁷.

After these language shifts from Arawakan languages to Nheengatu, there was an intensification of the use of Portuguese in the region. As a result, the Baré people in the Lower and Middle Rio Negro, who had previously switched from Baré to Nheengatu, switched from Nheengatu to Portuguese. Nheengatu was more resistant to Portuguese in the Upper Rio Negro region, where Nheengatu remains the first language of the Baré, the Warekena, and the Baniwa of Lower Içana. These three groups have different degrees of bilingualism: while Baré and Baniwa communities tend to be bilingual in Nheengatu and Portuguese, the Warekena communities are predominantly monolingual in Nheengatu.

SOURCES FOR THIS STUDY

Even though the development of Tupinambá into Nheengatu is a continuous process, the analysis of the emergence of the number agreement system in the modern language can be simplified by the recognition of different stages of development. The periodization proposed here is based on the external history of Nheengatu, as studied by Bessa Freire (2004) and summarized in the previous section.

Four stages of development can be proposed. The first stage is Tupinambá, as it was spoken in the 16th and 17th century and described by Anchieta (1990). It corresponds to the initial stage, before its use in the colonization as *língua geral*. The sources used for the analysis of this stage are the grammar written by Anchieta (1990), and a collection of letters written by Diogo da Costa to Pedro Poti in the 17th century and edited by Sampaio (1906). Additional data is provided by Lemos Barbosa (1956) and Rodrigues (1996a).

The second stage corresponds to the period of expansion of Tupinambá as a *língua geral*, replacing other languages. To highlight the changes in the language and its use in colonization, it is called *Brasílica* (from “língua geral Brasília”). Data provided by Lemos Barbosa (1956) is used for observing this stage of development.

The third documented stage is 19th century Nheengatu. In this period there were different varieties of Nheengatu spoken in the Amazon. Since the focus of this paper is on the emergence of number agreement in the Nheengatu spoken in the Rio Negro, the only source used for this stage is the collection of texts gathered by Barbosa Rodrigues (1890) in the Rio Negro region.

The last stage, 21st century Nheengatu, consists of two varieties of currently spoken Nheengatu: the Negro-Içana variety spoken in Upper Rio Negro and Lower Içana rivers by Baré and Baniwa groups; and the Xié variety spoken in Xié river by Warekena people. The data from the Xié-variety was collected during fieldworks conducted in 2007 and 2014, while the data from the Negro-Içana variety was collected in 2007, 2008, 2009 and 2013. Table 1 summarizes the historical stages of development of Tupinambá into Nheengatu, indicating the name that the language received in each stage and the sources used for the analysis of each stage.

THE GRAMMATICALIZATION OF A PLURAL WORD IN THE LÍNGUA GERAL BRASÍLICA

Languages of the Tupi-Guarani family are well known for not having morphology to express number distinctions (Jensen, 1999, pp. 151-152). Tupinambá, considered one of the most conservative Tupi-Guarani languages, represents

⁷ According to UNESCO, Baré is an extinct language. Warekena is highly endangered, since it is spoken by only approximately 20 speakers (Moseley, 2010). Baniwa is spoken in Upper Rio Içana by approximately 5000 speakers.

Table 1. Sources used in this study.

Stage	Name of the language	Sources used
16 th century	Tupinambá	Anchieta (1990) Sampaio (1906) Rodrigues (2010) Lemos Barbosa (1956)
17 th and 18 th century	Brasílica	Lemos Barbosa (1956)
19 th century	Nheengatu (from Rio Negro)	Barbosa Rodrigues (1890)
21 st century	Nheengatu (from Upper Rio Negro and Lower Içana) Nheengatu (from Xié)	Cruz (2011) and fieldwork from 2007 – 2014

an initial stage in which no number distinctions were grammatically expressed: “plural can be understood by the context” (Anchieta, 1990, p. 8v)⁸. In typological analysis of the number category (Corbett, 2004), Tupinambá can be classified as a language in which there is ‘general number’. In this type of language, the meaning of a noun can be expressed without reference to number. For instance, a noun such as *oka* expresses ‘house’, but the use of this form does not commit the speaker to refer to a unique house: there can be one or more houses. In order to refer explicitly to one or to more than one entity, the quantifiers *ojepé* ‘one’ and *c-eta* ‘many’ are available. Anchieta (1990, p. 8) explicitly mentions that *eta* is used only to indicate a great number and translates it as “todos, tantos, quantos, muitos” [all, many, lots]⁹.

The word *ojepé* is translated by Anchieta (1990, p. 42) as both ‘one’ and ‘once’, as shown in (3), (4) and (5). In example (3) the quantifier *ojepé* occurs with a verb, and it is translated as ‘once’. Notice that in example (4), the word *ojepé* ‘one’ occurs before the noun, whereas in example (5) it occurs postposed to the noun.

(3) TUPINAMBÁ <ojepêaço>
ojepé a-ço
 one 1SG.I-go
 ‘once I went’ (Anchieta, 1990, p. 10v)

(4) TUPINAMBÁ <ojepê abá>
ojepê abá
 one man
 ‘one man’ (Anchieta, 1990, p. 9v)

(5) TUPINAMBÁ <cunhã ôyepê>
kuñã ojepê
 woman one
 ‘one woman’ (Anchieta, 1990, p. 9v)

Although a more thorough analysis is needed, the translation given by Anchieta, and its occurrence before and after a noun seems to indicate that the word *ojepé* does not function as a modifier in the NP, but rather as an adverbial quantifier¹⁰. Evidence supporting this hypothesis comes from the analysis of genetically related languages. In several descriptions of

⁸ In the original words: “O plural se entende pelo que se trata.”

⁹ To the best of our knowledge, there is no study that investigates what leads a speaker to choose implicit or explicit number.

¹⁰ Nheengatu differs from Tupinambá in the categorization of numerals. While numerals in Tupinambá are adverbs; in Nheengatu, numerals have turn into modifiers that occur inside NPs (Cruz, 2013).

languages of the Tupi-Guarani branch of the Tupian family, numerals share their properties with adverbs. For instance, in Kamaiurá numerals can be combined with nominalizers that are specifically used with adverbs (Seki, 1990, p. 78). In Apyãwa (Tapirapé), numerals, as adverbs, do not occur as nominal modifiers, and may occur in any position of the clause (Praça, 2007, p. 149). In Ava, numerals have the same properties as words with adverbial function (Borges, 2006, p. 181). The adverbial properties of numerals can also be found in languages of other branches of the Tupian family. In Karitiana, from the Ariken branch of Tupian family, numerals are analysed as adverbial quantifiers (Müller; Sanchez-Mendes, 2008). In that language, numerals may take scope over NPs or VPs. The fact that numerals occur with NPs and VPs is presented as evidence that they do not function as modifiers inside an NP, but rather as adverbs. According to the authors, numerals in Karitiana do not express multiplicity of entities, but the idea that an event occurs more than once.

