

Publication status: This preprint has been published elsewhere.

DOI of the published preprint: <https://doi.org/10.25189/2675-4916.2025.v6.n1.id806>

# A Sociophilological Account of the Formation and Evolution of the term *Língua Geral*, with Emphasis on Amazonia

Thomas Finbow

<https://doi.org/10.1590/SciELOPreprints.11130>

Submitted on: 2025-01-23

Posted on: 2025-01-29 (version 1)

(YYYY-MM-DD)

## **A sociophilological account of the formation and evolution of the term Língua Geral, with emphasis on Amazonia**

**Thomas Finbow**

**Departamento de Linguística, Universidade de São Paulo**

thomas.finbow@usp.br

<https://orcid.org/0000-0003-4803-6603/>

### **ABSTRACT**

This chapter critically reevaluates the role of structural change in the development of the concept of *Língua Geral*, as proposed in Rodrigues (1986,1996, 2010; see also EDELWEISS, 1947, 1969; DIETRICH, 2014). In Rodrigues's model, two languages, *Língua Geral Paulista* and *Língua Geral Amazônica*, emerged from the Tupi and Tupinambá languages spoken by bi- or multilingual 'Mamluks' (Luso-Amerindian mestizos) over the 16<sup>th</sup> and 17<sup>th</sup> centuries.

However, a socio-philological analysis (WRIGHT, 1982, 1991, 1994, 2002) applied to a broad sample of Luso-Brazilian colonial sources, reveals that contemporaries neither attributed a characteristic variety of *Língua Geral* to the mamluks nor classified *Língua Geral* diatopically. Moreover, in 18<sup>th</sup> c. Amazonia, the linguistically non-Tupi-Guarani *Tapuia* peoples, rather than the Mamluks, were explicitly identified as agents of linguistic change (Daniel, 2004 [1757-1776]); however, no new name emerged. Thus, although language contact and shift undeniably occurred and undoubtedly contributed to structural change, Rodrigues' hypothesis regarding the Mamluks' central role in the formation of the *Língua Geral* cannot be sustained.

We therefore develop a revised trajectory for the concept of *Língua Geral* in which we propose that perceived changes in function rather than observed structural divergence were responsible the shift from using '*Língua Brasileira*' in 17<sup>th</sup> c. Jesuit publications to '*Língua Geral*' in the 18<sup>th</sup> c. Our research highlights the need for a re-assessment of the term *Língua Geral*, and, in particular, the anachronistic *Ausbau* of periodizations and varieties on the basis of later perceived structural *Abstand* (KLOSS, 1967, 1976, 1978) that do not correspond to contemporary usage.

**Keywords: Língua Geral, Old Tupi, Nheengatu, Sociophilology, Aryon Rodrigues**

**Palavras-chave: Língua Geral, Tupi Antigo, Nheengatu, Sociofilologia, Aryon Rodrigues**

### **Introduction**

This chapter investigates the development of naming patterns for the so-called *Língua Geral* ('general language', i.e., 'lingua franca'), a Tupi-Guarani language spoken very extensively in Portugal's South American colonies. An Amazonian variety of the *Língua Geral* – nowadays mostly called *Nheengatu* ('good language/speech') – is still spoken natively today

along the Rio Negro in Brazil and Venezuela, and residually in Colombia. Nheengatu is also being revitalised as a heritage language in several Amazonian communities.

We focus on the divisions that linguists and historians have proposed, particularly Rodrigues (1986, 1996), which heavily emphasises ethno-racial miscegenation as the main catalyst for structural linguistic changes that in turn stimulated the invention of new names. We argue that, although much structural diversity existed and was occasionally commented on, such *Abstand* “[structural] distance” (KLOSS, 1978, 1967, 1976, see also GOEBL, 1989; BOSSONG, 2008, p. 25-28) was not the basis for the perceived changes in nomenclature that modern researchers have identified in Old Tupi (henceforth, OT) and in LG, especially the shift in Jesuit publications from using *Língua Brasileira* (‘Brazilian language’, henceforth LB) in the 17<sup>th</sup> century to *Língua Geral* in the 18<sup>th</sup> century (henceforth, LG) (EDELWEISS, 1969; LEE, 2005, 2014; RODRIGUES, 1984/5, 1986, 1996, 2010; DIETRICH, 2010; ARGOLO, 2011a/b, 2012a/b, 2016; VIEIRA; ZANOLI; MÓDOLO, 2019).

Instead, developing Finbow (2022), we sustain that the evidence indicates that (at least) non-indigenous people between the 16<sup>th</sup> and 18<sup>th</sup> centuries saw the language called both ‘the LB’ and ‘LG’ as essentially a single language complex, in spite of extensive diatopic variation. Rather than the conscious perception of structural differences, the name shifted primarily from changes in the language’s function between the 16<sup>th</sup> and 18<sup>th</sup> centuries. Thus, the idea that the terms LB and LG refer to ‘different languages’ or to different diachronic phases of the same language is the result of modern *Ausbau* ‘elaboration’ (KLOSS, 1967, 1976, 1978) and therefore, anachronic.

To analyse the linguistic consciousness of Portuguese America’s colonies regarding LB and LG, we draw on sociophilological frameworks, developed to understand how the conceptual distinction between ‘Latin’ and ‘Romance’ arose in the early Middle Ages.

## **Sociophilology**

Sociophilology is essentially the diachronic investigation of linguistic metalanguage (HERMAN 1996, p. 30, WRIGHT 1996, p. 31-44, 277-287, 2002a, p. vii). Bringing the insights of sociolinguistics into Romance historical linguistics and philology, Wright transformed the understanding of emergence of the Romance languages, convincingly showing how the classic models’ notions of natural, gradual, structural evolution producing ever greater *Abstand* as vernacular Latin diverged from an artificially maintained, hyper-archaic variety cannot explain the emergence of ‘Romance’ as a metalinguistic category distinct from the

traditional denominations ‘Roman’ and ‘Latin’ (WRIGHT, 1982, p. i–ii, 1994: 27-28, 2003, p. 676-677, see also FINBOW, 2011, 2012). Instead, Wright demonstrates the impact of conscious interventions in formal written style in the form of a novel spelling pronunciation and insistence on profoundly archaic lexis and grammar in the 8<sup>th</sup> and 9<sup>th</sup> century Carolingian reforms in the Frankish Empire and the introduction of this novel Carolingian Latin into the Christian Iberian kingdoms over the 12<sup>th</sup> century during the Gregorian reforms (WRIGHT, 1976, 1982, 1994, 1996, 2000a/b, 2002). Thus, the distinction we make between ‘Latin’ and ‘the (Romance) vernaculars’ as distinct languages is a case of *Ausbau* which evolved fairly gradually from the Carolingian period and not consolidated until almost a century after the respective ecclesiastical reforms in each region, as the artificially induced *Abstand* with the regional vernaculars impeded vertical communication, i.e., the act of reading aloud to an audience (BANNIARD, 1992, 2013), which implies the illiterate could understand reading aloud in traditional written styles before the reforms, despite numerous archaisms, but struggled to understand the reformed variety subsequently (WRIGHT, 1994, p. 3, 126-27, 1999, p. 506-07, see also BANNIARD, 1992). Consequently, sermons in the vernacular were permitted. However, as the vernacular no longer possessed any written form, since the traditional spellings now had reformed pronunciations, clerics used the novel, direct grapho-phonemic correspondences of reformed spelling as models for spellings that represented vernacular phonology. The term “Romance” (< *romanice* “Romanly”, i.e., “vernacularly”) referred to this new modality of vernacular writing while “Latin” (*latine* “Latinly” > “formally”, “properly”) became restricted to the non-natively acquired, ecclesiastical norm (WRIGHT, 1982, 1996, 2002; see also MÜLLER, 1963). In time, the existence of these two written norms for increasingly mutually incomprehensible varieties in each Late Latin/Early Romance-speaking community gradually altered peoples’ linguistic consciousness (WRIGHT, 1991, 2003, 2008; WOOLARD; GENEVESE, 2007; TEJEDO-HERRERO, 2009). The written modalities came to be seen as two separate languages rather than the formal and the vernacular written varieties of the same language (WRIGHT, 2002, p. 262-273, LLOYD, 1996, JENSEN, 1996).

Applying sociophilology to the emergence and development of the concept of LG and other associated terms, e.g., Língua Brasílica, Tupi, Tupinambá, Old Tupi, etc., we can demonstrate that, just as Wright showed that “Latin” and “Romance” were not distinguished before the Carolingian and Gregorian Reforms, contemporaries did not see ‘the LB’ as something different to ‘the LG’ or that “Tupi” and *Lingua Geral Paulista* (LGP) were

conceived of as different languages to “Tupinambá” and *Língua Geral Amazônica* (LGA), as proposed by Rodrigues (1986, 1996, 2010).

