THE NATURE AND EMERGENCE OF THE *LÍNGUA GERAL AMAZÔNICA* ACCORDING TO MUFWENE'S LANGUAGE ECOLOGY MODEL

Thomas FINBOW[1]

Abstract: This paper addresses two aspects of the *Língua Geral Amazônica* (LGA). First, we propose that *Língua Geral* as a technical term in Portuguese was inspired by administrative practices in Spain's American colonies and that, given contemporary usage, the term should be understood functionally as any Tupi-Guarani variety broadly mutually comprehensible with the colonial Old Tupi, rather than as a structurally modified variety or the speech of a particular ethnic or social group, as has been claimed. We then briefly analyse two recent hypotheses that treat the emergence of LGA as creolisation (VIEIRA; ZANOLI; MÓDOLO, 2019; NOBRE, 2019). Next, we apply key notions of the Language Ecology creolisation model (MUFWENE, 2003, 2008) to the formation of LGA and suggest a periodization.

Keywords: Língua Geral Amazônica. Nheengatu. Salikoko Mufwene. Language Ecology. Creolization.

---

1 Universidade de São Paulo (USP), São Paulo, São Paulo, Brasil; thomas.finbow@usp.br; https://orcid.org/0000-0003-4803-6603;

A NATUREZA E O SURGIMENTO DA LÍNGUA GERAL AMAZÔNICA, SEGUNDO O MODELO DE ECOLOGIA LINGUÍSTICA DE MUFWENE

Resumo: O presente artigo aborda dois aspectos da *Língua Geral Amazônica* (LGA). Primeiro, propomos que "língua geral", como termo técnico em português, foi inspirado pelas práticas administrativas nas colônias da América espanhola. Dado o uso contemporâneo, esse nome deveria ser entendido de modo functional como referindo-se a qualquer variedade tupi-guarani mutuamente compreensível com o Tupi Antigo colonial, antes que, designando uma variedade, tenha sofrido certa modificação estrutural ou tenha sido a norma expressiva de determinado grupo étnico ou social, como outros pesquisadores já defenderam. A seguir, analisamos duas hipóteses recentes que tratam o surgimento da LGA como crioulização (VIEIRA; ZANOLI; MÓDOLO, 2019; NOBRE, 2019). Apresentamos as noções centrais do modelo de crioulização da Ecologia linguística (MUFWENE, 2003, 2008) e as aplicamos ao surgimento da língua geral amazônica proferindo um esboço de periodização.

Palavras-chaves: Língua Geral Amazônica. Nheengatu. Salikoko Mufwene. Ecologia Linguística. Crioulização.

We begin by discussing the definition of *Língua Geral* (henceforth LG) in colonial Brazil. Focussing on the Amazonian region, we analyse some well-known proposals regarding the nature and formation of LG. After critically analysing two recent proposals that consider the formation of *Língua Geral Amazônica* (LGA) to be a case of creolisation, we present the central pillars of Mufwene's Linguistic Ecology model of creolisation (MUFWENE, 2003, 2010). We conclude a broad periodization of the formation of LG in the Amazon as a Mufwenean creole.

What is a *Língua Geral*?

Língua geral, in Portuguese, or *lengua general*, in Spanish, means "lingua franca", i.e., a language used for communication amongst groups of people who do not share a common first language. Such a language may or may not be the first language of one of the interacting parties, or it may be the native language of neither group.

A Spanish idea?

The instrumentalization of indigenous *lenguas generales* in Spain's American colonies was shaped by Jerónimo de Loayza's *Instrucción de la orden que se a de tener en la Doctrina de los naturales* (1545-49) (ALFARIO LAGORIO, 2003, p. 43), the decisions of the *Junta Magna* of 1568 (PÉREZ PUENTE, 2009, p. 47-48; RAMOS PÉREZ, 1986, p. 7-9), and the three Councils of Lima (1552, 1567-68, 1582-83) (ALTMAN, 2003; ALFARO LAGORIO, 2003). The outcome was that catechesis would be performed in certain indigenous languages, not in Castilian. Thus, missionaries had to learn these languages and compose catechisms in them, supported by chairs in *lenguas generales* established in 1580 at the University of Lima and the city's Jesuit College (PÉREZ PUENTE, 2009, p. 47).

In the Spanish empire, pre-Columbian indigenous lingua francas already existed, e.g., Quechua and Nahuatl, which were adapted to colonial projects (ALTMAN, 2003, 2011; ALFARO LAGORIO, 2003). Other indigenous *lengaus generales* officially fostered by the Spanish Crown included Aymara, Chibcha/Muisca, Puquina, Guarani, and Mapuche. Unlike Quechua and Nahuatl, these languages were not originally super-ethnic lingua francas but rather formed dialect continua that were spoken over extensive territories, often by diverse polities (ALTMAN, 2003, p. 60). The Spanish aimed to form koines from these languages for the purposes of evangelization and administration (ALTMAN, 2003, p. 64; ALFARO LAGORIO, 2003, p. 46-50).

Unlike the Andean region or Mexico, the territory the Portuguese Crown received under the Treaty of Tordesillas (1494) contained no large-scale indigenous states. Contemporary European accounts, e.g., Thévet (1555-57), Gândavo (1576), Cardim (1584), Soares (1587), Léry (1578), Anchieta (1596), and Vasconcellos (1663), describe the predominant native population as divided into culturally similar, rival groups that spoke varieties of the same language, in what appears to have been a dialect continuum.

> *Desde o rio do Maranhão, que está além de Pernambuco para o norte, até a terra dos Carijós, que se estende para o sul, desde a Lagoa dos Patos até perto do rio que chamam de Martim Afonso, em que pode haver 800 léguas de costa, em todo o sertão dela que se estenderá com 200 ou 300 léguas, tirando o dos Carijós, que é muito maior e chega até as serras do Peru, há uma só língua.*
> Anchieta, *Arte de grammatica da língoa mais usada na costa do Brasil* (1596)

> *Em toda esta província há muitas e várias nações de diferentes línguas, porém uma é a principal que compreende umas dez nações de índios; estes vivem na costa do*

mar, e em uma grande corda do sertão, porém são todos estes de uma só língua ainda que em algumas palavras discrepão e esta é a que entendem os Portuguezes. [...]

Todas estas nações acima ditas, ainda que diferentes, e muitas delas contrarias umas das outras, têm a mesma língua, e nestes se faz a conversão.

(CARDIM, 1584, p. 101 *apud* ROSA, 2003, p. 136)

Tupinaês é uma gente do Brazil semelhante no parecer, vida e costumes dos tupinambás, e na linguagem nõ tem mais diferença uns dos outros, do que tem os moradores de Lisboa dos de entre Douro e Minho; mas a dos tupinambás é a mais pulida; e pelo nome tão semelhante d'estas duas castas de gentio se parece bem claro que antigamente foi esta gente toda uma, como dizem os índios antigos d'est nação; mas tem-se por tão contrarios uns dos outros que se comem aos bocados [...].

(SOUZA, 1938 [1587], p. 406 *apud* ALTMAN, 2003, p. 67, n. 8)

Sadly, the opinions of native speakers of OT regarding linguistic variation have not been documented. However, given the contemporary Europeans' accounts of such linguistic uniformity, we favour Navarro's vision of a single "Old Tupi" language (henceforth OT) exhibiting mutually comprehensible regional variation (2008, p. 11-13).

Like Edelweiss (1969, p. 69-109), we find the use of "Tupinambá" as a generic term for the coastal OT-speaking peoples, following the practice of the anthropologist Alfred Métraux, e.g., Métraux (1979), and the Brazilian sociologist Florestan Fernandes, e.g., Fernandes (1948), to be inappropriate. Not only did many non-Tupinambá peoples speak the language, e.g., Potiguara, Caeté, Tamoio, Marakajá, Tupiniquim, Amoipira, etc., but also because contemporary usage does not support the generic use of Tupinambá, as we shall demonstrate below.

We also disagree with Rodrigues' treatment of the speech of the "Tupi" people of São Vicente and São Paulo as a subbranch of his Tupi-Guarani Group III, separate from the larger "Tupinambá" subgroup proposed to extend from Rio de Janeiro to Paraíba and in the Amazon (RODRIGUES, A.,1985, see also JENSEN, 1999, p 130-131; DIETRICH, 2010, p. 25).

The southern variety or varieties did exhibit a few structural differences, e.g., the suffix *-(r)amo* for the circumstantial indicative/indicative II in stative predicates and /-i/ (/ C+__), /-w/ (/ V + __) for active predicates. Rodrigues' "Tupinambá" variety exhibits uniform /-i/ and /-w/ for both predicate classes (NAVARRO, 2008, p. 121, §263). Another

southern feature was consonantal apocope, reminiscent of the neighbouring Guarani sub-branch of Tupi-Guarani (RODRIGUES, A., 1958, p. 232), classified as Group I in Rodrigues (1984/85). This is best accounted for as an areal phenomenon, for the Tupi of what would become Sao Vicente and São Paulo were in direct contact with the "Carijó" Guarani, before and during the Portuguese colonization.

Thus, Rodrigues (1996)'s regionalised nomenclature for both the precolonial autochthonous varieties and the colonial *Línguas Gerais*, i.e., *Língua Geral Paulista* (LGP), descended from his "Tupi", and *Língua Geral Amazônica* (LGA), descended from his "Tupinambá", like Nobre's *Língua Geral da Sul da Bahia* (NOBRE, 2011, 2012, 2016), are pure anachronisms. No such divisions in terms of separate varieties or languages are recognised within the *Língua Brasílica* ("Brazilian language", i.e., the codified missionary variety, henceforth LB) or LG in historical sources. Indeed, Anchieta's poetry mixes features of southern and northern varieties (NAVARRO, 2004, p. 88, n. 10, 164, n. 97). Additionally, in the 17th century, *Paulistas* were sought to work in the Jesuit missions in the colony of *Maranhão e Pará*, precisely because their fluency in their variety of LG made them especially successful at converting the Amazonian Tupinambá (BARROS, 2010). This is scarcely what one would expect from speakers of different languages.

It is therefore unsurprising that the Portuguese did not use the term "LG" initially, for in the 16th century, OT was not an interethnic lingua franca but merely the "most spoken language" of the Atlantic seaboard, as Anchieta noted in his grammar's title (*Arte de grámatica da língua mais falada na costa do Brasil*). The lingua franca role would emerge as the colonial programme expanded, especially in Amazonia.

The commonest term employed initially by the Jesuits in the Portuguese territories was *Língua Brasílica*, e.g., Araújo: *Catecismo na Lingoa Brasílica ...* (1618, 1678); Figueira: *Arte da Língua Brasílica* (1621), *Arte de grammatica da Lingoa Brasílica* (1687); Bettendorf: *Compendio da Doutrina Christãa na língua portuguesa e Brasílica* (1687). However, the term '*Língua Geral*' also appears almost simultaneously, e.g., in the letters of Manoel Gomes written between 1616 and 1621 (SANTOS, 2011, p. 10), in Antônio Vieira's writings in *Maranhão e Pará* (1653-61) and in Brazil (1681-97), and in Bettendorf's *Crônica dos padres jesuítas no Maranhão* (1687). Thus, the traditional chronology for the linguistic nomenclature in Portuguese, i.e., approximately LB (17th c.) > LG (18th c.) > NHG (19th c.), is problematic since the use of the terms LB and LG overlap considerably.

We suspect, given the dates of the councils in Spanish America mentioned above, that the change in linguistic nomenclature may, at least in part, be due to closer ties between Portugal and Spain during the Iberian Union (1580-1640). However, as yet, the precise connection eludes us.