In Tupinambá nouns can be combined with the form *etá* to explicitly express multiple entities, as given in examples (6), (7) and (8) below. The formal and functional characterization of *etá*, translated as 'many; many times' by Anchieta (1990), in Tupinambá is still not clear. Syntactically, *etá* occurs immediately after the nucleus of the NP, as illustrated in example (7) below. Morphologically, the form *etá* can be inflected with person markers', as illustrated in example (8) below. Since it can be inflected, *etá* has the status of an independent word in Tupinambá.

- | | | | |
|-----|-----------|---|--|
| (6) | TUPINAMBÁ | <oca> | <óçetà> |
| | | <i>oka</i> | <i>ok-eta</i> |
| | | house | house-many |
| | | 'house or houses' | 'houses' (Anchieta, 1990, p. 8v) |
| | | | |
| (7) | TUPINAMBÁ | <caraibe.á paraibiguara> | |
| | | <i>karayb-etá</i> | <i>paraybi-guara</i> |
| | | non.indigenous-many | paraiba-origin |
| | | 'the non-indigenous from Paraíba' (Sampaio, 1906, p. 294) | |
| | | | |
| (8) | TUPINAMBÁ | <ndetá> | <çetá> |
| | | <i>nd-etá</i> | <i>s-etá</i> |
| | | 2II.many | 3.II-many ¹¹ |
| | | 'yours' | 'his / hers/ its / theirs' (Sampaio, 1906, p. 296) |

The classification of *etá* as a noun or a verb is still a matter of debate¹². According to Rodrigues (1996b), in Tupinambá, only nouns can be inflected with person prefixes of set II. If this is the case, *etá* is classified as a 'noun'. Leite (2003, p. 18) shows evidence that Tupinambá should be classified as an active language. In her analysis, the set II cannot be used as a property that defines nouns and verbs in Tupi-Guarani languages. Instead, nouns and verbs can occur as predicates and arguments, but only verbs can occur with nominalizers. If this is the case, *etá* can be classified as a stative verb.

¹¹ In this work *s-* is analyzed as third person, since it occurs in the same paradigm as other personal prefixes.

¹² For further discussion on the distinction between nouns and verbs in Tupi-Guarani verbs, see Queixalós (2006), Meira (2006) and the collection of papers, organized by Queixalós (2001).

Evidence from Kamaiurá supports the analysis of *etá* as a stative verb. According to Seki (1990, p. 456), in that language, the cognate word (*h*)*etã* ~ (*r*)*etã* 'be a lot' is classified as a 'descriptive verb'.

According to Lemos Barbosa (1956), during the colonial stage of development of this language (called *Brasílica*) the use of the word *etá* 'many' had become a strategy for highlighting plurality¹³. Rodrigues (1996a, p. 13) analysed this change as grammaticalization of the word *etá* 'many' into a 'plural word'. In (9), examples given by Lemos Barbosa for the use of *etá* in *Brasílica* are provided:

(9)	BRASÍLICA	<pirá>	'peixe'	<pirá etá>	'peixes'
		<syry>	'siri'	<syry etá>	'sirís'
		<paka>	'paca'	<pak' etá>	'pacas'
		<gûyratinga>	'garça-branca'	<gûyrating' etá> ¹⁴	'garças-brancas' (Lemos Barbosa, 1956, p. 45)

The process of grammaticalization of *etá* from 'many' to a plural word had continued in the development of *Brasílica* into Nheengatu. Alongside with the increase in its use for indicating plurality, the word *etá* had lost semantic content: from a lexical word meaning 'many', to a grammatical word used to mark the general category of plural. Evidence that the grammaticalization of the plural word has resulted in loss in semantic content can be found in data provided by Barbosa Rodrigues for the 19th century Nheengatu spoken in the Rio Negro. In this variety, the plural word *etá* can be used even with numerals that indicate low quantity, such <mocoin> 'two', thus, it no longer means 'lots', as it was the case in Tupinambá. Examples (10a) and (10b) shows the variation of use of the plural word with numerals:

(10) 19th CENTURY NHEENGATU

(a) <nhaan mocoin taira étá u çu Curupira irumo>

<i>nhaã</i>	<i>mukoim</i>	<i>taira</i>	<i>eta</i>	<i>u-su</i>	<i>Kurupira</i>	<i>irumo</i>
DEM	two	son	PL	3.A-go	Kurupira	COM

'Those two sons went with Kurupira.' (Barbosa Rodrigues, 1890, p. 53)

(b) <mucoin curumí u caima caá açú pitêra rupi>

<i>mukoim</i>	<i>kurumi</i>	<i>u-kaima</i>	<i>kaa</i>	<i>asu</i>	<i>pitêra</i>	<i>rupi</i>
two	boy	3.A-get.lost	forest	big	middle	PERL

'Two boys got lost in the middle of the forest.' (Barbosa Rodrigues, 1890, p. 49)

Moreover, while in the 16th century Tupinambá, the word *etá* could be inflected by prefixes belonging to person markers of set II, the inflection is no longer available in Nheengatu. Grammaticalization also caused the loss of phonetic substance, from *etá* into *ita* ~ *ta*.

¹³ In Lemos Barbosa's words: "No tupi colonial, passou-se a empregar *etá* 'muitos' para realçar a pluralidade" [in the colonial Tupi, *etá* 'many' started to be employed for highlighting the plurality] (translation added).

¹⁴ Lemos Barbosa uses the symbol (') to indicate that the non-stressed final vowel *a* is suppressed when combined with *etá*.

GRAMMATICALIZATION INTO A THIRD PERSON PLURAL PRONOUN

The grammaticalization of a plural word creates phonological and morphological material for the emergence of a third person plural pronoun, the source for the development of number agreement in 21st century Nheengatu. In order to understand the emergence of number agreement in the current stage of development of this language, it is necessary to describe the situation of the language in the 19th century. In the first subsection below the restructuration of the pronominal system from Tupinambá to 19th century Nheengatu is analysed. Then, in the next subsection, the new pronominal system is discussed in relation to other grammatical properties of 19th century Nheengatu.

RESTRUCTURATION OF THE PRONOMINAL SYSTEM: FROM TUPINAMBÁ TO 19TH CENTURY NHEENGATU

Languages of the Tupi-Guarani family generally do not use number as a grammatical category in their pronominal system. Tupinambá has a series of free pronouns and four series of person markers for cross-reference on verbs and nouns. The pronouns and person prefixes distinguish speaker and listener. Additionally, the pronominal system distinguishes the conjunction of speaker and listener, called first person inclusive, from the conjunction of speaker and a third participant, called first person exclusive. The expression of external participants is made by demonstratives, without differentiating singular from plural. In Table 2 the pronominal system of Tupinambá is presented.