### **A note on nomenclature**

We do not employ ‘Tupinambá’ as a generic term for the language of the coastal polities, as has been typical of anthropologists since Métraux and Fernandes (MÉTRAUX, 1948, p. 95, 96-98, 128, see also MÉTRAUX, 1928, 1979; FERNANDES, 1948), and linguists, following Rodrigues, e.g., Rodrigues (1958, 1958/59, 1984/1985, 1986, 1996). Instead, following Navarro (2008), we call the TG linguistic and cultural complex of the Atlantic coast ‘Old Tupi’ (henceforth, OT) to emphasise the perceptions of contemporary Europeans, such as Anchieta (1596), Cardim (1584), and Soares de Souza (1587). This interpretation also seems to reflect better the way that the indigenous groups perceived themselves in relation to other peoples, insofar as the scant evidence permits, as we shall see.

However, in Maranhão, Portugal’s Amazonian colony, founded in the 17<sup>th</sup> century, unlike the *Estado do Brasil*<sup>1</sup>, ‘Tupinambá’ is appropriate because it was the endonym of the largest polity the Portuguese initially encountered around the island of São Luís and along the coast westwards. These Tupinambá had migrated from the northeastern coast, as the Jesuit Manoel Gomes related in 1616, soon after the Portuguese conquered the region (GOMES, 1904, p. 329; see also HEMMING, 1995, p. 213). This migration probably occurred sometime around 1535, as the Portuguese occupation intensified. We use ‘Maranhão Tupinambá’ (henceforth, MATB) to distinguish this group from the ‘Bahian Tupinambá’ (henceforth BATB), spoken on the eastern seaboard, and from the ‘*Tamoio* Tupinambá’ (henceforth, TTB), spoken between the Guanabara Bay (Rio de Janeiro State) and Ubatuba (São Paulo state).

Our preference for OT as a generic is threefold. First, only a subset of historical speakers identified themselves as Tupinambá (NAVARRO, 2008, p. 11, see also

---

<sup>1</sup> The *Estado do Brasil* was established in 1548 by king João III of Portugal to centralise the administration of unsuccessful the hereditary captaincies granted by the Crown to private individuals from 1534, which were placed under the *Governador Geral* based in the first colonial capital, *São Salvador da Bahia de Todos os Santos*, around the midway point on the coast. Thus, in 1616, Gomes was not in ‘Brazil’ but in the recently conquered ‘Maranhão’, that had not yet been colonised by Portugal. The *Estado do Maranhão* was formally separated from the *Estado do Brasil* in 1621. The successive titles of the administrative units are the captaincy of Maranhão (1621-1654), the state of Maranhão and Grão-Pará (1654-1751), the state of Grão-Pará and Maranhão (1751-1772/74), the state of Grão-Pará and Rio Negro (1774-1823) and the State of Maranhão and Piauí (1774-1811), the provinces of Grão-Pará, São José do Rio Negro (1823-1832) and States of Maranhão and Piauí (1811-1850), the province of Grão-Pará, which contained the district (*comarca*) of *Alto Amazonas* (1832-1850), and the provinces of Pará, Amazonas, Maranhão, and Piauí (1850-1889). The modern title of ‘state’ was reapplied on the proclamation of the Republic in 1889.

EDELWEISS, 1969, p. 69-111). Secondly, the modern descendants of non-Tupinambá OT-speaking peoples such as the Potiguara and Tupiniquim dislike Tupinambá being employed as an umbrella term for the linguistic and cultural tradition they identify with (Romildo Araújo Guyraakanga Potiguara, p.c., Tiago Matheus Kaûê Tupinakyâ, p.c.), preferring ‘OT’. Thirdly, Anchieta’s use of ‘Tupi’ (*Na aldeia de Guaraparim*, ll. 183-189) shows it could be used generically, despite referring to the São Vicente variety too (NAVARRO, 2008, p. 11-12, see also EDELWEISS, 1969, p. 69-108; RODRIGUES, 1984/5, RODRIGUES; CABRAL, 2002).

### **The textual record**

The nomenclature changes in texts written between the 16<sup>th</sup> and the 18<sup>th</sup> century (AVILA, 2021, AYROSA, 1950, p. 9-16; BARROS; MONSERRAT, 2015, pp. 239-40; DIETRICH, 2014, p. 596-8; EDELWEISS, 1969, p. 138-165; RODRIGUES, 1985, p. 96). Earlier works, except Anchieta (1596), who uses ‘the language most spoken on the coast of Brazil’, use LB, e.g., Figueira (1621 [1686]), Araújo (1618 [1687]), Bettendorf (1687), the anonymous *Vocabulário na Língua Brasílica* (AYROSA, 1935; NAVARRO, 2013). Later works mostly use LG, e.g., *Dicionário Português - Língua Geral e Língua Geral - Português*. (Anonymous, 1756, Trier, 1136/2048), *Doutrina christãa em lingoa geral dos Indios do Estado do Brasil e Maranhão, composta pelo padre P. Philippe Bettendorf, traduzida em lingoa geral e irregular, e vulgar uzada nesses tempos*, (Anonymous, 1750s), the anonymous *Diccionario da Lingua Geral do Brasil* (1771) and the anonymous, undated *Gramática da Língua Geral do Brazil com hum Diccionario dos vocabulos mais uzuaes para a inteligencia da dita Lingua* (Biblioteca da Universidade de Coimbra, ms. n<sup>o</sup>. 69).

Two other 18<sup>th</sup> century texts use *língua* ‘language’, probably with the adjective ‘general’ understood: João de Arronches’ *Caderno da língua* of 1739 (AYROSA, 1935), the anonymous *Prosódia. Dicionário da língua falada por índios do Brasil* (Academia de Ciências Lisboa, ms. n<sup>o</sup>. 569) from the 1750s, as is Anselm Eckhart’s *Vocabulario da lingua. Brazil* (Biblioteca Nacional de Lisboa, códice 3143) and his *Specimen da lingua brasilica vulgaris*, from 1778 (ECKHART, 1994) in which the Latin adjective *vulgaris* suggests Portuguese ‘geral’, i.e., ‘vulgar’, ‘popular’. Frei Onofre’s *Diccionario Portuguez-Brasiliiano e Brasiliiano-Portuguez* of 1751 is uniquely named (DICIONÁRIO, 1896; PRAZERES, 1891). However, LB occurs in the four texts preserved in the British Museum manuscripts of 1757 (British Library, ms. 223,

FRANÇA, 1859; AYROSA, 1950)<sup>2</sup> and in the *Diccionario da Lingua Brazilica* (ms. 94, *Biblioteca da Universidade de Coimbra*).

Rodrigues (1986, p. 96, 1996, p. 5, 8) claims that the term LG emerged the 16<sup>th</sup> century and in Brazil its meaning changed over the 17<sup>th</sup> and 18<sup>th</sup> centuries, when it substituted LB. For Rodrigues, the adjective ‘general’ (*geral*) initially referred to the geographical extent of the language. Then, after new varieties arose from contact with Portuguese amongst the emergent class of European-Amerindian mestizos (*mameluco* ‘mamluk’), the expression LG came to refer to those novel varieties (RODRIGUES, 1996, p. 6, see also ARGOLO, 2016, p. 90-93). In Amazonia, LG was also used, according to Rodrigues (1986, 1996), to refer to other TG languages that were structurally similar to the LG. Additionally, Rodrigues (1986, 1996) divides LG into ‘São Paulo/Southern’ (LGP), descending from the ‘Tupi’ spoken on the coastal plain of São Vicente and the interior plateau of São Paulo and ‘Amazonian/Northern’ (LGA), later known as Nheengatu (NHG), which evolved from ‘Tupinambá’, i.e., the varieties spoken along the Atlantic coast from the present-day state of Rio de Janeiro to around the state-line between modern Ceará and Paraíba and in Maranhão and Grão-Pará. This ‘Tupinambá’ formed the base for Jesuit norm, although Anchieta first learned OT in the south, he later moved north to Bahia and Espírito Santo, where other Jesuit linguists were already working (EDELWEISS, 1969).

Rodrigues claims no LG developed in the central coastal region between Rio de Janeiro in the south and Ceará in the north because the Amerindian population was rapidly either killed, driven away, or died from disease. This resulted in few Mamluks being born and enslaved Africans and their descendants with Europeans soon became the dominant demographic segment in the region, which prevented the formation of an OT based LG<sup>3</sup>.

Rodrigues (1986, 1996) essentially reworks Edelweiss (1969), in which (pre-contact) ‘Tupi’, i.e., Rodrigues’s ‘Tupi’ and ‘Tupinambá’, was codified as *Língua Brasileira* (1500-1700) and then suffered structural changes from language contact and racial miscegenation to become ‘*Brasiliano*’ (‘Brazilian’) or ‘*Tupi médio*’ (‘Middle Tupi’) or ‘LG’ (1700-1800), before becoming NHG (1800–present) (EDELWEISS, 1969, p. 44-45, 111, 123-158, 158-204).

---

<sup>2</sup> (1) *Doutrina e perguntas dos mistérios principais de nossa Santa Fé na Língua Brasileira*; (2) *Diálogo da Doutrina Cristã pela Língua brasileira*; (3) *Diálogo da Doutrina Cristã pela Língua Brasileira, composto pelo M. R. P. Marcos Antônio*; (4) *Compêndio da Doutrina Cristã que se manda ensinar com preceito, Ano de 1740* (AYROSA, 1950, p. 10).