The shift in the predominant variety of written OT from southern Vicentine/Paulista, where Anchieta began his missionary career (1553-1565), to that of the Tupinambá of Bahia and Espírito Santo, where he subsequently moved (1569, 1577-1587, 1587-96), may also be related to the Spanish practice of adopting a geographically central variety of a language intended as a lingua franca in the – fairly well-founded – belief that such a variety would be more likely to permit successful communication across a wider region than a peripheral variety (ALTMAN, 2003, p. 66-67; ALFARO LAGÓRIO, 2003, p. 49). Further investigation of Jesuit documentation is required to confirm if such a plan did indeed exist.

It is also worth bearing in mind that the other Jesuit, besides Anchieta, to describe OT, Luís Figueira, despite his role in the early Amazonian missionary projects, actually learned the language at the Jesuit College in Salvador da Bahia (1602-1607) and in the captaincy of Pernambuco at the Jesuit College in Olinda (1610-1619), as well as amongst the Caeté people at the missionary village of Our Lady of Escada (1619-1621), where he wrote his famous LB grammar. Thus, neither of the Jesuit *Artes* directly reflects the language of the Tupinambá of *Maranhão e Pará* that was LGA's direct ancestor (EDELWEISS, 1969, p. 69-81), although both grammar books, especially Figueira's, and Araújo's catechism, were employed in the Jesuit's Amazonian colleges and missions.

The Portuguese Crown only made LG an official administrative vehicle in 1689, responding to the *Regimento das Missões* (1686), which made the Jesuits responsible for overseeing the implementation of Portugal's indigenous policy. However, the practice of using LB for catechesis had been in place since the 1650s in *Maranhão e Pará* and in the *Estado do Brasil* it was already over 150 years old.

"*Língua Geral*" and "*Nheengatu*"

From the 17th to the 19th century, Portuguese writers almost exclusively use the term *Língua Geral*. One 18th-century vocabulary employs *brasiliano* "Brazilian" (EDELWEISS, 1969, p. 123-133). *Nheengatu* (henceforth NHG) is first attested in Seixas (1856), and today is widely understood as synonymous with Rodrigues (1996)'s term *Língua Geral Amazônica* (LGA).

Today, the Baré, Baniwa, and Warekena peoples of the Rio Negro call their language *Nheengatu*, as do the indigenous peoples revitalising the language on the lower Tapajós, in the state of Pará, and in the Baixo Amazonas region of Amazonas state. Older people, both native speakers and Brazilian Portuguese monolinguals, may also refer to the language as *Língua Geral*, in Portuguese, and occasionally as *gíria* "slang".

Edelweiss (1969, p. 198) states that Couto de Magalhães' *O Selvagem,* published in 1876, popularised the term *Nheengatu*, which the author claimed to be a *Tapuia*, i.e., a detribalised indigenous/mestizo, endonym (MAGALHÃES, J., 1876, p. 38). However, Edelweiss is categorical about no colonial writer employing *Nheengatu* to refer to LG or LB, and no conclusive evidence exists that the indigenous population did so either before the 19th century. Edelweiss further observes that Hartt (1875) includes only *nheenga katu* "good language/speech"*,* and *tapuya/tapya nheenga* "Tapuia language" as endonyms for the language called "LG" in Portuguese (HARTT, 1875, 1938, ex. 684, 686, 687).

The word *Nheengatu* comes from the OT roots *nhe'eng* /jẽʔẽŋ/ "speech", "language", "vocalization", and *katu* /kaˈtu/ "good", "well". OT morphemes can only exhibit CVC syllables in word final position. Compounding causes apocope whenever a consonant-final morpheme precedes a consonant-initial morpheme, i.e., C → Ø /__+C, e.g., /jẽʔẽŋ/ → [ɲẽʔẽ]. Tupi-Guarani (henceforth TG) nasal harmony (ESTIGARRIBIA, 2020, p. 39-44) means the phonologically nasal root /jẽʔẽ(ŋ)/ causes the phonologically oral root /kaˈtu/ to exhibit /k/'s nasal allophone, i.e., /k/ → [ŋᵍ] / V[+nasal] + __ V[-nasal], i.e., *nhe'ẽ-ngatu* / jẽʔẽ-kaˈtu/ [ɲẽʔẽŋgaˈtu].

Despite Edelweiss' justifiable reticence, it is curious that the word *Nheengatu* cannot come from any modern or attested 19th-century variety of that language, since none possess the requisite nasal harmony, exhibiting only /jẽʔẽ kaˈtu/ "speak well" and /jẽʔẽŋa kaˈtu/) "good speech/language", "well-spoken" (HARTT, 1938; COSTA, 1909 EDELWEISS, 1969, p. 201, n.17). However, since both Seixas (1856) and Magalhães (1876) independently record the nasal variant twenty years apart and claim it to be autochthonous, it seems unlikely that either writer coined the expression. Thus, it is still possible that the term *Nheengatu* may be much older than its first written attestation.

Significantly, *nheengatu*'s antonym, *nheengaíba* (< /jẽʔẽŋ/ "speak/speech" + /aˈiβ/ "bad" + /-a/ "NMZR") is amply documented from the 17th century onwards, in Amazonia, as a generic term for non-LG-speaking indigenous peoples, often called *línguas travadas* ("blocked/stuck/locked tongues") in contemporary Portuguese, e.g., Vieira in the 1650s (*Cartas*, v. 1, p. 556, *apud* EDELWEISS, 1969, p. 198, n. 5). *Nheengaíba* was also the LG exonym for a large, non-TG confederacy on the island of Marajó that exhibited great animosity towards the Portuguese before being missionized and converted in 1658-59.

Therefore, when Edelweiss affirms that *[...] Se [o nome] nheengatu não chegou a usar-se, foi simplesmente porque todos os bemfalantes eram parentes e tinham nome próprio conhecido por todos, que tornava supérfluo qualquer denominação global; [...]"* (1969, p. 198), he overlooks the fact that those who wrote in and about LG, in the 18th and 19th centuries

were addressing non-native speakers who were largely uninterested in native-speakers' opinions. Moreover, as we shall develop below, by the mid-18th century, "well-spoken" indigenous and mestizo Amazonians were far from being closely related to one another.

Theories about the emergence of the *LGA*

Most previous analyses of LGA tend to associate its emergence with particular social groupings[2]. Such theories are not new. The Jesuit missionary João Daniel wrote that:

> *Nesta língua [Tupinambá] se composeram ao princípio pelos primeiros missionários jesuítas o catecismo, e doutrina; e a reduziram a arte com regras, e termos fáceis de se aprender. Porém, como os primeiros, e verdadeiros topinambares já quase de todo se acabaram, e as missões se foram restabelecendo com outras mui diversas nações, e línguas, se foi corrompendo de tal sorte a língua geral topinambá, que já hoje são raros, os que a falam com a sua nativa pureza, e vigor; de sorte, que já os mesmos índios não percebem o catecismo, nem os que estudam a arte se entendem com os índios especialmente no Amazonas, como muitas vezes tem experimentado, e confessado os mesmos missionários, e índios, de sorte está viciada, e corrupta que parece outra língua diversa; mas a qual é a que se usa em todas as missões portuguesas do Amazonas, e a que aprendem as novas nações, que vão saindo dos matos, e a que estudam os missionários brancos, que tratam com índios não com regras, e preceitos da arte, mas pelo uso, e trato dos mesmos índios [...].*

> Daniel, *Tesouro descoberto no máximo Rio Amazonas*
> (2004 [1757-76], v. 2, p. 365)

Edelweiss (1969, p. 156-157) stressed the "simplification" of the 'pure', i.e., precolonial, Tupi and the largely unaltered LB into what he called *Brasiliano* through acculturation and miscegenation on a massive scale. Similarly, Rodrigues (1996) emphasizes miscegenation, picking out the *mameluco* class as instrumental in developing and generalising LGA[3].

---

2 Mattoso Câmara is a notable modern exception to the contact-based account, in attributing the LG's emergence to a deficient analysis of the original Tupi by the Jesuit grammarians (1977, p. 101). This notion is not original to Câmara. Francisco Xavier de Mendonça Furtado, governor of Grão-Pará (1752-1759) and the prime mover in implementing the anti-Jesuit and anti-*LG* policies of his brother, the future marquis de Pombal, accuses the Jesuits of inventing the LG. Rosa (1990, 1992) convincingly disproves this hypothesis.

3 The term *mameluco* is usually used in modern Brazilian Portuguese to refer to anyone of European and Amerindian ancestry. However, as Monteiro (1994:166-67) emphasizes, in the 17th century, *mameluco* referred specifically to those whose 'white' paternity was publicly recognized. Such individuals enjoyed greater social freedoms. Unrecognised offspring of 'white' men and 'indian' women were called *bastardos,* a term that extended its range over the course of the 18th century, making *mameluco* obsolete.

Dietrich (2014, p. 594) contrasts Nahuatl as an "authentically" indigenous lingua franca in Mexico with Rodrigues (1996)'s "mestizo" LG in Brazil. Freire (2011) gives a central role to language contact. Lee (2005), Nobre (2011a, 2011b, 2012a, 2012b, 2015, 2016), and Oliveira, Zanoli and Módolo (2019) base their discussions on pidginization and creolisation. All associate miscegenation with structural change that caused LB/Tupinambá/OT to *evolve into* LG.

Above, we have shown that the names LB and LG appear in writing at almost the same time in the 17th century. We argue below that there is no direct relationship between the patterns of metalinguistic change and observable structural evolution. Nor can contemporary usage link LG to any particular social category in Portugal's South American colonies. Instead, the key factor was the language's function in colonial society.

Function, not structure or group

As Wright (1982, 1994, 2002) has stressed for those who seek to locate the moment when "Latin" became "Romance" in terms of structural modifications rather than because of some change in speakers' perceptions brought about by the development of novel scripto-linguistic norms, we argue in a similar vein that those who try to distinguish "LG(P/A)" and "OT"/"Tupinambá" on the basis of structure do so largely in vain, for it seems that in the seventeenth century in the colony of *Maranhão e Pará*, any indigenous group speaking a TG language that was broadly intelligible with what the Luso-Brazilians, their mestizo descendants and their Tupinambá vassals and allies spoke qualified as being "*índios de língua geral*".

We do not deny that sociocultural and demographic changes and language contact and diversity were crucial factors in promoting language evolution in the mission villages and the colonists' households and farms. However, the way that the term LG was used in the 17th and 18th centuries shows that it was above all a functional category. Although typically referring to OT and varieties derived from it, the term LG was not limited to that language.

For instance, already in 1616, Manoel Gomes wrote: "*Ha muitos tapuyas [i.e., non-OT-speaking peoples] de muitas nações, das quaes quatorze fallão a língua geral dos Tupynambás, que é quase commum no Brazil. [...]*" (GOMES, 1904, p. 334). Bettendorf's *Crônica da missão dos padres da Companhia de Jesus no Estado do Maranhão* (1698) states that the Guajajara of the Pindaré river in the present-day State of Maranhão are "*de Língua Geral*" and "*parecidos com os Carijós*", i.e., a Guarani-speaking TG people located thousands of miles to the south (*apud* LEITE, 1943, p. 185).

Even before their missionisation, the "*multidão de gente*" living on the lower Tocantins and Araguaia were also described as "*quase todos da Língua Geral*" (BETTENDORF *apud* LEITE, 1943, p. 316), despite these groups including many non-Tupinamabá, e.g., amongst those still living in the region today, the Guajajara and Tembé, Urubu Kaapor, Guajá, Anambé, Takunyapé, Turiwara, and Amanayé, who live around the Maranhão/Pará state line, and the Tocantins Asurini, Tocantins Suruí, and Parakanã. The Emerillon and Wãjãpi formerly inhabited the lower Tocantins before undertaking a large-scale migration to what is now Amapá State and French Guiana in response to Portuguese predation. All these peoples belong to quite separate sub-branches to OT within the Tupi-Guarani branch of the Tupian language family (JENSEN, 1999, p. 126, 130-132; RODRIGUES, A., 2010; GALÚCIO *et al.,* 2015; MICHAEL *et al.*, 2015; MELLO; KNEIP, 2018). Even allowing that four centuries of diachronic change may have obscured greater proximity in the past, none of these languages can be considered structurally identical to any OT variety: LB, LG, or NHG (see also CABRAL; OLIVEIRA, 2013). Nevertheless, in the second half of the 17th century, such peoples were referred to as being speakers of LG, even before missionisation.