According to Jensen (1998), set I occurs exclusively with verbs. Set II occurs with nouns, postpositions, and verbs. Set III occurs with intransitive verbs to indicate that its unique argument is co-referential to the subject (or unique argument) of the verb of the main clause. Set IV has two portmanteau markers, *oro-* and *opo-* that refer simultaneously to agent and patient (Rodrigues, 2010, p. 34). The prefix *oro-* refers simultaneously to first singular agent and second singular patient, or to first plural agent and second singular patient. For instance, *orosúβ* is translated both as 'I visit you (SG)' and as 'We visit you (SG)'. The prefix *opo-* refers simultaneously to first singular agent and second plural patient, or to first plural agent and second plural patient. For instance, *oposúβ* is translated both as 'I visit you (PL)' and as 'We visit you (PL)'¹⁵.

During the development of Nheengatu from Tupinambá, sets III and IV disappeared. In Nheengatu set I occurs with transitive and active intransitive verbs to indicate the subject. Set II indicates complements of nouns and postpositions and the unique argument of inflectional stative verbs. In the description of Nheengatu, the traditional terminology of the four sets of prefixes is not used. Instead, the paradigm cognate to set I is called (A) for agentive; whereas the cognates of set II are called non-agentive (NA).

Table 2. Tupinambá person markers and pronouns, adapted from Jensen (1999, 1998)¹⁶.

	Pronouns	Set I	Set II	Set III	Set IV
1SG	ijε	a-	se-	wi-	
1PL.INCL	jandε	ja-	jane-	ja-	
1PL.EXCL	ore	oro-	ore-	oro-	
2SG	ndε	ere-	ne-	e-	oro-
2PL	peʔε	pe-	pe-	pe-	opo-
3	Demonstratives	o-	i- ∞ s-	o-	

¹⁵ For an overview of the function of set IV in Tupi-Guarani languages, see Monserrat; Facó-Soares (1983)

¹⁶ In this analysis, the prefixes *i-* and *s-* are considered third person markers, and not relational prefixes.

Additionally, the forms for first person exclusive disappeared, and the formerly first person plural inclusive morpheme is used as first person plural whether referring to the speaker and listener, or to the speaker and a third participant. Furthermore, two third person pronouns emerged: *aʔɛ* 'third person singular' and *aeta* ~ *aita* 'third person plural'. The Nheengatu pronoun *aʔɛ* evolved from the Tupinambá demonstrative *aʔɛ*, described by Rodrigues (2010, p. 27) as 'far from the speaker; non-visual; possessed'. The properties of being far from the speaker and of being non-visual have been lost. The Nheengatu pronoun *aeta* ~ *aita* 'third person plural' results from a fusion of *aʔɛ* 'third person singular' and *eta* 'plural word'. Stradelli (1929, p. 363) mentions that the intermediary form *aeta* was still found in the 19th century, although he explicitly mentions that *aita* was already the most common form. The schema in (11) illustrates the fusion of *aʔɛ* 'demonstrative' with *eta* ~ *ita* 'plural'.

- (11) xvi > xviii > xix
aʔɛ eta > *aʔɛ eta* > *aeta* ~ *aita* ~ *ainta*
 DEM many > DEM PLURAL.WORD > 3PL

In the 19th century, the pronominal system of Nheengatu radically differs from the original pronominal system of Tupinambá. In Table 3, the pronominal system and the respective person markers of 19th century Nheengatu are presented. Notice that Nheengatu has only one form for indicating first person plural, and has two pronouns to indicate the third person: a singular and a plural one.

Even though 19th century Nheengatu has free pronouns to indicate the opposition between third person singular and plural, the same opposition does not occur in the prefixes of the active and non-active series. The prefix *u-* indicates 'third person, active', unspecified for number, thus, it occurs either when the subject is singular, as illustrated in (12), or when the subject is plural, as illustrated in (13) and (14) below. The category of plural is indicated in the subject NP by the grammatical word *etá* modifying its nucleus, such as in (13), or by the newly emerged pronoun *aitá* 'third person plural', such as in (14).

The prefixes *i-* and *s-*¹⁷ mark 'third person singular, non-active'. They are exclusively used to indicate third person singular. For indicating plural, the 19th century Nheengatu uses a genitive construction, in which *aita* occurs as the possessor and any other noun occurs as the possessed, as shown in (15). Structurally, *aita* occurs in the same position as any other noun in the position of possessor. For instance, in (16), the noun *makaka*¹⁸ 'monkey' occurs as possessor and the noun *peá* 'heart' occurs as possessed.

Table 3. 19th century Nheengatu pronouns and person markers.

	Pronouns	A Set	NA Set
1SG	<i>ijɛ</i>	a-	se-
1PL	<i>jandɛ</i>	ja-	jane-
2SG	<i>indɛ</i>	re-	ne-
2PL	<i>peʔɛ</i>	pe-	pe-
3SG	<i>aʔɛ</i>	u-	i- ∞ s-
3PL	<i>aeta</i> ~ <i>aita</i>		-----

¹⁷ The choice between *i-* and *s-* is lexically determined.

¹⁸ At this point of the analysis of 19th century Nheengatu, it is not yet possible to verify if *makaka* should be analyzed as *makak-a* 'monkey-REFER' or as *makaka* 'monkey'. Since this analysis is not relevant for the description of the rise of number agreement in Nheengatu, the referrers will not be taking into consideration in the gloss.

- (12) 19TH CENTURY NHEENGATU <yepé, cunhan, paá, uricó, paá, mocoin membira>
jepe cunhã paa u-riko paa mokoin membira
 INDF woman REP 3.A-have REP two offspring
 'They say that a woman had two children' (Barbosa Rodrigues, 1890, p. 53)
- (13) 19TH CENTURY NHEENGATU <amu etá unhehi>
amu eta u-nhehi
 other PL 3.A-say
 'Others said.' (Barbosa Rodrigues, 1890, p. 29)
- (14) 19TH CENTURY NHEENGATU <aitá u çu>
aita u-su
 they 3.A-go
 'They go.' (Barbosa Rodrigues, 1890, p. 42)
- (15) 19TH CENTURY NHEENGATU *aitá makira*
 they hammock
 'their hammock' (Barbosa Rodrigues, 1890)
- (16) 19TH CENTURY NHEENGATU *makaka peá*
 monkey heart
 'the heart of the monkey' (Barbosa Rodrigues, 1890)

THE IMPACT OF THE THIRD PLURAL FORM TO THE NULL SUBJECT PARAMETER IN THE 19TH CENTURY NHEENGATU

As shown in Table 3 in the previous section, the grammaticalization of *aita* 'third person plural' does not coincide with an emergence of prefixes of third person plural neither in active (A) series nor in the non-active (NA) series. The pronoun *aita* 'third person plural' occurs as subject of transitive and agentive intransitive verbs, but the verb is inflected by the prefix *u-* 'third person', which does not make reference to the grammatical category of number.

The variety of Nheengatu spoken in Rio Negro in the 19th century allows null subject. Thus, transitive and agentive intransitive verbs occur inflected by prefixes of set A with no necessity of a pronominal NP in the position of subject. The examples in (17) illustrate the use of verbs inflected by the prefixes of set A with no NP in the position of subject.