<sup>3</sup> For an alternative analysis, see Argolo (2011, 2012a/b, 2016) and Finbow (2022). Additionally, the term LG has also been applied to the ‘*Língua Geral d’El Mina*’ (‘Elmina lingua franca’) based on West African FonGbe languages that was spoken across the Minas Gerais goldfields in the 17<sup>th</sup> and 18<sup>th</sup> centuries, and in modern times to refer to a Guarani variety spoken in São Paulo and restructured Portuguese varieties influenced by African and Amerindian languages (ARGOLO, 2016, p. 11, see also MATTOS E SILVA, 2004, p. 78).

Argolo (2016) essentially follows Rodrigues (1986, 1996) but adds creolisation to the mix in the case of LGA. As we shall see, these classifications are all founded on perceptions of *Abstand*.

### ***Abstand* in OT, LB and LG**

Rodrigues' separation of 'Tupi' and 'Tupinambá' arises from Anchieta's statement that lexical roots in the region south of Rio de Janeiro did not exhibit final consonants that existed in northern varieties (RODRIGUES, 1958/59, 1985, 1986, 1996, 2010), e.g., *Apâb* ~ *Apâ*, *Acêm* ~ *Acê*, *Apên* ~ *Apê*, *Aiûr* ~ *Aiú*, i.e., /a-/ 'first-person singular subject, active class', plus the lexical roots /pa(β)/ 'terminate' /sẽ(m)/ 'leave', /pẽ(n)/ 'break', /ju(r)/ 'come' (ANCHIETA, 2014 [1596], p.1, see also EDELWEISS, 1969, p. 76 ff.). Such consonantal apocope was almost certainly a southern areal phenomenon for Montoya records synchronic alternation of apocoped and unapocoped allomorphs in old Guarani, e.g., (*h*)*endu(v)*- /(*h*)*e'nu(β)*/ 'hear (sth.)', *a(r)* /*a(r)*/ 'take (sth.)', equivalent to the northern OT variants (*s*)*endub* and (*t*)*ar* i.e., /(*s*-)*e'nuβ*/, /(*t*-)*ar*/ (MONTOKYA, 2011 [1639], p. 163; NAVARRO, 2013, p. 102, 160).

In São Paulo OT, as in Old Guarani, suffixed lexical roots retained their ancestral final consonants, e.g., *Ubatuba* /uʔuβa-tiβ-a/ arrow-abundant-REF 'abundance of arrow (cane)', *Itatiba* /ita-tiβ-a/ stone-abundant-REF 'abundance of stones' (toponyms from São Paulo State). However, all modern Guarani varieties have reanalysed their lexical roots as possessing only open syllables, e.g., (*Tekoa*) *Itaty* /ita-ti/ stone-abound 'abundance of stones' (Morro dos Cavalos, municipality of Palhoça, Santa Catarina State). On the other hand, NHG, the descendant of MTB, generally retains the ancestral final consonants through fossilising the OT referential suffix /-a/ or vocalic paragoge, e.g.,

(1) OT *taba* /taβ-a/ 'village' > NHG *tawa* /'ta.wa/,

OT *ygara* /i'ar-a/ 'canoe' > NHG *igara* /i'ga.ra/'canoe',

OT *pa'i* /pa'ʔi/ 'father!' (voc.) ~ PT *pai* /'pai/ 'father' > NHG *paia* /'pa.ja/ 'father'

(cf., PT *mãe* /'mãi/ > NHG, *manha* /'mãja/),

OT *sem* /'sem/ 'leave' > NHG *sému* /'se.mu/ 'leave',

OT *syk* /'sik/ 'arrive', 'approach' > NHG *sika* /'si.ka/.

Northern OT used different morphemes on stative and active verbs to express the 'circumstantial indicative' or 'indicative II', that marks subordination with a third-person subject when non-arguments undergo topic-focus (VIEIRA, 2014; NAVARRO, 2008, p. 191;

RODRIGUES, 2010b, p. 25, 38, 41). In the north, *-i/-j/* appeared on consonant-final roots and *-û /-w/* followed vowel-final roots on both stative and active verbs. Southern varieties marked this construction with *-i/-û* on active verbs, but stative verbs employed *-(r)amo*<sup>4</sup>.

A further difference that has been suggested in the LG phases is in the articulation of the high, central, unrounded vowel phoneme /i/. In the Rio Negro LG, the Arawak substrate caused this vowel to merge with /i/ (or occasionally /u/), e.g., NHG *pisirũ* /pisi'rũ/ 'help' < OT *pysyrõ* /pisi'rõ/. In the south, the scant data suggests that /i/ merged with /u/, e.g., *putúnami* 'become night' (PT *enoitecer*, GER *Nacht werden*) < OT /pi'tun-eme, pi'tũ-reme/ night-at (time) 'at night[time]' (MARTIUS, 1867, II, p. 190-122), cf., NHG *pituna ramé* 'at night', Nhandewa Guarani, *pyntũ ramõ*, Mbyá Guarani, *pytũ ramõ*, Paraguayan Guarani, *pyhare ramõ* 'idem'; *oçuca*, *açuc* < OT /o-, a-'sik/ 3A-, 1SG.A-reach 'suffice' (PT *bastar*, GER *genügen*), also 'come close to', 'arrive'. However, clear evidence of /i/ > /u/ as a systematic process is scanty, as northern varieties also exhibited /pu'tũ(n)/ for 'night'. In toponyms, the OT morpheme *tyb* /tiβ/ 'abound', appears written *-tiba*, *-tiva*, *-ndiba*, *-ndiva* and *-tuba*, *-tuva*, *-nduba*, *-nduva* in the north and the south<sup>5</sup>.

Thus, structural differences certainly did exist between two broad diatopic blocs of OT. However, it is unclear how such (fairly minor) structural divergence was conceptualised by contemporary speakers, both native and non-native, which will be investigated in the following section.

## Contemporary perceptions of linguistic diversity in Old Tupi

Sadly, no contemporary testimony exists of any OT-speaking people's opinions about any linguistic similarities or differences they detected. However, the fundamental distinction

<sup>4</sup> Thus, stative verbs' Indicative II was formally identical to the enclitic translative postposition and the gerund for stative verbs. In TG, 'gerunds' express focus/theme/comment, referencing events with the same subject as the main predicate (topic/theme).

<sup>5</sup> The alternation between <t> and <nd> reflects the perception of TG nasal harmony in the phoneme /t/: [t] / [-nasal] + \_\_, [n<sup>d</sup>] / [+nasal] + \_\_. The alternation of <v> and <b> is the result of the uncertainty in representing the OT bilabial fricative or approximant ([β], [v]) in Portuguese orthography, in which <b> manages to capture the bilabial feature but lacks the frication, as Portuguese /b/ is a stop, and <v> registers the fricative articulation but loses the bilabiality, as Portuguese /v/ is labiodental. Examples of the geographical extent of these variants can be seen in *Catanduva* /kaʔa-ãtã-'tiβ-a/ 'abundance of hardwood' (Minas Gerais State), *Curitiba* /kuri-'tiβ-a/ 'abundance of araucaria pines' (Paraná State), *Itatiba* /ita-'tiβ-a/ 'abundance stones' (São Paulo State), *Ubatuba* /uʔuβa-'tiβ-a/ 'abundance arrow [cane]' (São Paulo State), *Taquarenduva* /tak<sup>w</sup>ar-ẽʔẽ-'tiβ-a/ 'abundance of sugarcane' (São Paulo State), *Comandatuba* /komana-'tiβ-a/ 'abundance of beans' (Bahia State), *Brejetuba* /maraja-i(β)-'tiβ-a/ 'abundance of brejaúba [*Astrocaryum aculeatissimum* (Schott) Burret and *Toxophoenix aculeatissima* (Schott)] palms' (Espírito Santo State), *Aratuba* /ara-'tiβ-a/ 'abundance of macaws' (Bahia and Ceará States) and *Jiribatuba* /jara-iβa-'tiβ-a/ 'abundance of jerivá (*syagrus romanzoffiana*) palms' (Bahia State), *Ibicuituba* /iβi-kuʔi-'tiβ-a/ 'abundance of sand' (Maranhão State).

that was recorded as having been made between themselves and the ‘*Tapuia*’, i.e., peoples considered non-OT by culture and/or language (CARDIM, 2009 [1584], p. 205-206), TG-speaking or otherwise, would suggest that OT-speakers perhaps perceived themselves rather like the Germanic and Hellenic peoples in ancient Europe did, i.e., as a broad cultural and linguistic bloc whose members opposed themselves to other cultures on the basis of certain shared characteristics, despite sociocultural and linguistic diversity, and even internecine warfare, between in-group polities.