Moreover, the Jesuit missionary João Daniel (1722-1776), whose *Tesouro descoberto no Rio máximo Amazonas* (DANIEL, 1976) is cited by all who seek proof of a Tupinambá pidginization or creolisation amongst the missionized *Tapuia*, actually reveals that what he calls the "true", "Tupinambá LG", i.e., the codified LB, was essentially a high diglossic code that only the missionaries maintained - non-natively - and possible relic groups of "Tupinambá" spoke natively: '*[...], que já hoje são raros, os que a falam com a sua nativa pureza, e vigor; [...]*'. Thus, Daniel's comment, quoted above, reveals that the *Tapuia* actually speak the colonial vernacular of the non-missionized, which had evidently undergone considerable restructuring with regard to the codified variety, given Daniel's comments: '*[...] já os mesmos índios não percebem o catecismo, nem os que estudam a arte se entendem com os índios [...] como muitas vezes tem experimentado, e confessado os mesmos missionários, e índios*'. It is important to observe that Daniel does not claim that a radical difference existed between the speech of the missionized *Tapuia* and that of the rest of the colony. In fact, he states that *everyone* in the missions speaks the "corrupt" LG in day-to-day interactions: '*[...] é a que se usa em todas as missões portuguesas do Amazonas e a que aprendem as novas nações, que vão saindo dos matos, e a que estudam os missionários brancos, que tratam com índios [...] pelo uso, e trato dos mesmos índios [...]*'. Note also that both the H and the L varieties are referred to as "LG" by Daniel, i.e., he does not go so far as to classify them as actual separate languages, although he does claim the "corrupt" LG "appears to be another language" (*[...] parece outra língua diversa*).

Daniel's comments regarding the Amazonian colonial vernacular appear to be in a similar vein to the bishop of Pernambuco's observations regarding the *Paulista* Domingos Jorge Velho: "*este homem é um dos maiores selvagens com que tenho topado: quando se avistou comigo trouxe conisgo língua, porque nem falar sabe, nem se diferença do mais bárbaro tapuya, mas que em dizer que é Cristão*" (*apud* MONTEIRO, 1994, p. 164). In fact, the bishop's prejudice was groundless: Domingos Jorge not only spoke Portuguese but was literate. Similarly, Mufwene (2003, p. 118-123) highlights the risks of employing mutual intelligibility as a determinant of separate language status for creoles, for it depends overmuch on which variety is compared to the alleged creole.

As a non-native speaker of "true", "Tupinambá" LG, i.e., LB, Daniel takes the codified variety as his standard and castigates the "corrupted and error-strewn" (*corrupta e viciada*) speech of the *Tapuia*, despite his employing it every day in all probability outside his catechesis classes. However, missionaries from other orders, e.g., Veloso, Avronches, Onofre, and others, continued to compile vocabularies and write in a structurally cohesive and typologically recognisable TG language they call "LG" into the second half of the 18th century (RODRIGUES; CABRAL, 2011; RODRIGUES, A., 2010; MONSERRAT, 2003; MONSERRAT; BARROS; MOTTA, 2010; EDELWEISS, 1969), see also Rosa (1990, 1992), the manuals of Navarro (2004) and Lemos Barbosa (1956). Even when someone, probably Anselm Eckhard (1721-1809), took Bettendorf's *Compêndio da doutina crista* (1687) and had it "*traduzida em língua irregular, e vulgar uzada nestes tempos*" (ms nº 1.089, Universidade de Coimbra) the changes made are slight (EDELWEISS, 1969, p. 138; MONSERRAT, 2003, p. 185-194; MONSERRAT; BARROS; MOTA, 2010; FREIRE, 2011, p. 121; RODRIGUES; CABRAL, 2011, *passim*).

Some issues with recent hypotheses

Several studies of the LGA draw on pidginization and creolization to explain the structural difference between LB/LG (17th and 18th c.) and NHG/LG (19th c.). The most detailed recent treatments of LG-as-pidgin/creole are Vieira, Zanoli, and Módolo (2019) (henceforth, VZM) and Nobre (2016), which we discuss here.

VZM propose that LG begins as "Tupinambá", which is a pidgin. This pidgin expands into the "*Língua Brasílica*" and then creolizes to become the "*Língua Geral do Brasil*" (2019, p. 330). Sadly, this vision is utterly incompatible with the known historical data.

Firstly, as we have shown above, LB is amply documented as the vehicle of Anchieta's poetry, plays, and catechism, published by Araújo (1618, 1686). It is the language of Figueira's *Arte* and of Bettendorf's *Compêndio.* As Rosa (1990, 1992) has convincingly demonstrated,

the language described in Anchieta's and Figueira's grammars, i.e., LB, cannot be classed, either structurally or functionally, as anything other than a fully-fledged TG language, comparable in every sense to those described by 20th century linguists, e.g., Kamayurá (SEKI, 2000), Guajá (MAGALHÃES, M., 2007), Paraguayan Guarani (ESTIGARRIBIA, 2020), Tapirapé (PRAÇA, 2007), among many others (see also RODRIGUES, A., 2010; NAVARRO, 2008; LEMOS BARBOSA, 1956).

The locus of acquisition of LG is another issue. VZM refer to the *Tapuia* learning LG *"nos colégios dos jesuítas"* on eleven occasions (2019, p. 320, 326, 329). However, this factually inaccurate, for few – if any – *Tapuia* ever saw the inside of a Jesuit school as students. *Tapuia* LG acquisition occurred through unstructured immersion in the mission villages, royal factories, and in the colonists' residences and farms. The LG studied at Jesuit colleges was the codified variety, i.e., LB, which, by the 1750s, was essentially the preserve of the missionaries.

Moreover, we believe a generalised pidgin could not have endured any length of time on a colony-wide scale, even in the later, multi-lingual, multi-ethnic mission villages, because the forced labour-service gave missionized adolescent boys and men ample access to the colonial vernacular norm in the form of annual six-month immersions. Similarly, household slaves of either sex were in constant contact with native LG speakers. Thus, if they survived the epidemics, general mistreatment, and dangerous activities, an initially pidgin-like speaker could reasonably be expected to gain fluency constantly over the course of their lifetime of immersion and approximate ever closer to the community's vernacular norm. Such an individual's descendants, if they had any that survived (see below), would have grown up in constant contact with native LG speakers and, thus, would have received ample primary linguistic data, favouring native or near-native competence in the local LG variety, probably in conjunction with knowing other indigenous languages.

Only the kind of cultural prohibition that João Daniel reports for married Nheengaiba women would realistically limit the use of LG amongst the missionized. Even in this case, however, the women reportedly knew LG, having grown up surrounded by the language in their mission villages, and merely refused to exercise competence in it once they married, supposedly forbidden to do so by their husbands. Daniel states that the missionaries resorted to acts of violence in order to force the women to speak LG in the confessional (DANIEL, 1976, I, p. 272, *apud* BARROS, 2003, p. 95-96).

The data, therefore, seem to point to a restructured vernacular variety, rather than the emergence and generalisation of a pidgin or subsequent basilectal creole that expanded functionally and complexified structurally through innovative grammaticalization. That is

to say, there have always been native speakers of LG at every stage in its history. What seems to have changed over time is the constitution of the communal norm, which lost or partially altered some features present in the ancestral system, but preserved many others, rather than being a wholly novel communal system built up out of very incompletely acquired linguistic data, as most creoles are.

Although doubtless much influenced by substrate and other L2 features, rather like vernacular Brazilian Portuguese, nonetheless, mid-18th century vernacular LG would still have shown much continuity with previous stages of the language, as even the 19th and 20th century language still does, since there has never been any interruption in its natural, intergenerational transmission (see also RODRIGUES; CABRAL, 2011, p. 635-636). We shall develop this vision further in later sections. We now pass to Nobre (2019).

Nobre's discussion of LG (2019) is historically detailed and bibliographically dense. However, his linguistic analysis suffers from several severe shortcomings, which we outline below.

Firstly, we question the model of creolization. Specifically, we note Nobre's quotation of Lucchesi, which states that *'(...) a referência nas situações típicas de crioulização seria a proporção de pelo menos dez indivíduos dos grupos dominados para cada indivíduo do grupo dominante"* (LUCCHESI s/d: 25, *apud* NOBRE, 2016, p. 33; see also BAXTER; LUCCHESI, 2009). Below, we show that, for Portuguese Amazonia, the probable ratio of *Tapuia* to non-*Tapuia* stood at around 3:1. Neither the number of Jesuits, amongst whom many were not even Portuguese, let alone native speakers of LG, nor the number of Portuguese colonists (HONAERT *et al*., 2008, p. 83, *apud* NOBRE, 2016, p. 34) should be confused with the numbers of LG-speakers in the colony, because the large number of *mamelucos* and *bastardos* would have greatly increased the proportion of LG speakers, particularly in the earlier phases of colonization.

Nobre (2016, p. 34) reuses Lucchesi's model, which restates a traditional relationship progressing from a jargon to a pidgin, which Nobre defines according to Holm (2000, p. 5, *apud* NOBRE, 2016, p. 35, n. 9), adding that a pidgin may expand to respond to greater functional demand. In this classic model, a pidgin becomes a creole when children acquire the pidgin used non-natively by the community as a primary means of communication (LUCCHESI s/d, 3 *apud* NOBRE, 2016, p. 35).

However, the missionized and enslaved Amazonian populations do not appear to have grown vegetatively after the early years of missionary expansion, being maintained rather by replacing the dead or escapees with new "descents" (*descimentos*) of indigenous peoples from the interior or by "ransoms" (*resgates*), i.e., purchasing alleged potential

victims of cannibalism to enslave. Thus, the primary restructuring context for LG must have been adults learning LG from other adults, as Mufwene claims was the predominant creole-formation context in the Caribbean (1993, p. 39-45, 131).

Nobre's use of Freire (2003, 2004) regarding the indiscriminate mixing of indigenous peoples in mission villages (NOBRE, 2016, p. 37) may be problematised in our Expansion Period (1653-1660/70, see below), when early mission villages were often established with a sole or heavily predominant indigenous nation, which often gave its name to the village, e.g., *aldeia dos Tupinambá de cima*, *aldeia dos Tupinambá de baixo*, *aldeia dos Bocas*, *aldeia dos Nheengaíbas*, etc. (LEITE, 1943). Later "descents", however, did not respect ethnolinguistic differences, as Eckert recorded (*apud* EDELWEISS, 1969, p. 114, n. 6).

Nobre has claimed the LGA should be classed as a 'mesolectal' creole (2016, p. 48; see also 2011a, 2012b) for exhibiting only limited structural divergence from the lexifier. However, the categories Nobre (2016) proposes that LG has lost on the basis of his analysis of Magalhães' 19th-c. description of Nheengatu are deeply problematic. For instance, it is not at all reasonable to define OT as a language with morphological cases, as Rodrigues (2010) does.

Rodrigues proposes five flexional cases for Tupinmabá: "argumentative" (*-a*), "punctual locative" (*-pe*), "diffuse locative" (*-βo*), "partitive locative" (*-i*), and "translative" (*-(r) amo*). The total absence of morphological case distinctions for core argument functions, e.g., subject, object, recipient, benefactive, possessive, etc., is deeply suspicious (see also SCHMIDT-RIESE, 2012, p. 151-156). Moreover, the lack of morphological cases in any non-TG Tupian language is also very conspicuous.