Since 19th century Nheengatu is a pro-drop language, transitive and agentive intransitive verbs do not require the presence of an NP in the position of subject. However, the lack of a prefix to indicate 'third person plural' gives rise to the necessity of expressing the subject by a lexically manifested NP or by the pronoun *aeta* ~ *aita* 'third person plural'. In the discursive sequence in (18), the pronoun *aeta* ~ *aita* occurs before verbs in which the *aita* refers to the subject, creating the structure *aita u-VERB*. This is the case for the verbs <*u-çu*> 'go', <*u-acema*> 'find' in the first line; <*u-iuúca*> 'take', <*u-çaam*> 'wait' in the second line; <*u-acema*> 'find' and <*u-iapy*> 'shoot' in the third line. In the fourth and fifth lines, the subjects of the verbs <*u-euêo*> 'fly', <*u-iure*> 'go back' and <*u-iumu*> 'shoot' refer to a third singular participant

<uhiua> 'arrow'. In these verbs, the inflection with the prefix <u-> is sufficient to encode the subject. In the last sentence in(18) the verbs <u-are> 'fall' and <u-manu> 'die', the subject is manifested by a postposed NP *nhaan curumi* <that boy>.

(17) 19TH CENTURY NHEENGATU

(a) <yá çu ana>

ja-su ana

1PL.A-go PFV

'Let's go! (Barbosa Rodrigues, 1890, p. 28)

(b) <re ricu uhiua re-iucá arama ne rembiara>

re-riku uhiua re-juca arama ne-rembiara

2SG.A-have arrow 2SG.A-kill to 2SG.NA-food

'Do you have an arrow to kill your food?' (Barbosa Rodrigues, 1890, p. 28)

(c) <cuire pe ruiare ana>

kuire pe-ruiare ana

now 2PL.A-believe PFV

'Now you already believe' (Barbosa Rodrigues, 1890, p. 46)

(d) <U munuca macaco pô>

u-munuka makako po

3A-cut monkey hand

'He cut the monkey's hand.' (Barbosa Rodrigues, 1890, p. 24)

(18) 19TH CENTURY NHEENGATU

<Coema aetá u çu aap. Aitá u acema uhiua.>

koema aeta u-su aap. Aita u-asema uhiua

morning they 3A-go there they 3A-find arrow

<Aé uana aitá u-iuúca ihia, aé uana aetá u çam uirá recé,>

ae uana aita u-juuca ihia ae uana aeta u-saam uira rese

DEM PFV they 3A-take arrow DEM PFV they 3A-try bird about

<u-euêo uhiua u-iure cecé u-iumu cecé,>

u-eueo uhiua u-jure sese u-jumu sese

3A-fly arrow 3A-go.back 3NA:about 3A-hit 3NA:about

<u are catu, aé uana u-manu nhaan curumi>

u-are katu ae uana u-manu nhaã kurumi

3NA-fall be.well DEM PFV 3NA-die DEM boy



'In the morning they went there. They found an arrow. Then, they shot the arrow, then they tried it on the bird, they found the bird flying, they shot, the arrow flied and went back to him, and hit him, he fell, then that boy died.' (Barbosa Rodrigues, 1890, p. 30)¹⁹.

In pro-drop languages the expression of the subject by a pronoun is rare, occurring only in emphatic situations. The occurrence of the pronoun *aita* in every sentence of example (18) is, thus, unexpected. However, in 19th century Nheengatu there is no prefix to indicate third person plural. Instead, prefix *u-* 'third person' indicates both singular and plural. In order to make clear that the subject is plural, the speakers make frequent use of the subject NP *aita*. The high frequency of the structure *aita u-VERB* provides the morphological material for the emergence of prefixes indicating a third person plural. The emergence of such prefixes will be shown in the next section.

THE RISE OF NUMBER AGREEMENT IN THE 21ST CENTURY NHEENGATU

In conservative varieties of Nheengatu, the third person plural pronoun is found as a free form (*ainta*, *aita*, *aetá*), as illustrated by (19) and (20); or as a proclitic *ta=*, as illustrated by (21). Example (19) was registered by Moore, Facundes, and Pires (1993). Examples in (20) and (21) were produced by an 81 year-old woman, who is bilingual in Warekena and Nheengatu.

- (19) *presizo aētá u-ıştuda pohtugéiš upé*
 need they 3-study Portuguese in

'It is necessary that they study in Portuguese' (Moore, Facundes, & Pires, 1993, p. 106)

- (20) *puru werekena panhe ainta membira ita aiwā ainta u-su=wā*
 pure warekena all they offspring PL then they 3-go=PFV
 All of their children were pure Warekena. Then, they went way.'

- (21) *amu ita ti=ta=u-putari*
 other PL NEG=PL=3-want
 'Others didn't want.'

The clitic form *ta=*, which combines with the verb inflected with the third person prefix *u-*, unspecified for number, has provided the phonological material for the emergence of a third person plural. Phonetically, both the clitic *ta=* and the prefix *u-* are pronounced by the older speakers. Thus, a verb such as *ta=u-putari* in the example (21) above is realized as [taw.pu.'ta.ri] by older speakers.

In the younger generation a phonological fusion between the earlier clitic *ta=* and the prefix *u-* gives rise to a third person plural prefix, which distinguishes two dialects of Nheengatu. In the Negro and Içana variety the third person plural prefix is *ta-* 'third person plural, active', as illustrated in (22) below. In the Xié variety the combination of clitic *ta=* and the prefix *u-* gives rise to the prefix *tu-* 'third person plural, active', as illustrated in (23) below ²⁰.

¹⁹ The expression <aé uana> 'demonstrative perfective' is translated by Barbosa Rodrigues as *logo* 'soon', or as *então* 'then'. In 21st century Nheengatu, the expression is found grammaticalized as a connector: /aiuã/ [aj.'wã] 'then'.

²⁰ Similar processes are discussed by Mithun (1991) for various languages of North America.

(22) NEGRO-IÇANA VARIETY *ape* *paa* *ta-su=wā*
 CONJ REP 3PL.A-go=PFV
 'Then, they say, they went away.'

(23) XIÉ VARIETY *mairame* *tu-sika* *kua=kiti*
 when 3PL.A-arrive DEM=ALL
 'When did they arrive here?'

Notice that, in contrast to 19th century Nheengatu, verbs in the 21st varieties verbs can be inflected with a third person plural prefix and, consequently, do not require the subject NP to be expressed.

The process of grammaticalization of the free pronoun *ainta* into a prefix has also occurred with nouns. In example (20) above, the NP *ainta membira*, structured as $[[ainta]_{\text{POSSESSOR}} [membira]_{\text{HEAD-NP}}]$, indicates the relation between a noun and its possessor. The frequency of the use of *ainta* as the possessor in the genitive construction seems to have contributed to the emergence of the prefix *ta-* 'third person plural non-active' in both Içana-Negro variety and Xié variety. In examples (24) and (25) below, the use of *ta-* 'third person plural non-active' is shown. Notice that in Negro-Içana variety, the prefix *ta-* 'third person plural, non-active' is homophonous with the prefix of *ta-* 'third person plural active', whereas in the Xié variety, the prefix of 'third person plural non-active' is *ta-* and the prefix of 'third person plural active' is *tu-*.