Contemporary reports demonstrate that the diatopic variation detected by Europeans was not seen as different languages amongst the main peoples on the coast. Anchieta (1989 [1584], p. 59) separates the Carijó (a Guaranian people) from the speakers of the ‘language most spoken on the coast of Brazil’ but mentions no regional subdivisions beyond the reference to consonantal apocope discussed above. This is despite his having lived with the Tupi in São Paulo and São Vicente for many years, amongst the Tamoio (TTB) and BATB, as well as the Tupiniquim. His writings exhibit both southern and northern features. Cardim (2009 [1584], p. 101) states flatly that the ten coastal (OT) peoples, i.e., Tupi, Tupiniquim, Tamoio, Temiminó, Marakajá, Tupinambá, Tupinaé, Caeté, Tobajara, Potiguara, all speak the same language. For Soares de Souza (2010 [1587], p. 406), the linguistic differences between the Tupinaé and the BATB were comparable to Coimbra and Beira Portuguese, i.e., diatopic varieties of the same language.

*Língua Brasileira* (LB) is the term first used in grammars and catechisms of OT, but it is important to note that LB referred to any Brazilian indigenous language, not just OT or even TG. For example, in the interior of the Northeast, where there were few (if any) OT speakers historically, catechesis was carried out in two closely related ‘Kariri’ or ‘Kiriri’ languages (< OT *kyriri* ‘be silent’) known today as Kipeá (MAMIANI, 1698, 1699) and Dzubucua (NANTES, 1709; QUEIROZ, 2008, 2012). Mamiani’s catechism and grammar refer to ‘the LB of the Kiriri nation’: *Catecismo da doutrina christãa da língua brasileira da nação Kiriri...*, *Arte de grammatica da língua brasileira da nação Kiriri*. Nantes referred to his catechism in Dzubucua as an ‘Indian catechism’ (*Catecismo índico*), demonstrating the original use of the term LB.

In 1605, Pero Rodrigues, Jesuit Provincial of Brazil from 1594 to 1603, recorded that ‘This language is the general (one), beginning above the Maranhão River... as far as Paraguay’<sup>6</sup>. (EDELWEISS, 1947, p. 29, RODRIGUES, 1996, p. 7). Once again, Guarani is excluded from

---

<sup>6</sup> ... esta língua he a jeral comensando arriba do rio Maranhão e correndo por todo o distrito da Coroa de Portugal atee o Paraguay.

LG, perhaps because at the time it was located mainly beyond the Portuguese domain. Shortly afterwards, the Jesuit Manoel Gomes wrote from Maranhão in 1616 that, ‘There are many *tapuias* [i.e., non-OT peoples] of many nations, of which fourteen speak the Tupinambá lingua franca, which is almost universal [*comum*] in Brazil’ (GOMES, 1904 [1616], p. 334; SANTOS, 2011, p. 10)<sup>7</sup>. Gomes, and later Antônio Vieira, compare the Guajajara people with the Guaranian Carijó, showing that structural parallels were noted between the speech of non-OT TG-speaking peoples in Maranhão and the main languages of the *Estado do Brasil*. Note, however, that even before any Jesuit publications with LB in the title, Rodrigues and Gomes use LG. This proves that the Jesuits already saw the largest indigenous language (LB) on the coast, i.e., OT, as an LG despite the titles of their 17<sup>th</sup> century publications. Thus, the idea that the term LG arose much later is not borne out.

Rodrigues, Gomes and Vieira use LG as Rodrigues (1986, 1996) proposes was its original meaning, i.e., a large geographical area occupied by a language or dialect continuum. This was the primary criterion for selecting a language to be a diocese’s administrative lingua franca according to the practices developed over the mid-16<sup>th</sup> century in Spain’s American conquests (MADUREIRA, 1977, ZAVALA, 1977, RAMOS PÉREZ, 1986, ALTMAN, 2003, ALFARO LAGORIO, 2003, PÉREZ PUENTE, 2009, DIETRICH, 2014). In the Portuguese sphere, the Jesuit norm was from the outset a hybrid of formalised aspects of diverse OT diatopic varieties (EDELWEISS, 1969, p. 73-79) following the standard Jesuit practice of lingua franca creation, probably thought to be intelligible on the broadest scale (EDELWEISS, 1969, p. 72-79). This is the systematization published by Figueira in 1621, which became the main teaching grammar, in conjunction with Araújo’s catechism of 1618 and the anonymous *Vocabulário na Língua Brasileira*, also from 1621. It is possible that increased institutional contacts during the Iberian Union (1580-1640) may have spread the Spanish model of colonial linguistics amongst the Jesuits sent to Brazil with the University of Salamanca as a major centre of diffusion (FINBOW, 2022; BARROS, 2023, p.c.).

Despite there being a very large region between central Maranhão and the easternmost part of the northeast coast that was inhabited many non-TG peoples, and where no OT speakers are recorded, the missionaries appear to have envisaged the linguistic situation in Maranhão as essentially a continuation of that found in the *Estado do Brasil* (FREIRE, 2011, p. 43), i.e., BATB, TTB and other OT-speaking peoples. Their perception was justified: the MATB and

---

<sup>7</sup> *Ha muitos tapuyas de muitas nações, das quaes quatorze fallão a língua geral dos Tupynambás, que é quase commum no Brazil.*

the Tobajara of the Ibiapaba hills in Ceará were recent incomers who had fled from the eastern coast a few of generations previously.

Rodrigues' idea that there were two OT-based LGs is dealt a severe blow by Antônio Vieira's attempts in the 1650s and 1660s to recruit native LG-speaking *Paulistas* to work in Amazonia, presumably because of their fluency resulted in greater success in converting indigenous peoples (BARROS, 2003, 2010). This strongly suggests that southern LG was not seen as a distinct language from what was spoken in the early Amazonian missions, despite the existence of some structural differences, as we have seen.

Gomes' reference to LG-speaking *Tapuias* in Maranhão probably refers to the non-OT TG-speakers, who were very numerous on the southern shore of the lower Amazon. The Jesuits classified these peoples as '*índios de língua geral*' (lingua franca-speaking Indians), despite their diverse ethnic affiliations<sup>8</sup>, as illustrated by the list of ethnonyms in Johann Philipp Bettendorf's *Crônica da missão dos padres da Companhia de Jesus na Província de Maranhão* (BETTENDORF, 2010 [1698]), e.g., *Guajajara, Juruna, Curuba, Tocantim, Naimiguara, Usaguara, Pacajá, Nambiquara, Coatinga, Guauara, Poquiguara, Guaiapi, Taconhapé, Aruaqui*.

Because the groups mentioned by Gomes and Bettendorf were still unmissionized and had had little, if any, contact with the colonists, it is highly unlikely that they already knew mamluk MATB or the Jesuit norm. Thus, LG in the 17<sup>th</sup> c. Amazon was definitely not identified with the speech of the mamluks and other racial, ethnic, or structural factors did not define it. Instead, as Finbow argues (2022, p. 82-85), LG seems to encompass any TG language that permitted communication with the missionaries and the East-coast OT-speaking settlers and their slaves and allies. Indeed, given the diversity of TG-speaking peoples that were brought as slaves and allies to Belém and São Luís from 1616 and the mainly TG-speaking peoples 'descended' into the missions as catechumen labourers from 1650 onwards, a TG-koine, with MATB as the major contributing variety would have been the most likely outcome (FINBOW, 2022; NOLL, 1999, 2008; see also BARROS, 2003; MUFWENE, 2003, 2008).

### ***Mamelucos or Tapuias?***

---

<sup>8</sup> The Guajajara are conspicuously non-OT speaking to this day. Indeed, none of the TG languages spoken today in the Brazilian states of Pará and Maranhão, e.g., Guajá, Zo'ê, Anambé, Xingu Asuriní, Tocantins Asuriní (Suruí), Tapirapé, Araweté, Avá Canoeiro, Apiaká, Kayabi, etc., belong to the same branch as OT in any phylogenetic classification of TG, e.g., Rodrigues (1985), Jensen (1999, p. 126, 130-132), Rodrigues; Cabral (2002), Galúcio et al. (2015), Michael et al. (2015), Mello; Kneip (2017).

Rodrigues uses ‘mamluk’ as a synonym of ‘mestizo’ in the modern Brazilian manner. However, this is an anachronism, for *mameluco* in 16<sup>th</sup> and 17<sup>th</sup>-century Brazil meant someone born outside official matrimony whose mother was classed as an “Indian” and whose father was legally “white” and publicly recognised his offspring. This meaning dropped out of circulation the 18<sup>th</sup> century (MONTEIRO, 1994, p. 166-167). Such paternal recognition guaranteed certain freedoms that were unavailable to unrecognised illegitimate offspring, who were simply called *bastardos* (‘bastards’). Thus, on the one hand, the mamluks’ (limited) social privileges probably did allow greater access to Portuguese and therefore did favour bilingualism, as Rodrigues (1996) claims, which can be an important factor in stimulating language change. On the other hand, Rodrigues’s indiscriminate use of *mameluco* hides fact that it was not employed in the same way in the past. Moreover, no contemporary sources identify the mamluks as speaking a characteristic variety.