Consequently, /-a/, /-pe/, /-βo/, /-i/, /-(r)ãmo/ are far better accounted for as nominaliser (*-a* ~ *-Ø*), since only ever creates nominal arguments out of inherently predicative roots, and four unstressed, enclitic postpositions. The lack of "argument case" /-a/ before the three locative forms, which seems to underpin Rodrigues' classification of them as case-inflexions, can be better accounted for as the result of their conversion into enclitic postpositions, as anaptyxis with /i, ɨ/ is the rule in OT for enclitic particles with initial consonants that follow consonant-final roots, e.g., /-(ɨ)ne/ FUT, /-(ɨ)βa'e/3.REL, /-(ɨ)'pɨr-a/ PAT (NAVARRO, 2008, p. 138, §182; p. 223, §307; p. 268, §374).

The final "translational" case, i.e., /-(r)ãmo/, does not exhibit either /-a/, being vowel-final. Nor is there an epenthetic /ɨ ~ i/, as it belongs to a "pluriform" class in which /r-/ appears initially when the root is vowel-final and is absent following consonant-final stems. Other examples in OT include the stressed postposition /-(r)i're/ "after(wards)" and unstressed /-(r)(e)me/ "at the time of", which Rodrigues significantly does not

classify as cases, perhaps owing to their temporal meaning. Rodrigues leaves /-(r)i're/ as a postposition, but analyses /-(r)(e)me/ as "subjunctive mood" (2010, p. 29, 31), despite it also exhibiting indicative temporal function. Seki (2000, p. 127, 130-131, 193-194) treats both these inflexions as moods: "posterior" and "subjunctive", respectively[4].

Thus, the supposed loss of case-inflexion between OT/LB and 19th century LG is best accounted for as a levelling of a tiny minority of exceptional forms, rather than the wholesale elimination of a fully-fledged flexional case system. Locative /-pe/ becomes a free stressed postposition: *upé*. Hart (1872, p. 314-315) records possible cases of /-βo/ fused with personal pronouns to express dative/benefactive function, e.g., *ixéu < ixébo* / i'se=βo/, *ixébe* /i'se=pe/, *in(d)éu < endébo* /e'ne=βo/, *endébe* /e'ne=pe/, *yan(d)éu < îandébo* / ja'ne= βo/, *îandébe* /ja'ne=pe/, *penemo < peẽmo* /pe'ẽ=βo/. These forms are do not occur in modern NHG. The other enclitic postpositions have been lost. The /-a/ nominaliser fuses with the lexical root as part of a generalised process of paragoge, e.g., /'sem/ "leave" > /'semu/, /'sik/ "arrive" > /'sika/, /jor/ "come" > /'juri/, /'ok-a/ "house" > /'uka/, /kop-i'saβ-a/ "place for cultivation" > /kupi'ʃa/, /ja'w-ar-a/ "jaguar" > /ja'wara/ "dog", port., /'pai/ "father" > /'paja/, port., /'mãi/ "mother" > /'maɲa/.

Similarly, Nobre's analysis of the supposed erosion of OT TAM-marking inflections (2016, p. 39-42) is deeply flawed because =*ne* FUT/IRREALIS is actually a clause- or sentence- enclitic particle in OT (NAVARRO, 2008, p. 138, §182). Thus, 19th-century LG has merely exchanged a TAM particle (=*ne*) for a TAM adverb (*kuri*), which was always available for time marking in OT as an adjunct. Therefore, no loss of inflexion has occurred. Additionally, the tendency for Magalhães to include subject pronouns in his presentation of NHG conjugations (MAGALHÃES, J., 1876) belies the real pro-drop profile of Nheengatu then and now (CRUZ, 2011, *passim;* MAGALHÃES, J., 1876, *passim*), undermining Nobre's comparison with other creole languages (2016, p. 42).

Following the creole typological profile presented in Bakker et al. (2011), Bakker (2014) proposes four typological features for identifying creoles (p. 451-52): 1) indefinite article derived from "one", 2) no tense/aspect inflections, 3) a preverbal negative particle, and 4) a possessive "have" verb. These features are all found in modern NHG: *yepé* "one", "a(n)", absence of TAM-inflexion, *ti/umbaá* as NEG, and -*riku* for "have".

---

4 We cannot present our full analysis of the OT "cases" here for reasons of space but shall address the issue in another paper that is already in preparation.

However, NHG's parent, Tupinambá OT, also exhibited no tense or aspectual inflexions, as is the norm for TG languages[5]. So-called "nominal tense" is actually manifested through two stative roots /ˈpʷer/ "be an ex/a former …" (*[ser]* um *antigo* …) and /ˈram/ "potential", "will become" (*há de ser*) (ANCHIETA, 1595, p. 33v, *apud* NAVARRO, 2013, p. 408, 426). These roots may stand alone as stative predicates in OT, e.g*., xe pûer* "I was/am altered", "I am not myself"*, xe ram* "I have potential", "I can/may be". More often, they modify some other root in a compound noun, e.g., *ybyrapûera* /ɨβɨˈra-ˈpʷer-a/ tree-denatured-NMZR "former/ex-tree", i.e, "non-prototypical (dead, fallen, rotten, etc.) tree"[6], *ybyrarama* /ɨβɨˈra-ˈram-a/ tree-potential-NMZR, i.e., "potential tree" (a shoot, a seedling). Thus, the lack of tense inflection in Nheengatu is not an innovation but a retention from the uncreolised parent language.

 The OT etymon of NHG *yepé* "an"/"one" is /(o.)ieˈpe/, which also means "sole(ly)" or "once", rather than specifically being a numeral[7]. Significantly, the ordinal "first" is not related to /(o.)ieˈpe/ but is *ypy* /ɨˈpɨ/, an intransitive root meaning "begin/start/initiate". The negative *ti < niti(u) < nditybi* "there is not" is indeed an innovation via grammaticalization. "Have" (*-riku*) comes from OT *-(e)r(a)-ikó* -APL-exist "to be with something". It is not an innovation that is specific to Nheengatu but rather seems to be a pan-Tupi-Guarani solution to the Romance transitive "have"-verbs, cf., *-(gue)reko* "have" in Paraguayan Guarani, which no one claims to be a creole.

We therefore have difficulty accepting a creole status for LG or NHG based solely on comparisons with LB (of which more below). We also have reservations about Bakker (2014)'s use of a checklist of morphological features, given the typological profile of TG regarding TAM-marking, numerals, and the emergence of "have" verbs from ancestral morphology that remains productive in most modern TG languages, although not in modern NHG.

Given our reservations about earlier analyses that have treated LG and NHG as creoles, we now apply an as-yet unused model of creolisation to the historical data to suggest how a similar result might have arisen by a different evolutionary route.

---

5 Like most TG languages, Old Tupi did have a prefix-marked "permissive" mood, (*ta-*). It did not survive into Nheengatu.

6 In modern TG languages, including NHG, /ˈpʷer/'s cognates bear an added sense of "denatured", "ex-prototypical", that can also be read into OT usage.

7 The other four cardinal numerals in Old Tupi are verb roots or phrases: *mokõî* "geminate", "make a pair" (2), *mosapyr* perhaps "make a point/peak" (3), *(oîo)irundyk* "(they) arrive accompanying (each other)" (4), *mbó, xe pó* "(human) hand", "my hand".

Creoles and creolisation, according to Mufwene

Mufwene (2003, p. 157, p. 195) claims that languages are complex adaptive systems formed from a loose amalgam of idiolects with a certain 'family resemblance' that are never wholly identical. Crucially, language learners select variants in a piecemeal, blending manner from a *feature pool* created by extant idiolects they are in contact with, such that no idiolect is ever wholly identical to those that contribute material to the pool or to the idiolects that arise from selecting features from the pool (MUFWENE, 2003, p. 9; see also LASS, 1997, p. 12). Thus, synchronic and diachronic feature replication is always approximate in this model, causing extensive idiolectal and communal polyploidy, i.e., individuals and communities exhibit a multiplicity of synchronic variants (2003, p. 18, 2008, p. 19). Moreover, linguistic heritability is Lamarckian, for individuals are fundamentally free to incorporate and exploit any linguistic resources that they possess and encounter in others, thereby making borrowed features available to subsequent learners (2003, p. 32). Feature transmission is, therefore, as much horizontal amongst peers as it is inter-generational.

Individuals also constantly accommodate their speech to that of their interlocutors, aiming at a "fuzzy" target that will allow satisfactory communication (2003, p. 192). This necessarily approximative replication drives constant restructuring both synchronically and diachronically (2003, p. 26, p. 126-127, 2008, p. 18). Features may be preserved, or innovations can arise via reinterpretations of existing expressions or as genuinely unprecedented creations. Communal norms therefore evolve as emergent phenomena, aligned via the 'invisible hand' as individuals imitate one another`s spontaneous behavioural patterns to enhance effective communication in the moment (MUFWENE, 2008, p.3, p. 60, p. 62-68; see also KELLER, 1994). Favoured strategies become entrenched, i.e., normalised. However, speakers' immediate individual goals will frequently be at odds with coarser-grained collective tendencies (MUFWENE, 2003, p. 26).

The emergence and propagation of an innovation both individually and in a community depend on the highly complex interaction of the components of a speaker's linguistic repertoire and disposition ("internal ecology") with pressures from the extra-personal milieu, i.e., society, culture and technology, and the physical environment ("external ecology"). For instance, in a language contact situation, structural congruence is an important factor in favouring which elements will be selected from the feature pool (MUFWENE, 2003, p. 115)[8].

---

8 However, given an external ecology that exhibits social prejudice against a particular structurally congruent variant, perhaps because of it's being associated with some stigmatised group, the conscious desire to avoid discrimination may lead to avoidance, thus generating shibboleth taboos and potentially stimulating hypercorrections in the opposite direction.

Another important – and hotly disputed – strand of Mufwene's hypothesis is that creoles and pidgins are not qualitatively different from other natural languages and do not constitute a special linguistic type (see also ABOH, 2015, ABOH; DEGRAFF, 2017)[9], but emerge just as languages traditionally considered not to result from contact do (MUFWENE, 2003, p. 134). At the root of this statement is the observation that in ecologies considered monolingual, much diversity still occurs dialectally, sociolectally, and, as ever, between idiolects. The greater similarities between new and extant varieties of non-creole languages are because the feature pool constituted is less diverse and the ecology is less amenable to admitting exceptions to community norms, thus the margin for divergence from extant norms is smaller.

On the other hand, Mufwene argues that creoles' perceived structural divergence from their lexifiers emerges when the prevailing ecological conditions do not inhibit people from selecting variants from an especially rich feature pool that are less common or indeed absent from the contributing varieties. One such situation would be when a large number of adult, non-native speakers learns a language from another group of non-native speakers, often while undergoing language shift, or from a much smaller group of native speakers, swamping any original norms (2003, p. 76, 104, 106-107).

Another important strand of Mufwene's theory of creole formation is the lack of a pidgin ancestor. In the traditional model, a contact jargon becomes a pidgin as contact and usage patterns stabilise. Creolisation occurs when children acquire the pidgin natively, causing further large-scale, innovative structural expansion, typically through grammaticalization. As mentioned above, we do not see evidence favouring this route to restructuring in the emergence of LGA.

Instead of such "rapid, early basilectalization", followed by subsequent complexification and, finally, decreolisation, Mufwene proposes "late basilectalization" (2003, p. 10, 38-67), when the ecology, often as a result of demography, changes to favours greater structural divergence from the colonial koine vernacular that had arisen alongside the creole. In the context of Atlantic creoles, the typical shift towards basilectalization occurs when the economic model changes. An initial, homestead settlement colony, in

___

9 For an opposing view of pidgin and creole genesis and evolutionary relations, see, for instance, McWhorter (2020), Bakker et al. (2011) Bakker (2014). We concur with McWhorter (2020) that pidginization is a real linguistic process and that it is exceptional, inasmuch as it represents the extreme end of language contact and L2+ acquisition amongst adults under extreme deprivation of primary linguistic data spoken natively. Nevertheless, we argue that such a model of construction of a new language through wholesale innovative grammaticalization is not applicable to *Língua Geral Amazônica*. Rather, Mufwene's model is better suited to explain the divergence of a native-speaker community from their ancestral structural prototype under the influence of large numbers of non-native speakers.

which a small number of incomers have ample access to the local vernacular, gives way to a plantation economy, in which very large numbers of non-native adults are obliged to acquire competence in the local language through unstructured immersion, usually mediated by other adults, many of whom are also non-native speakers of the target variety.