(24) NEGRO-IÇANA VARIETY *ta-kupi* *ta-kupixa*
 3PL.A-clear 3PL.NA-cultivated.field
 'They cleared their field'

(25) XIÉ VARIETY *amu* *ita* *tu-rasu* *te* *ta-raira*
 other PL 3PL.A-bring FOC 3PL.NA-son
 'Others bring their children.'

With the emergence of the prefixes for third person plural, the pronominal system became more symmetrical in comparison to the pronominal system of the 19th century variety. The distinction between third person singular and third person plural is present among the pronouns, the prefixes of the active series, and the prefixes of the non-active series. Table 4 below shows the pronominal system of 21st century Nheengatu.

The restructuring of the pronominal system has provided the morphological material for the possibility of agreement between subject NP and verbs. In (26) the subject NP (*panhe ita musapi*) agrees in number with the verb *tariku*, and the object NP *tarimiku ita* (in conjunction with the object NP *taraira*) is marked by a prefix for third

Table 4. 21st century Nheengatu pronouns and person markers.

	Pronouns	Set A	Set NA
1SG	ijε	a-	se-
1PL	jandε	ja-	jane-
2SG	inde	re-	ne-
2PL	pe?ε	pe-	pe-
3SG	a?ε	u-	i- ∞ s-
3PL	aeta ~ aita	ta- ~ tu-	ta-



person plural non-active. In (27) the subject *ai jawuti irũ* has the following complex structure: NP followed by a postpositional phrase (*jawuti irũ*) as its adjunct. The idea of multiple entities (*ai* and *jawuti*) allows the verb *manduai* to be inflected by the third person plural active prefix.

(26) *panhe ita musapi paa ta-riku ta-rimiriku ita ta-raira*
 every PL three REP 3PL.A-have 3PL.NA-wife PL 3PL.NA-son
 'They say that all the three [brothers] have their wives and their sons.' (Negro-Içana variety)

(27) *ape paa ai jawuti irũ ta-manduai ta-minhã=rã jepe kupixa*
 CONJ REP sloth tortoise COM 3PL.A-remember 3PL.A-make=PROSP INDF field
 'Then, they say that the sloth with the tortoise has remembered to make a [manioc] field.' (Negro-Içana variety)

The agreement between subject NP and the verb is not obligatory. In (28) the subject NP does not occur with the plural marker *ita* and the verbs occur inflected by third person plural.

(28) *a=rire=wã kua kariwa tu-sika tu-munhã kua pista ike*
 DEM=after=PFV DEM non.indigenous 3PL.A-arrive 3PL.A-make DEM airport here
 'After that, that non-indigenous people arrived and made this airport here.' (Xié variety)

In examples (25) and (26), there is agreement between the subject-NP and the verb, whereas in (28) the verb *tu-sika* is marked by the plural prefix, but the subject-NP is left unmarked. On the current state of this research, it is not clear which factors are correlated to the presence of number agreement between subject and verb.

MULTIPLE INFLUENCES IN THE RISE OF NUMBER AGREEMENT IN NHEENGATU

As stated in the introduction, the development from Tupinambá to Nheengatu can be divided into four stages. The main changes in each stage are summarized in the diagram below.

In order to better understand the whole process of the rise of number agreement, the sociolinguistic background of each stage should be taken into account. The need of considering each stage individually comes from the fact that the history of the development from Tupinambá into *Brasílica* and then into Nheengatu is not homogenous: there are stages in which the potential influence of Portuguese is more prominent, and others in which the potential influence of other indigenous languages seems to be more prominent.

16TH-18TH CENTURY: THE GRAMMATICALIZATION OF THE PLURAL WORD AS A RESULT OF CONTACT WITH PORTUGUESE

According to Lemos Barbosa (1956), during the development from Tupinambá into *Brasílica*, the word *eta* was commonly used for highlighting plurality. This change is analysed here as a grammaticalization of the noun *eté* 'many' into a plural word that no longer accepts person markers. Based upon the social background in which *Brasílica* was spoken in the 18th century, the typological properties of the languages involved, and similar cases described in literature, it is possible to ascribe this change to the influence of Portuguese on *Brasílica*.

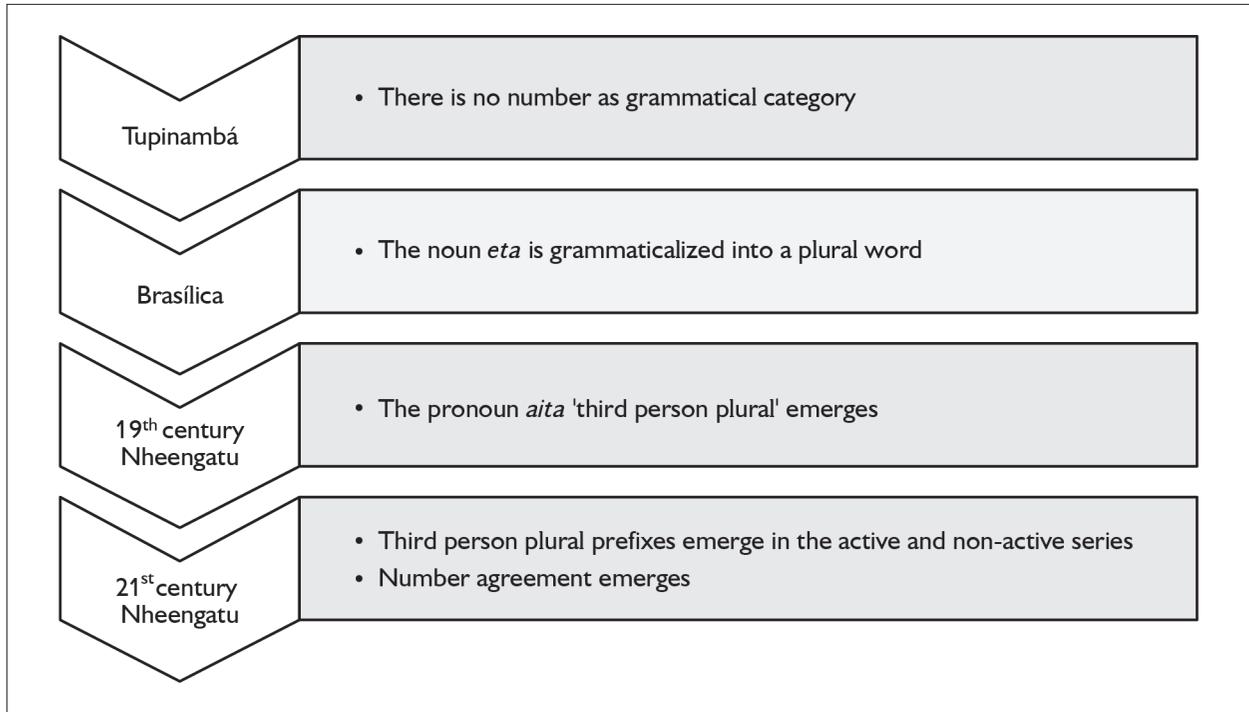


Figure 1. The stages of grammaticalization of number agreement from Tupinambá to Nheengatu.