Another issue with Rodrigues’ attribution of linguistic structural change that ‘created’ LG to the mamluks is that their emergence as a social category is much earlier than the emergence of the name LG in the *Estado do Brasil*. Moreover, in Maranhão, as we have seen, Gomes writes of *Tapuias*, not mamluks, speaking ‘Tupinambá LG’ in the year of the Portuguese conquest, too early for there to have been significant miscegenation of Europeans and Amerindians. Thus, with the exception of Rodrigues (1996, p. 5) and Carneira (2006), most treatments of LG, especially in Amazonia, e.g., Câmara Jr. (1978 [1972], p. 28)<sup>9</sup>, Freire (2008, 2011), Houaiss (1985), Noll (1991), Teyssier (2007, p. 96), , see also, Argolo, (2016), Lee, (2005, 2014), Monserrat; Barros (2015), Reich (2003), Schmidt-Riese (2003), identify the incorporation of non-TG-speaking *Tapuias* into the mission villages on a massive scale as the main catalyst for structural change. They see the *Tapuia* learning Jesuit norm and/or MATB as an auxiliary language in an unstructured manner as what ‘converts’ LB into LG and then causes LG to ‘evolve into’ NHG.

What is certain is that LG was acquired via two routes in the Amazonian missions from the 1650s: formal rote catechesis by the missionaries in the codified variety and immersion in the vernacular through cohabitation with speakers (BARROS, 2015, §37-40).

*[...] catecismo acabado se sentavão todos a ouvir uma pregação ou exortação, a qual se fazia na lingua geral dos índios como também o catecismo, e orações*

---

<sup>9</sup> NB Rosa (1990, 1992) and Edelweiss (1969, p. 38-54) for further criticisms of Mattoso Câmara’s characterisation of LG in *Introdução às línguas indígenas brasileiras* (Rio de Janeiro: Universidade do Brasil, 1965).

*eram compostas na mesma língua, e a dita exortação se acomodava sempre ao Evangelho*<sup>10</sup>.

*Resposta aos capítulos que deu contra os religiosos da Companhia em 1662 o procurador do Maranhão, Jorge de São Paio* (cit. BARROS; BORGES; MEIRA, 1996, p. 195)

The *Resposta* shows that LG refers to not only the language of catechesis – which we can be confident was the Jesuit written norm, given it would have been Araújo’s *Catecismo brasilico* of 1618 – but also to the language of the sermon.

Barros, Borges & Meira (1996) cite this as evidence for the use of ‘LB’, i.e., the Jesuit norm, regardless of the linguistic affiliation of a mission’s inhabitants. However, the *Resposta* is from 1662, when the missions were still predominantly inhabited by OT and other TG speakers, for the Portuguese were essentially only active below Gurupá at the mouth of the Xingu in the first decades of the colony (SARAGOÇA, 2000), where TG-speaking peoples were in the majority<sup>11</sup>. Given the structural proximity of TG languages, catechesis in the codified variety is unlikely to have been problematic. Indeed, we see considerable continuity in writing between the 16<sup>th</sup> and 17<sup>th</sup> centuries. Bettendorf’s *Compêndio* of 1687 diverges little from Araújo’s 1618 catechism (EDELWEISS, 1969, MONSERRAT, 2003; MONSERRAT; BARROS; MOTTA, 2010, MONSERRAT; BARROS, 2015, AVILA, 2021, see also RODRIGUES; CABRAL, 2010).

Around 1660-70, however, epidemics began and intensified over the rest of the century and into the following one, decimating the original inhabitants of the missions and households (HEMMING, 1987; FINBOW, 2022, p. 96-67). From the third quarter of the 17<sup>th</sup> c. as missionary activity also pushed beyond the Tapajós river and into the middle Amazon, where TG languages were largely absent. These factors combined to cause major demographic restructuring in the Amazonian colony.

By 1720, the mission population was estimated to be just under 55,000, with a further 20,000 ‘mamluks and slaves’ (RAIOL, 1900, p. 192; see also FREIRE, 2011, p. 68; HEMMING, 1995, p. 421; FINBOW, 2022, 97, 2023a). Another major territorial expansion occurred in the 1730s, after the Manao War (1728-1730), which saw missionaries and slavers

---

<sup>10</sup> ‘[When] Catechesis [is] over, everyone sits down to hear a sermon or exhortation, which was done in the lingua franca of the Indians as also the catechism and prayers were composed in the same language and said exhortation was always based on the Gospel’.

<sup>11</sup> In the large initial ‘descents’ into the early Amazonian missions, Bettendorf records Karajá (a Macro-Jê language), and ‘*joanes ou sacacas, aruãs, mapuases, mamaianazes, pauxis e bocas*’ (BETTENDORF, p. 105), that is to say, the so-called *Nheengaíba* (‘bad speech/language’) peoples, also called *línguas travadas* (‘trapped/jammed tongues’) in Portuguese, a linguistically unidentified confederation from Marajó.

bring huge numbers of non-TG-speaking indigenous peoples, especially from the Rio Negro, into the LG-speaking colonial centres on the lower Amazon. This mass of mostly Arawak-speaking L2 learners might possibly have initiated certain structural changes detected between early- and late-19<sup>th</sup> century Nheengatu, such as post-verbal object pronouns, although later influence from vernacular Brazilian Portuguese was probably much stronger (FINBOW, 2022, p. 98, 2023a). The result of these events led to next phase in the meaning of LG.

In Bluteau's *Vocabulário português e latino* (1712-28), which built on Jesuit descriptions of LG, he contrasts *Línguas Gerais*, 'spread by conquest and commerce', and *Línguas Particulares* (specific languages), spoken by 'isolated, barbarous nations' (BARROS, 2015). Moreover, the speakers of LG in the Amazon, according to Bluteau, are the *Tapuia*. This shows that in the 18<sup>th</sup> century, LG is not thought of as a geographically extensive language or dialect continuum but rather as a supra-ethnic lingua franca used as an auxiliary language by many peoples. Evidently, this change in status from communities of predominantly native-speakers, each using their own TG vernacular, to a society containing very large numbers of non-native speakers had structural repercussions.

### **LG as Creole?**

Clear proof that vernacular LG had evolved structurally comes explicitly from Daniel (2004 [1757-76]) and indirectly from the manuals written by the *tapuiatinga* ('white tapuia') missionaries, i.e., the northern and central European Jesuits, in the mid-18<sup>th</sup> century (DIETRICH, 2014; MONSERRAT, 2003; MONSERRAT; BARROS, 2015). However, none of these sources mention the mamluks. Instead, Daniel explicitly identifies the *Tapuia* as the group that 'corrupted' the 'true' 'Tupinambá' LG, which he equates with the 'Art', i.e, Figueira's 1621/1687 grammar. Nevertheless, Daniel also says that the 'corrupt' LG is spoken 'in all the Portuguese missions of the Amazon' and that few speak the 'true' LG of the Tupinambá 'in its native purity and vigour' because the 'first and true Tupinambá are already almost entirely extinguished' (DANIEL, 2004 [1757-76], v. 2, p. 365, see also FINBOW, 2022; MONSERRAT, 2003; DIETRICH, 2014). Thus, he reveals that even in old missions established in the mid-1600s with MATB and TG native speakers the vernacular is not like the missionaries' 'Art'. This suggests that 18<sup>th</sup> century Portuguese Amazonia should be seen as a diglossic society (LEE, 2005; FINBOW, 2022) because, the tiny contingent of missionaries sought to uphold and preserve their traditional codified variety, which had to be acquired

through formal instruction, for it was not acquired as a first language by any indigenous group (if it ever had been), while in day-to-day life everyone used the evolved vernacular.

Daniel's statement that the *Tapuias*' speech 'seems another, different language' to that the missionaries studied (2004, p. 365) and that the missionaries' codified variety was incomprehensible to them, has been used to suggest that pidginization and/or creolization has occurred in the missions, giving rise to a "new" language, e.g, Lee (2005, 2010, 2014), Argolo (2011a, 2012b, 2016), Oliveira, Zanolli & Modolo (2019)<sup>12</sup>, Castro (1991), Dietrich (2014). Dietrich, in particular develops Freire & Rosa (2003), Argolo (2011b) and Leite (2013), but he retains the emphasis on the mestizo class from Rodrigues (1996). He also discusses structural changes (2014, p. 613-617).

Yet, Daniel never exemplifies the kind of structural changes he is mentions. Thus, his testimony cannot be understood uncritically as evidence for creolization arising from a pidgin. We concur that a great deal of structural change certainly would be expected to arise in the context of 150 years of widespread unstructured, adult second-language learning accompanied by language shift, but there is also no evidence for a widespread stable pidgin variety of LG at any period. Vernacular LG would have certainly exhibited a very wide array of synchronic individual and collective linguistic competences. For example, the 'pidginised' speech of recent arrivals and infrequent users, both enslaved and free, Amerindian and European, at one extreme but, simultaneously, at the other, native speakers of the contemporary vernacular varieties whose ancestors had been MATB or speakers of other TG languages, even if they no longer identified with those traditional names, and fluent non-native speakers. Thus, lingua franca status should not be understood to mean that LG was always spoken non-natively.