In the following section, we apply Mufwene's "ecological" model of language evolution, e.g., Mufwene (2003, 2008), to analyse the emergence of LG in Amazonia.

Língua Geral as a "Mufwenean Creole"

Phase one: the founder effect of the Tupinambá koine

The first population to occupy a territory often influences the trajectory of a contact situation disproportionately, as later arrivals usually align themselves with the established norms they encounter. This kind of Founder Effect (MUFWENE, 2003, p. 25-27) can be clearly identified in the case of the LGA, which continues to exhibit a very markedly OT lexicon and grammatical typology.

The foundation and development of Portugal's Amazonian colony occurred under different conditions to the invasion and colonisation of the southern coast and interior over the previous century. Physical communication between the captaincy of Maranhão and the Amazon and the rest of the Atlantic coast was always difficult, hence the decision to separate the regions administratively in 1621 into the *Estado do Maranhão*, composed of the captaincies of Maranhão, Pará, and Ceará, with its capital at São Luís, and the *Estado do Brasil*, comprising the other captaincies, ruled initially from Salvador da Bahia and, from 1763, from Rio de Janeiro. In 1652, the captaincy of Ceará was transferred to the *Estado do Brasil* and in 1654 the remaining Northern captaincies were renamed *Maranhão e Pará*.

We identify the linguistic foundation period in Portugal's Amazonian colony as running from the initial capture of São Luís in 1615 and the foundation of Belém in 1616 to 1660, the year of the terrible 'pestilential catarrh' (HEMMING, 1995, p. 343) that opens the period of catastrophic epidemics.

The early 17th-century French colonists D'Abbéville and Évreux describe Tupinambá villages stretching from the island of São Luís far up the southern branch of the Amazon, perhaps as far as the Pacajá river (ABBÉVILLE/MARQUES, 1874, p. 220-221, EVREUX/DENIS, 1864, p. 27, EVREUX/OBERMEIER, 2014, p. 32-33).

The French chroniclers and the first Portuguese accounts note that the predominant Amazonian OT variety that would form the basis of the LG was a recent phenomenon, the result of migrations of Tupinambá that had been occurring from the north-eastern coast starting around 1530 as the Portuguese occupation intensified in that region. However, Bettendorf records that large groups of Tupinambá also inhabited the Tocantins and Araguaia rivers. It is unclear whether these groups were also recent immigrants from the coast or whether they were populations from which the coastal Tupinambá had separated when they first migrated across the *sertão* to the Atlantic. Other north-eastern OT varieties were brought by the hundreds of indigenous warriors who accompanied the Portuguese soldiers from the captaincies of Paraíba and Pernambuco, and with the slaves brought by the earliest settlers. The European Portuguese often knew OT as a second language, and their mestizo descendants typically spoke the language natively. Although expeditions upriver beyond the fort at Gurupá at the mouth of the Xingu were rare in the early years of the colony (SARAGOÇA, 2000), slavers frequently raided into the interior along the rivers Tocantins and Araguaia, where the populations were predominantly Tupinambá or spoke other TG languages. Tupi-Guarani is a very cohesive branch of Tupian, despite the large number of languages it contains. It is reasonable to compare TG with branches of IE such as Romance, Germanic, or Slavic, which have high degrees of intelligibility between varieties. The kind of interlanguage that is observed between such IE varieties is also a distinct possibility for TG.

Thus, the initial external ecology favoured the development of an OT koine, with local Tupinambá speech as the predominant variety. Given the considerable structural proximity of the OT varieties in contact initially, communication would not have been at all difficult between the Portuguese, their descendants, and indigenous allies, and the Amazonian Tupinambá. Therefore, the missionaries' grammars and catechisms, composed in the South, could be (and were) easily transferred to serve the Amazonian mission villages. The former texts were intended for the missionaries to acquire a working knowledge of the language, the latter for practical use with the villagers. The proximity of the TG varieties involved is highlighted by the head of the Maranhão mission seeking native LG-speakers from São Paulo, almost three thousand kilometres to the south, to work in Amazonia in the 1650s, as their language skills were much appreciated (BARROS, 2010).

The socioeconomic organization of *Maranhão* (1621-1654), with two small proto-urban nuclei, the indigenous villages, and the colonists' *sítios* (SARAGOÇA, 2000) resembles the "homesteads" in "settlement colonies", identified as the typical setting for forming colonial koines (MUFWENE, 2003, p. 34; see also CHAUDENSON, 1979, 1989, 1992). Any incomers, from Europe or from other indigenous groups, lived and worked in

close proximity with native Tupinambá speakers for extended periods, as was the case for enslaved Africans in the homestead phase of settlement colonies in North America and the Caribbean, where European languages were the targets (MUFWENE, 2003, p. 9), and indeed, in Brazil. These conditions would have given incomers ample opportunity to acquire the local vernacular reasonably well, depending on their individual linguistic abilities. Subsequent generations raised in such an ecology – and, prior to the great epidemics, the survival of children is much more likely – can reasonably be expected to have been native speakers, since contemporary sources indicate sufficient primary linguistic data was easily available.

> *Os moradores nascidos no Pará, sabem primeiro fallar a língua dos índios do que a portuguesa: por que como não bebem, nem se crião com outro leite mais, que com o das indias, com o leite bebem tambem a lingua, nem fallao outra, senão depois de andar alguns annos na escola e tratarem com os portugueses, que vão de Portugal, e com esta comunicação.*

> Jacinto de Carvalho (1677-1744)
> (*apud* MORAES, 1860, p. 322-323 *apud* BARROS, 2010, p. 150)

Similar comments exist for the south in the 1500s, e.g., Gândavo (1576 [1964, p. 33; *apud* LEITE, 2013, p. 17), Cardim (1584, p. 101; *apud* ROSA, 2003, p. 136), and later, in São Paulo, e.g., Fonseca (1752 [1932], p. 26, 40, *apud* BARROS, 2010, p. 148, 149).

Phase two: the early missions (1653-1660)

From 1653, the celebrated Jesuit preacher Antônio Vieira dynamized missionary activity in Maranhão and its subordinate captaincy, Pará, bringing down many LG-speaking Poquiguara from the Tocantins. In 1654-55, a thousand Tupinambá came to Belém from the upper Tocantins (LEITE, 1943, p. 337; HEMMING, 1987, p. 325), and in 1658, around a thousand Poquiguara were brought down to Belém, as well as many Karajá from the Araguaia. In 1659, 1200 Tupinambá and 800 Poquiguara entered the missions, along with 240 Inheguara slaves (LEITE, 1943, p. 337, n. 1, 339). Yet more Poquiguara and Tupinambá arrived in 1668 (LEITE, 1943, p. 340).

Most of the first wave of missionized peoples are identified as Tupinambá or are classified as '*índios da língua geral*', e.g., Poquiguara, Naimiguara (LEITE, 1943, p. 341, n. 1). However, some were not, e.g., Karajá (Macro-Jê), Nheengaíba (unidentified, perhaps

Arawak)[10]. Others are simply unidentifiable. However, the non-TG segments of the colonial population were still quite small, and although the missionized non-OT TG speakers would have contributed more variants to the Tupinambá koine, their inputs are unlikely to have altered the overall structural template, for 18th c. written LG had diverged very little from the 1596 and 1621 models (MONSERRAT, 2003, RODRIGUES; CABRAL, 2011), and 19th- and 20th-century NHG retains a markedly TG typology (see below).

Phase three: the demographic and linguistic watershed (1660-1700)

A few mission villages had been established close to Belém before 1653. However, from that year onwards, under Antônio Vieira, many more were built around Belém and São Luís. Large indigenous populations close to European settlements unleashed the same catastrophe that had engulfed the missions in the State of Brazil in the 1550s and 1560s (HEMMING, 1995, p. 343). São Luís had already seen smallpox outbreaks in 1621 and 1644, but now waves of epidemics ravaged the native population. The "pestilential catarrh" of 1660 (1995, p. 344) was followed by smallpox and measles in 1661-63 (1995, p. 345). Epidemics raged throughout the final third of the 17th century, with further devastating plagues in 1758, 1762, 1763-72, and 1776 (HEMMING, 1995, p. 662).

The colonial response to the demographic collapse was to "bring down" (*descer*) more indigenous groups from the hinterland. Unlike the founding populations, these groups were usually *Tapuias,* i.e., non-OT-speakers. They were "trained" or "instructed" (*praticados*) in basic notions of Catholicism and learned vernacular LG by unstructured immersion. Some perhaps acquired some of the codified variety, i.e., LB, by rote catechesis. They were not converted using their native languages, as was idealised, because the tiny number of missionaries was insufficient to learn all the highly diverse languages they encountered in the time they had available to replace the villages' populations.

Many authors, e.g., Lee (2005), Freire (2011), and Nobre (2016), have cited Baena's estimate of 184,000 indigenous people brought down '*só de alguns sertões dos rios Tocantins, Amazonas e Negro'* between 1687 and 1690 (BAENA,1831, p. 247). In 1719, Manoel de Seixas claimed that, over the previous century, the expeditions to collect jungle products each year also brought on average around a thousand indigenous slaves, i.e., 100,000 in total. However, like Vieira's boasts of 200,000 souls in the mission villages in the mid-1600s (HEMMING, 1995, p. 506), it is very likely that these figures are highly

---

10 The fact that these peoples of Marajó and Cabo do Norte were called *Nheengaíba* in LG or *linguas travadas* in Portuguese is precisely an indication that they did not speak a TG language. Vieira calls their languages "diferentes e dificultosas" (1660, p. 6 *apud* ROSA, 1992, p. 89).

exaggerated, as is João Daniel's estimate of 3 million people having been "brought down" from the Rio Negro (HEMMING, 1995, p. 459) between 1650 and 1750.

Our suspicions regarding the colony's demography come firstly from the fact that Baena's estimate is problematic, since the Rio Negro was not open to the *tropas de resgate* until the destruction of the Manao in 1730, who had acted as middlemen in slaving transactions until that time.

Secondly, Vieira himself recorded only some 3000 missionized Indians and 1800 slaves from eight expeditions between 1655 and 1670 (HEMMING, 1987, p. 333-334). A thirteen-fold growth in descents between 1670 and 1687 seems very unlikely, given that, just thirty years later, Pará's missions allegedly held 54,264 people (RAIOL, 1900, p. 132, *apud* FREIRE, 2011, p. 68; see also HEMMING, 1995, p. 421). These figures would require a population crash of 65-75% between 1690 and 1720, which seems excessive even given horrific indigenous mortality in the epidemics. Hemming estimates the enslaved population of Pará at 10,000 in 1700 (2004, p. 243), curiously undermining his own proposal of 100-200,000 slaves taken between 1620 and 1720, as this total would require the slavers to have bought or captured at least 10,000 individuals in every decade across the century analysed. As we have shown, Vieira recorded that fewer than 2000 indigenous slaves were captured in the first five years of the mission expansion. Before the 1650s, there were only a few hundred Luso-Brazilians in Maranhão. Even if every single freeman owned ten slaves, which is unlikely, the enslaved population would not have been more than a few thousand.

On the other hand, if one understands these figures to include the "free" victims of descents *and* the "ransomed", i.e., slaves, the totals seem more plausible. Hemming's 10,000 slaves in 1700, Raiol's 20,000 *mamelucos* and slaves in 1720, plus 54,216 mission villagers, give approximately 85,000 people in total for the first half of the 1700s, which is close to Hemming's lower estimate and roughly that of the contemporary evaluation by Seixas. One may, of course, still envisage the mortality-rate as running at perhaps some 15-25% of the total, which is a very considerable proportion of the population.