The social background, which shows that there was ample opportunity for language contact between Portuguese and Brasília, was discussed in detail in the overview of the social history of Nheengatu above. In this section, first, the typological properties of Portuguese and Brasília are presented in more detail. And then, similar cases found in literature are presented.

In current Brazilian Portuguese, plural is marked in the NP, and by agreement in the VP. In the NP, nouns, determiners, and adjectives carry plural markers in the standard variety of Portuguese. In non-standard variety of Brazilian Portuguese plural can be marked only in the determiner. In the VP, verbs may, depending on sociolinguistic factors, agree with their subject in number (Scherre; Naro, 2013; 2014). Due to the lack of sources of 18th century Portuguese spoken in the Amazon region, we can only assume that plurality in nouns is marked similarly for the variety of Portuguese that Brasília was in contact with.

Recall that before contact with Portuguese, Tupinambá is documented as a language in which plural is not marked (Anchieta, 1990). After two centuries of contact (in the 18th century), documents of Brasília (the descendent of Tupinambá) show a plural word. It is then plausible that this change (*i.e.* the grammaticalization of a plural word) is induced by contact with Portuguese.

The grammaticalization of *eta* into a plural word is not only a characteristic of Brasília, but is also observed in other Tupi-Guarani languages, such as Chiriguano, Tapiete and Ka'apor (Rose, 2012). In (30) the use of *-reta* ~ *-re* in Tapiete is exemplified. In (31) the use of the clitic *ta* in Ka'apor is exemplified²¹.

²¹ Ka'apor also has the word (*h*)*etá* 'be many' as a stative verb (for details, see Lopes 2009).

(30) TAPIETE (a) *karai-re*
 white.man-PL
 'white men' (González, 2005)

(31) KA'APOR (a) *pε fumukupape ta u-wir tĩ*
 and black.vulture Q 3-come also
 'The black vultures also came back' (Kakumasu *apud* Lopes 2009, p. 126)²²

The grammaticalization of cognates of *eta* as plural words in Chiriguano, Tapiete, Ka'apor and Nheengatu seems to be an independent instance of contact between these languages and the colonizers' languages: Portuguese, in the case of *Brasílica* and Ka'apor; and Spanish, in the case of Chiriguano and Tapiete.

THE 19TH CENTURY AND ONWARDS: THE EMERGENCE OF NUMBER AGREEMENT AS A RESULT OF THE CONTACT WITH ARAWAKAN LANGUAGES

In the 19th century, a pronoun indicating third person plural emerges in Nheengatu. The third person plural pronoun *aita* was registered in different varieties of the language by Hartt (1929), Tastevin (1923), Barbosa Rodrigues (1890), Stradelli (1929), among others. The fact that the same pronoun was found in different varieties indicates that this form had emerged in an earlier period.

As has been shown in the fifth section of this article, around the beginning of the 20th century, the 'third person plural' pronoun *aita* was grammaticalized into a clitic. From that moment until the 21st century, this clitic has evolved into prefixes indicating third person plural. The grammaticalization of the clitic *ta=* into prefixes is very recent, since the prefixes occur in the Nheengatu spoken by the younger generation, but not in the variety spoken by the older generation. Thus, to understand the grammaticalization of *aita* into prefixes, the social context in which the older generation acquired Nheengatu should be taken into account.

The older speakers of currently spoken Nheengatu were born around 1930. Most of these speakers do not speak any Portuguese. They have reported that they were born in bilingual societies; speaking Nheengatu and Warekena in the case of the inhabitants of Rio Xié, and Nheengatu and Baniwa in the case of the ones from Rio Içana. The subsequent generation, born around 1950, did not acquire the Arawak languages (Baniwa and Warekena)²³. Thus, it was around the 1950s that the language shift took place. According to the older speakers, the contact with Brazilian Portuguese took place later. A similar process of language shift is supposed to have taken place among the Baré people in an earlier period.

This historical background is important to understand the context in which the rise of number agreement occurred in Nheengatu. The grammaticalization of *aita* into the clitic *ta=* started with the old generation who speak both one of the north-Arawakan languages and Nheengatu, and do not speak Portuguese. Thus, this process cannot

²² Lopes (2009) refers to Kakumasu (1990) as the source of example (31). However, a bibliography entry for Kakumasu (1990) is not mentioned in Lopes's work, but Kakumasu & Kakumasu (1990) is. Hence we presume that Kakumasu & Kakumasu (1990) is the source of the example. In Lopes (2009), the abbreviated form Q means 'quantifier'.

²³ Some of them, however, can understand Baniwa or Warekena, due to the contact to the elderly people.

be only related to Portuguese, but may have occurred as a grammatical replication²⁴ of the languages from which they have shifted, *i.e.*, Baré, Baniwa and Warekena.

The description of these three languages supports the hypothesis that the rise of number agreement was a change induced by the contact with the Arawakan substratum (or even adstratum, since Warekena and Baniwa are still spoken in the region). Baré, Baniwa and Warekena all have third person plural prefixes and allow null subject, as given in (32), (33) and (34) below.

(32) BARÉ *id'ya awadari-nū me-nikā me-tjana-ka ukálisi*
 then cairara-PL 3PL-eat 3PL-stay-DECL inajá.fruit
 'Then cairara-monkeys were eating inajá.' (Aikhenvald, 1995, p. 18)²⁵

(33) BANIWA *na-dia*
 3PL.A-come
 'They came back.' (Ramirez, 2001, p. 105)

(34) WAREKENA *i-fumiawa e teJetfi fiani-pe ni-fa-wa ni-tsunua-hã*
 3SG.NF-wife.FEM and there child-PL 3PL-go-NONACC 3PL-tear-PAUS
 'His wife and three children went to tear (manioc)' (Aikhenvald, 1998, p. 246)²⁶

Even if the contact with Baré, Baniwa and Warekena was prominent as the trigger of the process, the contact with Portuguese cannot be disregarded. In the variety of Portuguese spoken in São Gabriel da Cachoeira, number agreement is often used²⁷. As stated in 5, number agreement between the verb and its subject NP is not obligatory. However, in contexts in which the speakers are more aware of the way they speak, they avoid sentences without agreement. For instance, in elicitation, speakers always give sentences with number agreement. Also, when the speakers write in Nheengatu, they consistently mark the plural in the subject NP and in the prefix of the verb; even when the *ita* 'plural' and third person plural prefix are contiguous²⁸, as illustrated in (35).

Based on the data from Arawakan languages and from Brazilian Portuguese, it seems reasonable to suppose that the emergence of number agreement is a result of the influence from both Arawakan languages and Portuguese.