As Finbow (2022, p. 87-90, 2023b) has shown regarding Argolo's (2011, 2012, 2016) and Oliveira, Zanolli & Modolo's (2019), attempts to identify structural features to prove the creolisation of LB into LGA generally suffer from a lack of knowledge of TG languages. and OT in particular. For example, Lee (2014) compares hypothetical sentences in LB with their equivalents in the 'Vulgar' language that she envisages emerging from it. However, her first example sentence – *Na eresendúipe?* – actually means "Don't/Can't you hear it?", not "Don't/Can't you hear me?". The correct sentence in OT is *Nda xe rendu(b)ipe iepé?* Her second sentence, *Ixé nde nheenga*, literally means "I am your speech/words" in OT and modern

---

<sup>12</sup> For the severe problems with the proposal in Vieira, Zanolli e MÓdolo (2019), i.e., *Tupinambá* (pidgin) > *Língua Brasílica* (expanded pidgin) > *Língua geral do Brasil* (creole), and the problems with Argolo's "mesolectal creole" (2016, p. 48; see also ARGOLO, 2011, 2012b), FINBOW (2022, 2023a).

NHG, not “I am speaking to you”, which is (*ixé*) *anhe’eng endébe*, (*gûi’ama* ~ *gûiûpa* ~ *gûitena* ~ *gûitekóbo* ~ *gûikupa*).

A further issue is that, like Rodrigues (1986, p. 104-109; 1996, p. 4-5) and Argolo (2016), Lee compares later 19<sup>th</sup> c. and modern NHG to “classic” Jesuit OT. The NHG spoken in the later 18<sup>th</sup> and early 19<sup>th</sup> century did not exhibit the same object-marking strategies (FINBOW, 2023a). Thus, the correct comparison between her two sentences across three diachronic phases of LG (16<sup>th</sup>-17<sup>th</sup> c., 18<sup>th</sup> c.-1850, 1850-20<sup>th</sup> c.) is

2a) “Aren’t you listening to me?”, “Can’t you hear me?”<sup>13</sup>

OT (1550-1750)	<i>Nda-xe-</i> NEG>1SG.P-	<i>r-endub-i</i> R1-hear<NEG	<i>=pe</i> =Q	<i>îepé</i> 2SG.A
‘Old NHG’ (1750-1850)	<i>Niti[u]=será</i> NEG=POL.Q	<i>se-</i> 1SG.P-	<i>r-endu</i> R1-hear	<i>indé</i> 2SG
Mod. NHG (1850–)	<i>Ti=será</i> NEG=POL.Q	<i>re-/pe-</i> 2SG.ACT-	<i>sendu</i> hear	<i>ixé</i> 1SG.P

2b) “I am speaking to you.”

OT (1550-1750)	<i>(ixé) a-nhe’eng</i> (1SG) 1SG.ACT-speak	<i>endé=be</i> 2SG=to		
	<i>gûi-</i> 1SG.ACT.GER-	<i>’am-a</i> stand-GER	<i>/-îup-a</i> /-lie-GER	<i>/-ten-a</i> /-sit-GER
			<i>/-tekó-bo</i> /-move-GER	
‘Old NHG’ (1750-1850)	<i>(ixé) a-nheẽ</i> (1SG) 1SG.ACT-speak	<i>indéu</i> 2SG=to	<i>a-</i> 1SG.ACT-be	<i>iku</i>
Mod. NHG (1850-)	<i>(ixé) a-nheẽ</i> (1SG) 1SG.ACT-speak	<i>a-</i> 1SG.ACT-be	<i>iku</i>	<i>indé arã[ma]</i> 2SG to

In (2a), despite the substitution of the special form for ‘1<sup>st</sup> person patient and 2<sup>nd</sup> person agent’ (*îepé* /je’pe/ for singular and *peîepé* /peje’pe/ for plural) by the generic second-person free personal pronouns *ine* ~ *indé* /i’ne/ and *penhé* /pe’je/ respectively, LG retains the classic TG person hierarchy with prefixes for the patient argument, as in Modern Paraguayan Guarani: *Nde cherendu* /ne se-r-e’nu/ ‘you.SG 1SG.INACT-POSSM-listen’ “you listen to me” (ESTIGARRIBIA, 2020, p. 139), i.e., analogical levelling has occurred across the paradigms.

<sup>13</sup> GLOSSES: = clitic morpheme boundary, – bound morpheme boundary, >... < circumfix morpheme boundaries, 1 first person, 2 second person, 3 third person, A agent, ACT active conjugation class, GER ‘gerund’ (same-subject, subordinate clause), Q interrogative, NEG negation, P patient, POL polar, PROG progressive aspect, RT root, SG singular, PL plural.

In (2b), even though four physical orientation auxiliary ‘gerunds’ (*îup* /-jup-/ ‘horizontal extension’, *am* /-ʔam-/ ‘vertical extension’, *en* /-en-/ ‘without extension’, *kup* /-kup-/ ‘in motion’) have not survived, the ‘be/exist’ root continues to express continuous aspect, i.e., OT *-ekóbo* /-e’ko-/ + /-aβo/ → /-e’ko-βo/ > Mod. NHG *-iku* /-i’ku/. The auxiliary verb has lost the characteristic OT gerund suffixes on intransitive active roots, i.e., /-a/ on consonant-final roots and /-aβo/ on vowel-final roots (NAVARRO, 2008, p.159-161, 172-173). The special intransitive active class number and person subject prefixes, i.e., *gûi(t)-* /wi(t)-/ ‘1<sup>st</sup> singular’, *e-* /e-/ ‘2<sup>nd</sup> singular’ (2008, p. 161, 176-179) have suffered analogical levelling by the standard 1<sup>st</sup> and 2<sup>nd</sup> person active class subject prefixes, i.e., *a-* /a-/ and *(e)re-* /-(e)re-/, falling in line with the rest of the active class gerund personal prefixes, which are identical to those of the indicative even in OT, i.e., *îa-* /ja-/ ‘1<sup>st</sup> person plural inclusive’, *(o)ro-* /oro-/ ‘1<sup>st</sup> person plural exclusive’, *pe-* /pe-/ ‘2<sup>nd</sup> person plural’, and *o-* /o-/ ‘3<sup>rd</sup> person subject-focus’ (FINBOW; O’NEILL, 2022). This is also attested in Paraguayan Guarani, although the root that survived was ‘sit’, ‘without extension’ (OT RT *-în* ‘sit’ → GER *-en-a*), rather than ‘be/in motion’. It became the progressive aspect particle *hína*, e.g., *rehai hína* /re-h-ai hina/ 2SG.ACT-3P-WRITE PROG “you are writing” (Estigarribia, 2020, p. 164).

Thus, later 18<sup>th</sup> century or early 19<sup>th</sup> century examples from “Old NHG” LG smooth the transition between OT and modern NHG. The changes and the structural parallels between the three diachronic ‘LG’ varieties are in line with tendencies observable in the other TG language in long-term close contact with Ibero-Romance, i.e., Paraguayan Guarani, which is not claimed to be a creole.

## Conclusions

Approaching *Língua Geral* sociophilologically reveals the problems of matching historical periodization and categorizations between languages/varieties on perceived structural differences, for there is no guarantee that modern perceptions of diversity correspond to those of the past (FINBOW, 2010, 2012, 2022). Modern researchers have identified subdivisions identified on the basis of the diatopic, diastratic, diamesic, and diachronic *Abstand* that they have detected in texts. However, although structural differences certainly existed, as is to be expected in a situation of very extensive linguistic and cultural contact and hybridization, we have shown that they were not important to the naming practices employed amongst at least the non-indigenous inhabitants of Portugal’s South American domains between the 16<sup>th</sup>, and 18<sup>th</sup> centuries. Consequently, the role of sociocultural mixing and,

especially in the Brazilian context, racial miscegenation, has been vastly overstated in changes in nomenclature.

Rodrigues (1996) claims that earlier 20<sup>th</sup> century writers had wrongly classified LG as either the language of the pre-colonial indigenous peoples or a new language grammaticised by the Jesuits out of the pre-colonial language or languages, or even a pidgin or creole that arose from contact between diverse indigenous peoples and Europeans (see also ARGOLO, 2011a, 2012b, 2016; LEE, 2005, 2014). However, the technical definition Rodrigues seeks to reserve for LG as the language of the mamluk class is as ‘unfounded linguistically and historically’ (RODRIGUES, 1996, p. 6) as the definitions he criticises. In fact, it seems likely that his proposals regarding the mamluks role as multilinguals in the ‘emergence’ of the LG, stimulated others to investigate the pidgin and creole avenue!

Modern proposals, however well founded in linguistic theory, must also be backed up by as solid evidence as can be obtained in order to avoid anachronisms or other kinds of misrepresentation. In the case of ‘the language most spoken on the coast of Brazil’ and its relatives and descendants, the silence of contemporary commentators on the role of the Mamluk class in generating structural change that caused people to exchange the name LB for LG is quite simply deafening. Thus, Rodrigues’s idea of employing LG to refer to the mamluk variety of OT is deeply unconvincing. An additional complication is there are no known historical records of mamluk speech, so even if one were to accept Rodrigues’ proposals, we would know nothing about the kinds of structural changes they introduced.