The mission-village ecology is not the same as the Caribbean plantations

The cycles of epidemics and repopulation from 1660 into the second half of the 18th century created ecologies that would in theory favour linguistic divergence as ever more non-TG speakers contributed more diverse features to the feature pool that LG speakers and learners drew on. Non-native LG speakers, themselves not long exposed to the language, would have provided the models for more recent arrivals. The destruction

of the Manao people (1728-30) is significant, for it opened the Rio Negro to unfettered "ransoming", resulting in tens of thousands of *Tapuias* being enslaved over the following half-century.

Such rapid replacement of the adult population should weaken communal norms and lead to violent linguistic restructuring as *bozal* (non-locally born) incomers' speech habits overwhelmed the *ladinos'* (fluent non-native) and natives' norms, generating basilectalization (MUFWENE, 2003, p. 31, p. 49-50, p. 209, n. 6, p. 213, n. 14). Nobre (2011a, 2012b) makes this claim for the later 17th- and especially 18th-century missions, modifying his position in 2016 to claim LG's restructuring made it a mesolectal creole.

Nevertheless, in Mufwene's model, ecology is everything. In the first half of the 18th century, there were perhaps around three to six indigenous villagers and slaves for every *mameluco,* i.e., 60,000:10-20,000 (FREIRE, 2011; HEMMING, 1995). Such a demographic imbalance between native and non-native speakers would certainly have favoured evolution away from LB and also from the founders' Tupinambá koine. Nevertheless, the disproportion is noticeably much lower than most Caribbean plantation colonies, e.g., 10:1 in Jamaica in the mid-1800s (MUFWENE, 2003, p. 39) and 17:1 in Suriname in 1700 (2003, p. 44, p. 45-50). The closest fit numerically and chronologically with a known case of creolisation is perhaps Barbados, where an initial 30-year European majority fell to three Africans per European (2003, p. 39-40). Similarly, in *Maranhão e Pará*, the Tupinambá majority lasted some fifty years (1616-1660/70).

Additionally, the traditional creolisation model whereby a pidgin develops amongst non-native speakers and is then learned natively by children, as borrowed by Nobre (2016) from Lucchesi and Baxter (2009), cannot be applied effectively to the LG in the Portuguese Amazon until the final stages of our chronology. The immediately preceding paragraph shows that population expansion in *Maranhão e Pará* was driven by slaving "ransoms" and missionary "descents", not by vegetative growth. The historical data, therefore, do not favour the existence of a large enough body of children in the missions to produce traditional-style creolisation. This is what Mufwene has proposed for the Atlantic creoles (2003, p. 50-51). In a plantation economy, birth rates are usually low amongst the enslaved, and interactions between adults are the principal drivers of language change. Such a model is also far more appropriate for the formation of LG in Amazonia after the foundation period and the collapse of the original mission populations. However, in *Maranhão e Pará*, there were also significant differences regarding restrictions on adults' access to native and fluent non-native models.

Both traditional creolistics and Mufwene stress that Caribbean plantation slaves often modelled their language use on non-native speakers. However, the Amazonian

colonial economic system depended on the circulation of the missionized *Tapuia*. The obligatory labour-service sent mission-dwelling males between the ages of thirteen and fifty for at least six months every year – and frequently much more – to work outside the missions alongside and under other LG speakers. The numerical dominance of such L2 LG speakers would have favoured particularly flexible, open-ended norms and substrate effects, but given their regular, extended immersions, mission-dwellers, especially men, had good access to native LG-speakers outside the mission villages.

Thus, aside from people ignorant of LG, there would have been individuals and communities, e.g., recent arrivals and groups, both indigenous and non-indigenous, with only sporadic contact with colonial society, whose performance in LG might indeed be considered "pidginized", i.e., strongly divergent L2+, with much heavier substrate influences[11]. However, such groups or individuals would constitute subgroups within the broader LG-using community. Continued, and especially increased, contact would have favoured ever greater structural approximation to the wider community norm. At the other extreme, native speakers of whatever passed as contemporary LG were always present in colonial society and were easily accessible to most non-native speakers, especially men, as we have shown. These men would have been the transmitters of LG from the wider society into the mission villages. Between these poles would have existed an immense range of non-native competencies reflecting the huge number of internal and external ecologies.

This conjunction of factors favours the idea of a process resembling a more gradual 'late basilectalization' of LG in the Amazon (MUFWENE, 2003, p. 10), rather than the traditional view of a rapidly formed pidgin basilect that creolises and then is gradually 'decreolised' through greater access to local koines and standard varieties. Indeed, if one accepts that LG was a creole, it could not have "decreolised" in this way in the Amazon, because by 1757 there were virtually no speakers of anything even approximating to the original OT koine or the LB norm, as João Daniel observed. Thus, the expulsion of the Jesuits in 1759 did not stop the restructuring processes and was largely irrelevant to the

---

11 A pidgin, in a Mufwenean model, is, in our view, a communal L2 norm of a language that only approximates native speakers' performance, despite serving as a satisfactory vehicle for communication, and that exhibits restructuring marked by innovations, substrates, and hybrid forms. So-called "expansion" occurs as the communicative fields and functions broaden, naturally favouring the creation of a larger expressive repertoire. In functional terms, a nativized, so-called "Expanded Pidgin", such as Tok Pisin, is scarcely distinguishable from a creole. Indeed, we believe that natively spoken expanded pidgins may in fact be indistinguishable from creoles in practical terms, save perhaps in relation to the continued autonomous existence of the pidgin's contributing languages.
We emphasise, however, that there is no *necessary* causal connexion between pidgins and creoles. A pidgin may arise and never undergo creolisation and a creole can arise from considerable restructuring of a communal norm without the prior existence of a generalised pidginized variety.

evolution of LG in the Amazon, since the missionaries' "true" Tupinambá-LG was not actually spoken natively by anyone and, in any case, was potentially barely comprehensible to the contemporary vernacular LG speakers.

The Pombaline reforms (1757-1796) did not abolish the mission villages but renamed and reclassified them as secular *vilas* ("towns"). The former priestly overseers were replaced with lay *directores*, who continued to ruthlessly exploit the indigenous population in much the same way as before (HEMMING, 1995, p. 499). The main effect of the reforms was to foster the emergence of the *Tapuia* as a distinct social class: the majority of the population of the Province of Grão-Pará, which was separated from Maranhão in 1772: nominally free subjects of the Portuguese Crown, nominally Catholic, LG-speaking and often multilingual, detribalised and partially deracinated, who continued to expand the LG-speaking zone and restructure the language into the 19th century.

Phase four: Expansion, contraction, and structural realignment towards Portuguese (1800-present).

This period is characterised by an initial moment of further growth of the LG-speaking community into the beginning of the 19th century, followed by a rapid loss of speakers to vernacular Brazilian Portuguese (VBP). In Maranhão, very significant quantities of African slaves were shipped under Pombal's economic reforms in the second half of the 18th century. Therefore, Portuguese almost certainly came to dominate earlier. In the Amazon proper, the Portuguese language was substituting NHG first on the lower reaches of the Amazon in the Province of Grão-Pará over the first half of the 19th century (FREIRE, 2010, p. 191-196), especially after the bloody repression of the Cabano rebellion (1839-45) in which some 40,000 people (around a third of the province's population) died. Subsequently, in the *província de Amazonas*, which emerged from Grão-Pará's subordinate captaincy of São José do Rio Negro in 1850, the language shift to VBP rapidly gathered momentum over the second half of the 19th century (FREIRE, 2010, p. 196-202, p. 203-207).

For reasons of space, we cannot deal with structural issues arising in this final evolutionary phase in detail here. However, there is good evidence that the NHG object pronoun system was still much closer to 18th-century LG at the beginning of the 19th century.

The archaic language of the *cantigas* in Rodrigues' *Poranduba Amazonense* (1890) (RODRIGUES, J., 2014) and the grammar section of Hartt's *Notes on the Lingoa Geral* (HARTT, 1875, 1938) record person hierarchy with stative indices for first- and second-person singular objects when the subject is third person. This is typical of TG, including OT. Therefore, the postverbal, free object pronouns that are identical with subject pronouns

and constitute the majority of late 19th century examples, and which are now universal in modern Nheengatu, almost certainly arose from alignment with VBP over the course of the 19th century (RODRIGUES, J., 2017, p. 497-578; HARTT, 1938, p. 314).

On the Rio Negro, language shift into NHG occurred simultaneously with the spread of VBP, which intensified in the upper Rio Negro over the second half of the 20th century. The watershed for NHG, when it lost its position as the interethnic lingua franca to VBP, seems to be marked by the knowledge of NHG as a second language amongst the Dâw people (Naduhup family). Dâw born before 1970 learned NHG – heavily influenced by their first language – through contact with the Baré, the largest native NHG-speaking ethnic group on the Rio Negro, who had mostly shifted from their ancestral Arawak language to NHG over the first half of the 20th century. The Dâw worked for the Baré and for other Neo-Brazilians as labourers, hunters, and fishermen, or were sent into the forest to harvest piassaba fibre and lianas, usually in conditions analogous to slavery. Dâw born after 1970 are bilingual in Dâw and Portuguese but know no NHG (FINBOW, 2020).

Conclusions

Basing our account on extensive contemporary usage, unlike previous investigations, which have tended to discuss LG on a more speculative, theoretical plane, we show that the employment of the term *Língua Geral* means that the language must be understood as a lingua franca composed of a very loose collection of broadly mutually intelligible TG varieties. We also speculate on how the concept may have emerged in Portuguese America.

Our treatment reconciles the issue of a possible LG in the south of Bahia (NOBRE, 2011a, 2011b, 2016; RODRIGUES, A., 1996; see also DIETRICH, 2014), in favour of Nobre. Although concrete evidence about the nature of the language spoken in the south of Bahia is slight, the racial and/or social composition of speakers emphasized by Rodrigues to preclude the existence of a LG between Rio de Janeiro and Ceará (1996, p. 11-12) is irrelevant to the contemporary LG definition, as we have shown. If the people spoke something that descended directly from earlier local OT or from their having acquired it during the evangelization period, such a variety would have qualified as LG under contemporary usage of the term.

Second, we analyse two recent theoretical models for the emergence of LGA as a creole and show how the historical data regarding the colony's ecology is also not at all conducive to Vieira, Zanoli, and Módolo (2019)'s proposed model of Tupinambá pidgin > LB > creole LG. Moreover, João Daniel's "corrupt" LG of the mid-18th century is not so

much a wholly new language, i.e., an emergent creole, as Vieira, Zanoli, and Módolo (2019) proposes, but rather a restructured variety of an existing, recognised language, i.e., LG, that is spoken by everyone.

Interestingly, Daniel's testimony also points to a diglossic LG, for he shows that the *Tapuia* he blames for debasing the "true" LG of the Tupinambá spoke the contemporary communal norm and that the structural divergence he observed was with respect to the centuries-old written norm and not between the *Tapuia* and the non-*Tapuia*.

Nobre's use of Lucchesi and Baxter (2009)'s "irregular transmission" model is also not borne out by the historical demography of the successive administrative units of *Maranhão* (1621-1652), *Maranhão e Pará* (1654-1737), and *Grão-Pará e Maranhão* (1751-1772).

The structural analyses in Nobre (2019), which seek to prove NHG's creole typology, are shown to be flawed for despite LGA having lost many LB structures and modified or innovated others, as well as moving closer to VBP, nevertheless, much has been retained. In fact, unlike many Atlantic creoles, which can be hard to recognise as Romance or Germanic at first glance, modern NHG's typology remains very distinctively TG. Aside from some regional loanwords, e.g., *darabi* "plate", *dabukuri* "festival", *kariamã* "initiation ritual", *kuradá* "soft manioc bread", and the Arawak phonological substrate of Rio Negro Nheengatu (MOORE, 2014, CRUZ, 2011, MOORE; FAGUNDES; PIRES, 1994, BORGES, 1991; TAYLOR, 1985), it is as yet unclear which grammatical elements in modern Nheengatu originated in indigenous languages besides the Tupinambá koine, although some undoubtedly did or were favoured by structural congruence with local substrates.