(35) <*u-puraki-sara ita tau-mujã mukúi mūti*>
 3SG.A-trabalhar-NMZ PL 3PL.A-make two group
 'The workers made two groups.' (Comunidade de Anamoim - Xié, ms.)

²⁴ The term is used in the narrow sense of the transference of grammatical meaning (including grammatical meanings or function) or combinations of meaning, as defined by Heine and Kuteva (2005, p. 2-3).

²⁵ In Aikhenvald (1995), the abbreviate form means 'declarative'.

²⁶ In Aikhenvald (1998), the following abbreviations are used: FEM 'feminine'; NF 'non feminine'; NONACC 'non-accusative'; PAUS 'pausal marker'.

²⁷ I am unaware of studies on number agreement in the variety of Portuguese spoken in São Gabriel da Cachoeira. During a trip to the region in April 2013, I observed that speakers of Nheengatu, when speaking Portuguese, mark agreement between the subject and the verb even in informal conversations.

²⁸ In writing, the Nheengatu speakers from Rio Negro and Rio Içana use the form <tau-> as a prefix of third person plural active. When asked why they do so, if they pronounce [ta], they said that in this way the prefix that combines with verbs <tau-> [ta] is kept apart from the prefix that occurs with nouns <ta-> [ta].

CONCLUDING REMARKS

The case of Nheengatu helps to fill a hiatus in the studies of grammaticalization by showing the first complete process of change: from a language in which no number as a grammatical category is found, into a system in which number is such a pervasive category that it has a role in agreement. In stark contrast with other cases of grammaticalization discussed in (Lehmann, 2002), the grammaticalization of Nheengatu is a remarkably short process. This is used as support for the idea that this change was induced by language contact.

The described result of the supposed contact with Portuguese was rather slow: four centuries were necessary for a rather small change, with relatively small typological impact, to take place. From the 19th century onward, the main typological changes have emerged: the rise of third person prefixes and the rise of agreement. This paper presents evidence that these more pervasive changes are a result of the language shift from Arawakan languages.

It is also important to understand that the rise of number agreement in Nheengatu is not an isolated fact in the language structure; it may indicate a complete reconceptualization of the number category, which includes changes in the patterns of the use of verbal number (Cruz, 2014), changes in the properties of numerals (Cruz, 2013), and the emergence of countable and uncountable noun distinctions (Cruz, 2011, pp. 164-5). However, these changes are left to future investigations.

The case presented in this paper helps to demystify the idea that grammatical transfer can take place only between languages that are very similar typologically: even though Tupinambá was typologically far from languages that have number system, a number agreement system has emerged in its descendent language, Nheengatu. We hope that by presenting the Nheengatu case, we inspire future researchers to take into account language contact as a possible way of explaining the development of grammatical structure over time.

ACKNOWLEDGEMENTS

Information on Nheengatu was obtained via fieldwork on the Rio Negro, Rio Içana and Rio Xié, São Gabriel da Cachoeira, Brazil, carried out in 2007-2014. I am grateful to my Baré, Warekena and Baniwa hosts, friends, and language teachers; and to FOIRN, CAPES (BEX 4267-05-2), and CNPq (482555/2013) for supporting this research. I am also grateful to the participants of the conference *Rethinking Contact Induced Change* (Leiden, 2011) for their helpful comments; to Wolf Dietrich, Sebastian Drude and the anonymous reviewers for further suggestions on this paper; and to Adriana Machado Estevam for her care to improve its English style. I am, of course, responsible for any remaining mistakes or improper argumentations.

ABBREVIATIONS AND CONVENTIONS

< >	spelling	CAUS	causative	NEG	negation
1	first person	COM	commitative	NMZ	nominalizer
2	second person	CONJ	conjunction	PROSP	prospective
3	third person	DEM	demonstrative	PERL	perlative
I	set I	FOC	focus	PFV	perfective
II	set II	INDF	indefinite	PL	plural
A	agent-like argument	LOC	locative	REP	reportative
ALL	allative	NA	non-agent-like argument	SG	singular



REFERENCES

- AIKHENVALD, A. Y. Warekena. In: DERBYSHIRE, D. C.; PULLUM, G. K. (Eds.). **Handbook of Amazonian languages** 4. Berlin: Mouton de Gruyter, 1998. p. 225-229.
- AIKHENVALD, A. Y. **Bare**. Munique & Newcastle: Lincom Europa, 1995.
- ANCHIETA, J. **Arte de gramática da língua mais usada na costa do Brasil**. São Paulo: Loyola, fac-simile from the first edition, 1990.
- BARBOSA RODRIGUES, J. Poranduba amazonense. **Anais da Biblioteca Nacional do Rio de Janeiro**, v. 14, n. 2, p. 1-337, 1890.
- BESSA FREIRE, J. R. **Rio Babel - a história das línguas na Amazônia**. Rio de Janeiro: Atlântica, 2004.
- BORGES, M. V. **Aspectos fonológicos e morfossintáticos do Avá-Canoeiro (Tupi-Guarani)**. Campinas: PhD dissertation, UNICAMP, 2006.
- CABALZAR, A., & RICARDO, C. A. **Povos indígenas do Rio Negro: uma introdução à diversidade socioambiental do noreste da Amazônia brasileira**. São Paulo & São Gabriel da Cachoeira: ISA & FOIRN, 2006.
- COMUNIDADE DE ANAMOIM - XIÉ (ms.). **Maye yamunhã bũgu - A construção do bongo**. (Organized by Florêncio Cordeiro).
- CORBETT, G. G. **Number**. Cambridge: Cambridge University Press, 2004
- COUTO DE MAGALHÃES, J. V. **O selvagem**. Rio de Janeiro: Typographia da Reforma, 1876.
- CRUZ, A. Reduplication in Nheengatu. In: GOODWIN GÓMEZ, G.; VAN DER VOORT, H. (Eds.). **Reduplication in indigenous languages of South America**. Leiden: Brill, 2014. p. 115-141.
- CRUZ, A. De múltiplos eventos a múltiplas entidades: mudança induzida por contato nos numerais em Nheengatu. **II Congresso Internacional da Faculdade de Letras da UFRJ - Simpósio Contato e Mudança Linguística**. Rio de Janeiro: EdUFRJ, 2013. p. 11-12.
- CRUZ, A. **Fonologia e gramática do Nheengatu - a língua geral falada pelos povos Baré, Warekena e Baniwa**. Utrecht: Netherlands Graduate School of Linguistics (LOT), 2011.
- DRYER, M. S. Plural words. **Linguistics**, v. 27, p. 865-895, 1989.
- FELIX, M. I. **A Língua Geral Amazônica: contribuição para o estudo de suas variedades dialetais faladas ao longo do Rio Amazonas e seus tributários no século XIX**. Belém: MA thesis, UFPA, 2002.
- GONZÁLEZ, H. A. **A grammar of Tapiete (Tupi-Guarani)**. Pittsburgh: PhD dissertation, University of Pittsburgh, 2005.
- HARTT, C. F. Notas sobre a língua geral ou tupi moderno do Amazonas. **Anais da Biblioteca Nacional do Rio de Janeiro**, v. 51, p. 305-390, 1929 [1938].
- HEINE, B.; KUTEVA, T. **Language contact and grammatical change**. Cambridge: Cambridge University Press, 2005.
- JENSEN, C. Tupí-Guaraní. In: DIXON, R. M. W.; AIKHENVALD, A. Y. (Eds.). **The Amazonian languages**. Cambridge: Cambridge University Press, 1999. p. 125-163.
- JENSEN, C. Comparative Tupi-Guarani morphosyntax. In: DERBYSHIRE, D. C.; PULLUM, G. K. (Eds.). **Handbook of Amazonian languages** 4. Berlin & New York: Mouton de Gruyter, 1998. p. 480-618.
- KAKUMASU, J. & KAKUMASU, K. **Outros textos urubu-kaapor**. Brasília: SIL, 1990.
- LEE, K.. **Conversing in colony: the Brasília and the Vulgar in Portuguese America, 1500-1759**. PhD dissertation, The Johns Hopkins University, 2005.
- LEHMANN, C. Thoughts on grammaticalization. **Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt**. Erfurt, 2002.