Rodrigues’ idea that the term LB was substituted by LG is only epiphenomenal. Pero Rodrigues and Manuel Gomes prove that the names circulated simultaneously, despite the Jesuits’ preferring LB in their earliest publications. Mamiani’s catechism and grammar of Kariri show that LB was still in use for indigenous languages at the end of the 17<sup>th</sup> century.

Where Rodrigues finds documentary support is in LG being used to refer to other TG languages besides OT, in the writings of Gomes, Vieira and Bettendorf. In this regard, Wright’s concept of ‘complex monolingualism’ (1982, p. xi, 1993d: 207-8) is potentially useful. In the Late Latin/Early Romance context, this meant that speech and writing constituted a single conceptual unit for Early Medieval Romanophone peoples, despite wide structural divergences between diaphasic and diamesic modalities. In Portugal’s South American colonies between the 16<sup>th</sup> and 18<sup>th</sup> centuries, the structural differences within LG were primarily diatopic initially, within TG languages. Bluteau and Daniel’s descriptions emphasize diamesic and diaphasic differences that are primarily the result of the diastratic limitation of the Jesuit norm to the missionary class. Thus, another of Kloss’s technical terms could be applied to the Jesuit LG:

*Dachsprache* (lit., ‘roof-language’), i.e., a norm that overarches varieties in a continuum. Such norms are typically deliberately elaborated, i.e., the result of *Ausbau*, e.g. standard ‘High’ German or standard Italian (KLOSS, 1967; MULJAČIĆ, 1989, p. 256 ff.; KREFELD, 2020).

Bluteau’s usage also shows that a semantic shift had occurred in the term LG between the 17<sup>th</sup> and the 18<sup>th</sup> century. LG still refers to the language of catechesis but does not identify a geographically extensive indigenous linguistic bloc of native speakers. Vernacular LG is thought of as a supra-ethnic lingua franca employed alongside the missionized indigenous communities’ native languages. As time went on, it substituted very many of them. This change in function arose from the largely forcible incorporation of innumerable non-OT and non-TG-speaking peoples into the colonial system in Maranhão and Grão-Pará. A similar situation occurred in the northeastern interior of Brazil around the same time. The missionaries’ decision not to catechise in each indigenous people’s native tongue according to the Jesuit ideal but use the codified variety for catechesis preceded the major structural changes that the new policy provoked. Ultimately, the semantic shift detected in LG ended up passing to the term *Tapuia*. From ‘non-OT Amerindian’ it came to mean the ‘detrribalised’ and ‘re-cultured’ LG/NHG-speaking indigenous or mestizo Amazonians of the 19<sup>th</sup> century (FREIRE, 2011). In this way, Hartt registered *Tapuia nheenga* ‘*Tapuia* language/speech’ as a synonym for *Nheengatu* in the 1870s (HARTT, 1938 [1875], ex. 684, 686, 687).

However, starting with speakers of the OT varieties in the *Estado do Brasil* and MATB in the North, through the koineization that arose from cohabitation, slaving, and missionization was an unbroken chain of intergenerational transmission within the oldest Amazonian colonial communities until their shift into Portuguese. The proportion of *Tapuia* learners to native speakers certainly rose very markedly over time (ARGOLO, 2016; FINBOW, 2022, 2023a; FREIRE, 2011; HEMMING, 1987, 1995; LEE, 2005), which would have strongly favoured restructuring, which is very evident in later, 19<sup>th</sup> century attestations of LG/NHG, e.g., Hartt (1875), Couto de Magalhães (1876). Nevertheless, it is hard to see LG/NHG as a wholly ‘new’ language produced by classical creolization processes out of a pidgin, for there were always ample opportunities to gain competency through contact with fluent speakers (FINBOW, 2022, p. 86, 97-100), and no evidence of a stable pidgin exists. Interestingly, the structural changes that can be adduced are more akin to what is seen in vernacular Brazilian Portuguese, which was the closest corollary in the *Estado do Brasil* to the ‘*Tapuia*-filtered’ LG in Maranhão and Grão-Pará (FINBOW, 2022, p. 102).

A modern definition that might be applied interestingly to LG is ‘macrolanguage’, which is defined by ISO 636-3<sup>14</sup> as ‘... closely related individual languages that are deemed in some usage contexts to be a single language’ (ETHNOLOGUE, 2009). The basic criterion for a macrolanguage is a close phylogenetic relationship, which is true of ‘LG as TG’. Additionally, a macrolanguage frequently possesses a classical standard which speakers of several closely related individual languages understand, or which is at least known to be the source of those individual languages, e.g., Arabic, which is divided into over thirty modern spoken varieties, but these exist alongside a common standard (Modern Standard Arabic) and also the classical Quranic norm. In the case of LG, this would be the Jesuit norm. Secondly, a macrolanguage can have a long-lasting and deep-rooted linguistic identity amongst the spoken varieties that have undergone separate developments because of sociopolitical factors, e.g., Serbo-Croat as a macrolanguage comprising the individual Croatian, Serbian and Bosnian languages. This aspect can be applied readily to the different coastal OT varieties with their opposition to the peoples they called *Tapuia*, *Kariri* or *Nheengaíba*. Thirdly, a macrolanguage may comprise several closely related languages which the specialist literature often treats as either a genetic unit of individual languages or which may be subsumed into one ‘language’ for other reasons, e.g., political or ethnic criteria, e.g., ‘Rajasthani’ as the macrolanguage of the individual Indo-Iranian languages Bagri, Gade, Lohar, Hadothi, Malvi and Wagdi. This is similar to the missionaries’ habit of calling all the TG-speaking peoples in Maranhão and Grão-Pará ‘*índios de Língua Geral*’, despite each having its own individual identity.

Even the criticisms made of the category of macrolanguage make it appropriate to apply to LG. For example, the inconsistency of the application by SIL on *Ethnologue* and the lack of a unifying definition to distinguish dialects, dialect clusters, and languages. No agreement or uniformity exists in the overarching linguistic standards in terms of linguistic criteria for identifying when one should retain non-linguistically determined divisions and when one should ignore them. Thus, any classification could very easily lead to problematic results and provoke controversy. This is precisely the same kind of difficulties as the concept of LG has posed to linguists and historians, which is why ‘macrolanguage’ is such an appropriate modern correlate.

**AUTHOR IDENTIFICATION:****NAME:** Thomas Finbow**DEGREE:** D.Phil. (University of Oxford)

---

<sup>14</sup> International Organisation for Standardisation (2023).

**AFFILIATION:** Departamento de Linguística, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo (São Paulo – SP, Brazil)

**E-MAIL:** [thomas.finbow@usp.br](mailto:thomas.finbow@usp.br)

**ORCID:** <https://orcid.org/0000-0003-4803-6603/>

**CRedit Roles:** The above-named author was the sole agent in all relevant areas, namely, Conceptualization, Methodology, Investigation, Resources, Writing – Original Draft. In the case of ‘Writing – Revision and Editing’, the author responded to suggested revisions from the special issue’s organizers.

**COMPETING INTERESTS:** The author declares no competing interests.

**RESEARCH PREREGISTRATION:** the conducted research was not preregistered in an independent, institutional registry.

**DATA ACCESSIBILITY STATEMENT:** The author confirms that the data, codes and materials supporting the findings of this study are available within the article [and/or] its supplementary materials.

**ETHICS AND CONSENT:** Not applicable. This research did not involve human participants.

**DISCLOSURE OF FUNDING SOURCES:** Not applicable.

## References

- ALFARO LAGORIO, M. A. C. Usos e funções das línguas na área andina: o Terceiro Concílio Limense (1582-1583). In: ALFARO LAGORIO; FREIRE; ROSA, p. 41-54, 2012.
- \_\_\_\_\_. Elementos da política lingüística colonial hispânica: o Terceiro Concílio Limense. In: FREIRE; ROSA, p. 43-56, 2003.
- ALFARO LAGORIO, C.; ROSA, M. C.; FREIRE, J. R. Bessa (org.). **Políticas de línguas no Novo Mundo**. Rio de Janeiro: EdUERJ, 2012.
- ALTMAN, C. As línguas gerais sul-americanas e a empresa missionária: linguagem e representação nos séculos XVI e XVII. In: FREIRE; ROSA, p. 57-84, 2003.
- AMMON, U. (Org.). **Status and Function of Languages and Language Varieties**. Berlin: Walter de Gruyter, 1989.
- ANCHIETA, J. de. **Arte de Gramática da lingoa mais usada na costa do Brasil (1595)**. 5<sup>th</sup> ed. facsimile, Salvador, BA: EdUFBA, 2014.
- \_\_\_\_\_. **Cartas. Correspondência ativa e passiva**. 2<sup>a</sup> ed., Sao Paulo: Loyola, 1984 [1553-1596].
- ARAÚJO, A. de. **Catecismo brasílico da doutrina crista. Edição fac-similar da 2. ed. de 1686, corrigida por Bartolomeu de Leão**. Edited by Júlio Platzmann, Leipzig: B. G. Teubner, 1898.
- ARGOLO, W. Línguas gerais na história social-lingüística do Brasil. **PAPIA**, São Paulo, v. 26, n. 1, p. 7-52, 2016.
- \_\_\_\_\_. História lingüística do Sul da Bahia: levantando hipóteses e iluminando caminhos. **Entrepalavras**, v.1, n. 2, p. 270-292, 2012a.
- \_\_\_\_\_. A língua geral da Amazônia como um sistema historicamente novo: jesuítas e tapuias na origem do contexto com interrupção de transmissão lingüística entre gerações. In: MATTOS E SILVA; OLIVEIRA; AMARANTE, p. 479-513, 2012b.
- \_\_\_\_\_. **Introdução à história das línguas gerais no Brasil: processos distintos de formação no período colonial**. Dissertação de mestrado. Universidade Federal da Bahia, 2011a.