Finally, we apply Mufwene's vision of creole formation to the historical and structural data and find it particularly well-suited to account for the observable phenomena.

We therefore define LGA/NHG as an adult-restructured, L2+-filtered variety employed as an autonomous vernacular norm in a community (MUFWENE, 2003, p. 106; see also CHAUDENSON, 1992). Its consolidation entailed the loss of many other communal languages but did not always involve communal language shift to the restructured variety because some communities were LG-native speakers from the outset and remained so, merely modifying their native norm. Thus, as Mufwenean creole, LG arose without a prior generalised pidgin phase, although some pidginised varieties certainly would have existed in some communities, before increased exposure to the majority norm resulted in ever greater approximations to native competence, often culminating in the community shifting language into LG. In such subgroups, one might propose a transition from pidgin to creole, however, the model for the creolised variety was supplied from outside, and children and young people acquired it through greater exposure rather than creating their own through large-scale innovations.

Moreover, the traditional creolisation model cannot be applied to those subgroups whose ancestors spoke the major contributing variety natively. In the latter case, their norm diverged from the ancestral variety, converging towards the L2+-speakers' norm in the guise of a restructured variety of Amazonian Tupinambá OT. Following Mufwene's observation that "[...] it is those who speak a language on a regular basis – and in a manner they consider normal to themselves – who develop the norms for their communities" (2003, p. 106; see also MUFWENE, 1997), we therefore feel that LGA is better accounted for as the result of an accommodation of norms rather than solely a process of children acquiring and expanding a pidgin.

We envisage native LG speakers adopting into their expressive repertoires many structures from the feature pool that had originated in the speech of non-native speakers. The latter also selected, with minimal modification, many elements that the former contributed to the feature pool. In this way, the overall community norm would diverge from the original vernacular Amazonian Tupinambá varieties, whose speakers founded the initial LG template but were decimated by epidemic disease, removing them as an effective brake on structural evolution. The communal norm would also diverge from the fossilised LB norm, always encountered only fleetingly.

The sociocultural and linguistic diversity and individual high geographical mobility in the Amazonian colonial system would favour flexibility and hybridism, rather than conformity to a single, conservative variety. Contrasting more innovative modern NHG with more conservative modern Paraguayan Guarani is instructive in this regard, despite Rodrigues (1996) classing "*Criollo* Guarani" as another TG *LG* which, as a colonial and post-colonial lingua franca, it undoubtedly is, although shaped by a very different ecology. The rapidly increasing presence of VBP as a prestige variety over the 19th century has redirected evolutionary trends towards ever greater structural congruence, as NHG object-marking strategies show.

Viewed as a Mufwenean creole, the Amazonian *Língua Geral* can be seen more clearly for what it is: a communal norm that has undergone significant restructuring in at least two distinct phases. Firstly, under poorly documented, largely non-native adult-to-adult acquisition strategies in the mission villages and the colonists' farms and households, and, subsequently, under the influence of vernacular Brazilian Portuguese as language shift from NHG occurred in the urban centres of the Brazilian Amazon. At all times, however, LG was the native language of a considerable number of people, and, thus, to see it as a "new" language arising from the forced integration of outsiders is to overlook the considerable structural continuities with preceding phases that have been and continue to be evident at all times.

References

ABBÉVILLE, C. d'. **História da missão dos padres capuchinos na ilha do Maranhão e suas circumvizinhanças**. [S. Luís do] Maranhão: Typographia do Frias, rua da Palma, n. 6, traduzido por Cézar Augusto Marques, 1874. Brasília: Biblioteca do Senado Federal, 1974.

ABOH, E. O. **The Emergence of Hybrid Grammars:** Language Contact and Change. Cambridge: Cambridge University Press, 2015.

ABOH, E. O.; DEGRAFF, M. A Null Theory of Creole Formation Based on Universal Grammar. *In*: ROBERTS, I. (ed.). **The Oxford Handbook of Universal Grammar**. Oxford: OUP, 2017. p. 401-458.

ALFARO LAGORIO, M. A. C. Usos e funções das línguas na área andina: o Terceiro Concílio Limense (1582-1583). *In*: ALFARO LAGORIO, C.; ROSA, M. C.; FREIRE, J. R. Bessa (org.). **Políticas de línguas no Novo Mundo**. Rio de Janeiro: EdUERJ, 2012. p. 41-54.

ALFARO LAGORIO, M. A. C. Elementos da política lingüística colonial hispânica: o Terceiro Concílio Limense. *In:* FREIRE, J. R. B.; ROSA, M. C. (org.). **Línguas gerais. Política lingüística na América do Sul no período colonial.** Rio de Janeiro: EdUERJ, 2003. p. 43-56.

ALTMAN, C. As línguas gerais sul-americanas e a empresa missionária: linguagem e representação nos séculos XVI e XVII. *In:* FREIRE, J. R. B.; ROSA, M. C. (org.). **Línguas gerais. Política lingüística na América do Sul no período colonial.** Rio de Janeiro: EdUERJ, 2003. p. 57-84.

ANCHIETA, J. de. **Arte de Gramática da língoa mais usada na costa do Brasil (1595).** 5th ed. facsimile, Salvador, BA: EdUFBA, 2014.

ARAÚJO, A. de. **Catecismo brasílico da doutrina crista**. Edição fac-similar da 2ª ed. de 1686, corrigida por Bartolomeu de Leão. Edited by Júlio Platzmann, Leipzig: B. G. Teubner, 1898.

AVILA, M. Twardowsky. **Proposta de dicionário nheengatu-português**. 2021. Tese (Doutorado em Estudos da Tradução) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2021.

BAENA, A. L. M. **Representação ao Conselho Geral da província do Pará sobre a especial necessidade de um novo regulamento promotor da civilização dos índios da mesma província.** [1831]. Belém: ABAPP, v. 2, p. 241-92, 1901.

BAKKER, P. Creoles and Typology. Problems of sampling and definition. *In:* MANFREDI, S.; TOSCO, M. (ed.). **Arabic-Based Pidgins and Creoles. Special issue of *Journal of Pidgin and Creole Languages* 29:2.** Amsterdam: John Benjamins, 2014. p. 437-455. DOI: http://dx.doi.org/10.1075/jpcl.29.2.

BAKKER, P.; DAVAL-MARKUSSEN, A.; PARKVALL, M.; PLAG, I. Creoles are typologically distinct from non-creoles. **Journal of Pidgin and Creole Linguistics**, v. 26, n. 1, p. 5-42, 2011. DOI: http://dx.doi.org/10.1075/jpcl.26.1.02bak.

BARROS, M. C. Drummond de Menezes. Notas sobre a política jesuítica da língua geral na Amazônia (séculos XVII-XVIII). *In*: FREIRE, J. R. B., ROSA, M. C. (org.). **Línguas gerais:** política linguística e catequese na América do Sul no período colonial. Rio de Janeiro: EdUERJ, 2003. p. 85-112.

BARROS, M. C. 'O uso do tupi na Capitania de São Paulo no século XVII. Indícios na vida de um jesuíta "língua"'. *In*: NOLL, V.; DIETRICH, W. (org.). **O português e o tupi no Brasil**. São Paulo: Contexto, 2010. p. 141-153.

BAXTER, A.; LUCCHESI, D. A transmissão linguística irregular. *In:* LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (org.). **O Português Afro-Brasileiro**. Salvador: EdUFBA, 2009. p. 101-124.

BETTENDORF, J. Ph. **Compendio da Doutrina Christãa na Lingua Portugueza, e Brasilica.** Lisboa: Offic. de Simão Thaddeo Ferreira, 1687. Reimpresso por Fr. José Mariano da Conceição Vellozo, 1800.

BORGES, L. C. **A lingua geral amazônica:** aspectos de uma fonêmica. 1991. Dissertação (Mestrado) - Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1991.

CABRAL, A. S. A. C.; OLIVEIRA, G. B. V. de. Sobre possíveis influências bilateriais entre a Língua Geral Amazônica falada no Baixo Amazonas e as línguas Tupí-Guaraní do subramo VIII, com foco especial na língua Zo'é. **Anais do SILEL.** Uberlândia: EdUFU, v. 3.1, p. 1-13, 2013.

CÂMARA JR., J. M. **Introdução às Línguas Indígenas**. 3. ed. rev. Rio de Janeiro: Acadêmica, 1977.

CARDOSO, A. A. Isidio. O mundo de Manoel Urbano da Encarnação: indígenas, regatões, migrantes e fugitivos no avanço rumo ao oeste amazônico no século XIX. Paper presented to the **XXVII Simpósio Internacional de História,** 2013.

CHAUDENSON, R. **Les créoles français**. Paris: Ferdinand Nathan, 1979.

CHAUDENSON, R. **Créoles et ensaignement du français**. Paris: L'Harmattan, 1989.

CHAUDENSON, R. **Des îles, des hommes, des langues:** essais sur la créolisation linguistique et culturelle. Paris: L'Harmattan, 1992.

COSTA, F. A. Elementos necessários para aprender o Nheengatú. *In*: COSTA, F. A. **Carta pastoral de D. Frederico Costa, Bispo do Amazonas a seus amados diocesanos**. Ceará: Typ. Minerva, 1909. p. 157-248.

CRUZ, A. The rise of number agreement in Nheengatu. **Boletim do Museu Paraense Emílio Goeldi. Ciências humanas**, v. 10.2, May-Aug. 2015.

CRUZ, A. **Fonologia e gramática do Nheengatu. A língua geral falada pelos povos Baré, Warekena e Baniwa**. Utrecht: LOT, 2011.

DANIEL, J. **Tesouro descoberto no Rio Amazonas**. 2 vols. Rio de Janeiro. Biblioteca Nacional, 1976.

DIETRICH, W. O conceito de "Língua Geral" à luz dos dicionários de língua geral existentes. **D.E.L.T.A.,** 30 especial, p. 591-622, 2014.

DIETRICH, W. O tronco tupi e as suas famílias de línguas. Classificação e esboço tipológico. *In*: NOLL, V.; DIETRICH. W. (org.). **O português e o tupi no Brasil**. São Paulo: Editora Contexto, 2010. p. 9-26.

EDELWEISS, F. G. **Estudos Tupis e Tupi-Guaranis:** confrontos e revisões. Rio de Janeiro: Livraria Brasiliana Editora, 1969.

ESTIGARRIBIA, B. **A Grammar of Paraguayan Guarani**. London: UCL Press, 2020.

ÉVREUX/OBERMEIER. **Voyage au Nord du Brésil (1615), edition critique de texte complet établi por Franz Obermeier.** Version en forme de livre 2012. Version électronique, 2014.

ÉVREUX/DENIS. **Voyage au Nord du Brésil fait pendant les années 1613 et 1614 par le père Yves d'Évreux. Publié d'après l'exemplaire unique conserve à la Bibliothèque impériale à Paris avec une introduction et des notes par M. Ferdinand Denis.** Leipzig/Paris: Librairie A. Franck, 1864.

FERNANDES, F. **A organização social dos tupinambá**. São Paulo: Instituto Progresso Editorial, 1948.

FIGUEIRA, L. **Arte da Lingua Brasilica**. Lisboa: Manoel da Silva, 1621.

FIGUEIRA, L. **Arte de grammatica da lingua brasilica**. Edição facsimilar, por Júlio Platzmann, da 2. ed. de 1687. Lisboa: Miguel Deslandes. Leipzig: B. G. Teubner, 1880.