- LEITE, Y. A Arte de gramática da língua mais usada na costa do Brasil e as línguas indígenas brasileiras. In: BESSA FREIRE, J. R.; ROSA, M. C. (Eds.). **Línguas gerais: política linguística e catequese na América do Sul no Período Colonial**. Rio de Janeiro: EdUERJ, 2003. p. 11-24.
- LEMOS BARBOSA, A. **Curso de Tupi Antigo: gramática, exercícios, textos**. Rio de Janeiro: Livraria São José, 1956.
- LOPES, M. A. **Aspectos gramaticais da língua Ka'apor**. Belo Horizonte: PhD dissertation, UFMG, 2009.
- MEIRA, S. Stative verbs vs. nouns in Sateré-Mawé and the Tupian family. In: ROWICKA, G. J.; CARLIN, E. B. (Eds.). **What's in a verb? - studies in the verbal morphology of the languages of the Americas**. Utrecht: Netherlands Graduate School of Linguistics (LOT), 2006. p. 189-214.
- MITHUN, M. The development of bound pronominal paradigms. In: LEHMANN, W. P.; HEWITT, H.-J. J. (Eds.). **Typological models in reconstruction**. Amsterdam: John Benjamins, 1991. p. 85-104.
- MONSERRAT, R. M.; FACÓ-SOARES, M. Hierarquia referencial em línguas Tupi. **Ensaios de Linguística. Cadernos de Linguística e Teoria da Literatura** 9, 164-187, 1983.
- MOORE, D.; FACUNDES, S.; PIRES, N. Nheengatu (LGA), its history, and effects of language contact. **Proceedings of the Meeting of the Society for the Study of the Indigenous Languages of the Americas**, v. 2, n. 4, p. 93-118, 1993.
- MOSELEY, C. **Atlas of the World's languages in danger** (3rd ed). Paris: UNESCO Publishing, 2010 Available at: <<http://www.unesco.org/culture/en/endangeredlanguages/atlas>>.
- MÜLLER, A.; SANCHEZ-MENDES, L. Pluractionality in Karitiana. **Proceedings of Sinn und Bedeutung** 12. Oslo: Department of Literature, Area Studies and European Languages, University of Oslo, 2008. p. 442-454.
- OLIVEIRA, R. L. **Natureza e direções das mudanças lexicais ocorridas na língua geral Amazônica do século XVII**. Brasília: Master's dissertation, UnB, 2008.
- PRAÇA, W. **Morfossintaxe da língua Tapirapé**. Brasília: PhD dissertation, UnB, 2007.
- QUEIXALÓS, F. The primacy and fate of predicativity in Tupi-Guarani. In: VAPNARSKY, V.; LOIS, X. (Orgs.). **Lexical categories and root classes in Amerindian Languages**. Frankfurt am Main: Peter Lang Publishing, 2006. p. 249 - 288.
- QUEIXALÓS, F. (Org.). **Des noms et des verbes en tupi-guarani: état de la question**. Munich: Lincom Europa, 2001.
- RAMIREZ, H. **Línguas arawak da Amazônia setentrional: comparação e descrição**. Manaus: Editora da Universidade do Amazonas, 2001.
- RODRIGUES, A. D. Estrutura do Tupinambá. In: RODRIGUES A. D.; CABRAL, A. S. (Eds.). **Línguas e culturas Tupi**. Brasília: LALI & Campinas: Editora Curt Nimuendajú, 2010. p. 11-42.
- RODRIGUES, A. D. As línguas gerais sul-americanas. **PAPIA - Revista de Crioulos de Base Ibérica**, v. 4, n. 2, p. 6-18, 1996a.
- RODRIGUES, A. D. Argumento e predicado em Tupinambá. **Boletim da Associação Brasileira de Linguística**, v. 19, p. 57-70, 1996b.
- ROSE, F. Borrowing of a Cariban number marker into three Tupi-Guarani languages. In VANHOVE, M.; STOLZ, T.; URDZE, A.; OTSUKA, H. (Eds.). **Morphologies in contact**. Berlin: Akademie Verlag, 2012. p. 37-69.
- SAMPAIO, T. Cartas tupis dos Camarões. **Revista do Instituto Archeologico e Geographico Pernambucano**, v. XII, n. 68, p. 201-305, 1906.
- SCHERRE, M. M.; NARO, A. J. Sociolinguistic correlates of negative evaluation: Variable concord in Rio de Janeiro. **Language Variation and Change**, v. 26, p. 331-357, 2014.
- SCHERRE, M. M.; NARO, A. J. Remodeling the age variable: Number concord in Brazilian Portuguese. **Language Variation and Change**, v. 25, p. 1-15, 2013.
- SEKI, L. Kamaiura (Tupi-Guarani) as an active-stative language. In: PAYNE, D. (Ed.). **Amazonian linguistics: studies in Lowland South American languages**. Austin: University of Texas Press, 1990. p. 367-392.

STRADELLI, E. Vocabulário da língua geral português-nheengatu e nheengatu-português, precedidos de um esboço de grammática nheenga-umbuê-sáua. **Revista do Instituto Histórico e Geográfico Brasileiro**, v. 158, n. 104, p. 9-768, 1929.

TASTEVIN, C. Grammatica da lingua tupy. **Revista do Museu Paulista**, v. 13, Separata, p. 1-151, 1923.

VON DEN STEINEN, K. Mitteilung über die Forschungsreise des Herrn Theodor Koch in Südamerika. **Zeitschrift für Ethnologie**, v. 36, n. 2, p. 293-299, 1904. Available at: <<http://recall.archive.org/details/zeitschriftre36berl>>. Accessed on: March 23th, 2013.