FINBOW, T. D. Nheengatu Dâw. **Cadernos de Linguística**, v. 1, 2020, p. 1-21.

FREIRE, J. R. B. **Rio Babel. A história das línguas na Amazônia**. 2. ed. Rio de Janeiro: Atlântica, 2011.

FREIRE, J. R. B.; ROSA, M. C. **Línguas Gerais. Política linguística e catequese na América do Sul no período colonial.** Rio de Janeiro: EdUERJ, 2003.

GALÚCIO, V.; MEIRA, S.; BIRCHALL, J.; MOORE, D.; DRUDE, S.; GABAS JR, N.; STORTO, L.; PICANÇO, G.; REIS RODRIGUES, C. Genealogical relations and lexical distances within the Tupian linguistic family. **Boletim do Museu Goeldi Ciências Humanas**, Belém, v. 10, n. 2, p. 229-274, maio-ago. 2015.

GOMES, M. Payno. Brazilian Portuguese and Nheengatu in São Gabriel da Cachoeira (AM): some effects of language contact. **Cadernos de Linguística,** v. 2, n. 4, p. e.480, 30 set. 2021.

GOMES, M. Carta que o Padre Superior Manoel Gomes escreveu ao Padre Provincial do Brasil. **Anais da Biblioteca Nacional do Rio de Janeiro**, v. 26, 1904.

HARTT, C. F. Notes on the Lingoa Geral or Modern Tupi of the Amazonas. **Transactions of the American Philological Association**, v. 3, p. 58-76, 1875.

HARTT, C. F. Notas sobre a lingua geral ou tupi moderno do Amazonas. **Anais da Biblioteca Nacional do Rio de Janeiro**, Rio de Janeiro: M. E. S. Serviço Gráfico, v. 51, p. 305-390, 1938.

HEMMING, J. **Amazon Frontier. The Defeat of the Brazilian Indians**. 2nd ed. London: Papermac, 1995.

HEMMING, J. **Red Gold. The Conquest of the Brazilian Indians**. 2nd ed. London: Papermac, 1987.

JENSEN, C. Tupi-Guarani. *In:* DIXON, R. M. A.; AIKHENVALD, A. (ed.). **The Amazonian Languages.** Cambridge: Cambridge University Press, 1999. p. 125-164.

KELLER, R. **On Language Change:** the Invisible Hand in Language. London: Routledge, 1994.

LASS, R. **Historical Linguistics and Language Change.** Cambridge: Cambridge University Press, 1997.

LEE, M. K. **Conversing in Colony:** The Brasílica and the Vulgar in Portuguese America 1500-1759. Ph.D. Dissertation, John Hopkins University, 2005.

LEITE, S. **História da Companhia de Jesus no Brasil**. v. 4. Rio de Janeiro: Instituto Nacional do Livro, 1943.

MAGALHÃES, J. V. C. de **O selvagem**. Rio de Janeiro: Typ. da Reforma, 1876.

MAGALHÃES, M. **Sobre a morfologia e a sintaxe da língua Guajá**. 2007. Tese (Doutorado em Linguística) – Instituto de Letras, Universidade de Brasília, Brasília, 2007.

MCWHORTER, J. **The Creole Debate.** Cambridge: Cambridge University Press, 2020.

MELLO, A. A. S.; KNEIP, A. Algumas evidências linguísticas (e algumas arqueológicas) que apontam para a origem dos povos tupi-guarani no leste amazônico. **Literatura e linguística** v. 36, p. 299-312, 2018.

MÉTRAUX, A. **A religião dos tupinambás.** São Paulo: Companhia Editora Nacional; Editora da Universidade de São Paulo, 1979.

MICHAEL, L.; CHOUSOU-POLYDOURI, N.; BARTOLOMEI, K.; DONNELLY, E.; WOUTERS, V.; MEIRA, S.; O'HAGAN, Z. A Baysian phylogenetic classification of Tupi-Guarani. **LIAMES**, Campinas, v. 15.2, p. 193-221, jul./dez. 2015.

MONSERRAT, R. O Tupi do século XVIII (tupi médio). *In:* FREIRE, J. R.; BESSA, R. M. C. (org.). **Línguas Gerais**, p. 185-194, 2003.

MONSERRAT, R.; BARROS, C. de, MOTA, J. Comparação entre dois diálogos de doutrina jesuíticos tupi: João Filipe Bettendorf (1687) e José Vidigal (1740). I: **Anais da XIII Jornadas Internacionais Missões Jesuíticas, Dourados. XIII Missões jesuíticas**, 2010.

MOORE, D. Historical Development of Nheengatu. *In:* MUFWENE, S. S. (ed.). **Iberian Imperialism and Language Evolution in Latin America.** Chicago: University of Chicago Press, 2014. p. 108-142.

MOORE, D.; FAGUNDES, S.; PIRES, N. Nheengatu (língua geral amazônica), its history and the effects of language contact. *In*: LANGDON, M. (ed.). **Survey Reports. Survey of California and Other Indian Languages. Proceedings of the Meeting of the Society for Indigenous Languages of the Americas**. Columbus, Ohio: Ohio State University Press, 1994. p. 93-118.

MUFWENE, S. S. **The Ecology of Language Evolution**. Cambridge: C.U.P., 2003.

MUFWENE, S. S. **Language Evolution. Contact, Competition and Change**. London & New York: Continuum, 2008.

MUFWENE, S. S. Native speaker, proficient speaker, and the norm. *In:* SINGH, R. (org.). **Native Speaker:** Multilingual Perspectives. New Delhi: Sage, 1997. p. 11-23.

NAVARRO, E. de Almeida. **Dicionário de tupi antigo. A língua indígena clássica do Brasil**. São Paulo: Editora Global, 2013.

NAVARRO, E. **Método novo de tupi antigo**. 3. ed. São Paulo: Editora Global, 2008.

NAVARRO, E. (org.). **Anchieta. Poemas. Lírica portuguesa e tupi**. 2. ed. São Paulo: Martins Fontes, 2004.

NOBRE, W. Línguas gerais na história social-linguística do Brasil. **PAPIA**, São Paulo, v. 26, n. 1, p. 7-52, 2016.

NOBRE, W. **História linguística do Sul da Bahia (1534-1940)**. 2015. Tese (Doutorado em Letras) - Universidade Federal da Bahia, 2015.

NOBRE, W. História linguística do Sul da Bahia: levantando hipóteses e iluminando caminhos. **Entrepalavras**, v.1, n. 2, p. 270-292, 2012a.

NOBRE, W. A língua geral da Amazônia como um sistema historicamente novo: jesuítas e tapuias na origem do contexto com interrupção de transmissão linguística entre gerações. *In*: MATTOS E SILVA, R. V.; OLIVEIRA, K.; AMARANTE, J. (org.). **Várias navegações:** português arcaico, português brasileiro, cultura escrita no Brasil, outros estudos. Em homenagem a Therezinha Barreto. Salvador: EdUFBA, 2012b. p. 479-513.

NOBRE, W. **Introdução à história das línguas gerais no Brasil:** processos distintos de formação no período colonial. 2011. Dissertação (Mestrado em Letras) - Universidade Federal da Bahia, 2011a.

NOBRE, W. Língua geral na Bahia: comarcas de Ilhéus e Porto Seguro. *In:* CARVALHO, C. dos Santos; ROCHA, F. Aninger de Barros; PARCERO, L. M. de Jesus (org.). **Discurso e cultura:** diálogos interdisciplinares. Salvador: EdUNEB, 2011b. p. 99-109.

PRAÇA, V. N. **Morfossintaxe da língua Tapirapé**. 2007. Dissertação (Mestrado em Linguística) – Instituto de Letras, Universidade de Brasília, Brasília, 2007.

PÉREZ PUENTE, L. La creación de las cátedras públicas de lenguas indígenas y la secularización parroquial. **Estudios de historia novohispana**, v. 41, p. 45-78, jul./dez. 2009.

OLIVEIRA, M. S. D. de; ZANOLI, M. de L.; MODOLO, M. O conceito de "Língua Geral do Brasil" revisitado à luz da linguística de contato. **Journal of Ibero-Romance Creoles**, v. 9.1, p. 306-333, 2019.

RAMOS PÉREZ, D. La crisis Indiana y la junta magna de 1568. **Jahrbuch für Geschichte Lateinamerikas/Anuario de historia de América Latina**, n. 23, p. 1-61, 1986.

RODRIGUES, A. Dall'Igna. Tupi, tupinambá, línguas gerais e português do Brasil. *In*: NOLL, V.; DIETRICH, W. (org.). **O português e o tupi no Brasil.** São Paulo: Contexto, 2010. p. 27-47.

RODRIGUES, A. Some cases of regrammaticalization in Tupí-Guaraní languages**. Revista Brasileira de Linguística Antropológica**, v. 2.2, p. 65-74, dez. 2010.

RODRIGUES, A. As línguas gerais sul-americanas. **Papia:** Revista Brasileira de Estudos Crioulos e Similares, v. 4, p. 6-18, 1996.

RODRIGUES, A. Relações internas na família tupi-guarani. **Revista de Antropologia**, v. 27/28, p. 33-53, 1984/85.

RODRIGUES, A. Classification of Tupi-Guarani. **International Journal of American Linguistics,** v. 24, p. 231-234, 1958.

RODRIGUES, A. D. I.; CABRAL, A. S. A. A contribution to the linguistic history of the Língua Geral Amazônica. **Alfa**, v. 55.2, p. 614-639, 2011a.

RODRIGUES, A. A contribution to the linguistic history of the *língua geral amazônica*. **Alfa**, v. 55.2, p. 613-639, 2011b.

RODRIGUES, J. B. **Poranduba amazonense, ou kochiyma-uara porandub**. 2. ed. Organizado por T. Telles. Manaus: Editora Valer, 2017.

ROSA, M. C. The 16th and 17th centuries: Tupi or Lingua Geral? The grammars of Anchieta and Figueira. I: **Actes: La "découverte" des langues et des écritures d'Amérique"**, 1990.

ROSA, M. C. Descrições missionárias de língua geral nos séculos XVI-XVII: que descreveram? **Papia**, v. 1, p. 85-98, 1992.

SANTOS, B. M. dos. As cartas do padre Manoel Gomes e suas representações em torno da incipiente missão jesuítica no Maranhão. Paper presented at **XVI Jornadas sobre Alternativas Religiosas en América Latina**. 2011. Accessed at: https://ptdocz.com/doc/424876/as-cartas-do-padre-manuel-gomes-e-suas-representa%C3%A7%C3%B5es. Acesso em: 3 maio 2022.

SARAGOÇA, L. **De Feliz Lusitânia aos Confins da Amazónia 1616-62.** São Paulo: Cosmos, 2000.

SEIXAS, M. J. J. **Vocabulário da língua indígena geral para o uso do Seminário Episcopal do Pará**. Belém do Pará: Typ. De Mattos e Comp., 1853.

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.

SCHMIDT-RIESE, R. Referencialidade ealinhamento em tupinambá: sobre a adequação de imprecisões missionárias. *In*: ALFARO LAGORIO, M. A. C.; ROSA, M. C.; FREIRE, J. R. B. (org.). **Políticas de línguas no Novo Mundo.** Rio de Janeiro: EdUERJ, 2012. p. 149-178.

TAYLOR, G. Apontamentos sobre o Nheengatu falado no rio Negro, Brasil. **Amérindia: revue d'ethnolinguistique amérindienne**, v. 10, p. 5-23, 1985.

WRIGHT, R. **A Sociophilological Study of Late Latin**. Turnhout: Brepols, 2002.

WRIGHT, R. **Early Ibero-romance:** Twenty-One Studies on Language and Texts from the Iberian Peninsula between the Roman Empire and the Thirteenth Century. Newark, DEL: Juan de la Cuesta, 1994.

WRIGHT, R. **Late Latin and Early Romance in Spain and Carolingian France**. Liverpool: F. Cairns, 1982.