- \_\_\_\_\_. Língua geral na Bahia: comarcas de Ilhéus e Porto Seguro. In: CARVALHO; ROCHA; PARCERO, p. 99-109, 2011b.
- ÁVILA, M. Twardwsky. **Proposta de dicionário nheegatu-português**. Tese de doutorado. Universidade de São Paulo, 2021.
- AYROSA, P. **O caderno da língua ou vocabulário portuguez-tupi de Frei João de Arronches, 1739: notas e commentarios á margem de um manuscrito do sec. XVIII**. São Paulo: Imprensa Official do Estado, 1935.
- \_\_\_\_\_. (Org.) **Orações e diálogos da Doutrina Cristã na língua brasílica – mss. do século XVIII, transcritos e anotados por Plínio Ayrosa**. São Paulo: Boletim CVI, Etnografia e Língua Tupi-Guarani Nº. 17, 1950.
- BANNIARD, M. **Viva voce: Communication écrite et communication orale du IVe au IXe siècle en occident latin**. Paris: Institut des études augustinienes, 1992.
- \_\_\_\_\_. The transition from Latin to the Romance languages. In: MAIDEN; SMITH; LEDGEWAY, p. 57-106, 2013.
- BARROS, M. C. Drummond de Menezes. Notas sobre a política jesuítica da língua geral na Amazônia (séculos XVII-XVIII). In: FREIRE; ROSA, p. 85-112, 2003.
- \_\_\_\_\_. ‘O uso do tupi na Capitania de São Paulo no século XVII. Indícios na vida de um jesuíta “língua”’. In: NOLL; DIETRICH, p. 141-153, 2010.
- \_\_\_\_\_. ‘Em razão das conquistas, religião, commercio’. Notas sobre o conceito de língua geral na colonização portuguesa da Amazônia nos séculos XVII-XVIII. **Mélanges de la Casa de Valázquez** 45.1, p. 99-112, 2015.
- BARROS, M. C. Drummond de Menezes; BORGES, L. C.; MEIRA, M. A língua geral como identidade construída. **Revista de Antropologia**, 39.1, p. 191-219, 1996.
- BETTENDORF, J. P. **Crônica da missão dos padres da Companhia de Jesus na Província de Maranhão**. Brasília: Senado Federal, Conselho Editorial, 2010.
- \_\_\_\_\_. **Compêndio da doutrina crista na língua portuguesa e brasílica, etc**. Lisboa: Officio de Miguel Deslandes (reimpressão por Frei José Mariano da Conceição Vellozo, Lisboa: Offic. de Simão Thaddeu Ferreira, 1800).
- BLUTEAU, R. **Vocabulário portuguez, e latino, ...: autorizado com exemplos dos melhores escritores portuguezes, e latinos; ...** 8 v; 2 Suplementos. Coimbra: Collegio das Artes da Companhia de Jesu: Lisboa, Officina de Pascoal da Sylva, 1712-1728.
- BOSSONG, G. **Die romanischen Sprachen. Eine vergleichende Einführung**. Buske, Hamburg 2008.
- CABRAL, A.S. Arruda Camara; RODRIGUES, A. Dall’Igna (eds.), **Línguas indígenas brasileiras: fonologia, gramática e história**, Belém: UFPA, 2002.
- CÂMARA JR., J. Mattoso. **História e estrutura da língua portuguesa**. Rio de Janeiro: Padrão, 1979.
- CARDIM, F. **Tratados da gente e terra do Brasil**. São Paulo: Editora Hedra, 2009.
- CARDEIRA, E. **História do português**. Lisboa: Caminho, p. 87-96, 2006.
- CARVALHO, C. dos Santos; ROCHA, F. Aninger de Barros; PARCERO, L. M. de Jesus (orgs.). **Discurso e cultura: diálogos interdisciplinares**. Salvador: EdUNEB, 2011.
- CASTRO, Ivo. **Curso de história da língua portuguesa**. Lisboa: Universidade Aberta, p. 46-63, 1991.

- DANIEL, J. **Tesouro descoberto no máximo Rio Amazonas**. Rio de Janeiro: Contraponto Editora, 2004.
- DICCIONARIO anonymo da lingua geral do Brasil, publicado de novo com seu reverso por Julio Platzmann. Edição facsimilar**. Leipzig: B. G. Teubner, 1896.
- 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.
- DIXON, R. M. W.; AIKHENVALD, A. (eds.). **The Amazonian Languages**. Cambridge: Cambridge University Press, 1999.
- ECKART, A. O Exemplário da língua corrente do Brasil. Tradução do latim de Carlos Antônio Kalil Tannus e Miguel Barbosa do Rosário. **Terceira Margem**, UFRJ, Rio. ano 2, n. 2, p.176-180, 1994.
- EDELWEISS, F. G., **Tupis and Tupi-Guarani: estudos de etnonímia e linguística**. Bahia: Secretaria de Educação e Saúde, 1947.
- \_\_\_\_\_. **Estudos Tupis e Tupi-Guaranis: confrontos e revisões**. Rio de Janeiro: Livraria Brasileira Editora, 1969.
- ESTIGARRIBIA, B. **A Grammar of Paraguayan Guarani**. London: UCL Press, 2020.
- ETHNOLOGUE, Lewis, M. Paul (Org.). **Ethnologue. Languages of the World**. 16<sup>th</sup> ed., Dallas, TX: SIL, 2009.
- FERNANDES, F. **A organização social dos tupinambá**. São Paulo: Instituto Progresso Editorial, 1948.
- FIGUEIRA, L. **Arte da Lingua Brasilica**. Lisboa: Manoel da Silva, 1621.
- \_\_\_\_\_. **Arte de grammatica da lingua brasilica**. Edição facsimilar, por Julio Platzmann, da 2. ed. de 1687. Lisboa: Miguel Deslandes. Leipzig: B. G. Teubner, 1880.
- FINBOW, T. D. A formação dos conceitos de “latim” e “romance”. In: LAGARES; BAGNO, p. 89-120, 2011.
- \_\_\_\_\_. As categorias metalinguísticas tupinambá/língua geral e latim/romance: alguns paralelos sociofilológicos. In: ALFARO LAGORIO; ROSA, FREIRE, 2012, p. 1-244.
- \_\_\_\_\_. The emergence and nature of the Língua Geral Amazônica’ in accordance with Mufwene’s Language Ecology model. **Revista do GEL**. v. 19, n. 2, p. 75-112, 2022.
- \_\_\_\_\_. **Old Nheengatu? A “lost” alignment from 19<sup>th</sup> century Amazonia and its implications for the evolution of the Língua Geral Amzônica**. Paper to Amazônicas XI, 5<sup>th</sup> – 9<sup>th</sup> June 2023. Universidad de los Andes, Bogotá, Colombia, 2023a.
- \_\_\_\_\_. Uma reanálise de dois casos de mudança estrutural entre o tupi antigo e a língua geral amazônica/nheengatu tidos como exemplos de crioulização e a apresentação de um caso de mudança gramatical no nheengatu do século XIX por contato com o português. In: LIMA-HERNANDEZ; SANTOS; ADRIANO, 2023b.
- FINBOW, T., O’NEILL, P. Koineization and language contact: the social causes of morphological change in and with Portuguese. In: LEDGEWAY; SMITH; VINCENT, p. 381-412, 2022.
- FRANÇA, E. Ferreira. **Chrestomathia da língua brasilica**. Leipzig: F. A. Brockhaus, 1859.
- FREIRE, J. R. Bessa. **Rio Babel. A história das línguas na Amazônia**. 2. ed. Rio de Janeiro: Atlântica, 2011.
- FREIRE, J. R. Bessa; ROSA, M. C. **Línguas Gerais. Política linguística e catequese na América do Sul no período colonial**. Rio de Janeiro: EdUERJ, 2003.

- GALÚCIO, A. V.; MEIRA, M.; BIRCHILL, J.; MOORE, D.; GABAS JR., N.; DRUDE, S.; STORTO, L., PICANÇO, G., REIS RODRIGUES, C. Genealogical relations and lexical distances within the Tupian linguistic family. **Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas**, Belém, v. 10, n. 2, p. 229-274, maio-ago. 2015.
- GOEBL, H. Quelques remarques relatives aux concepts *Abstand* et *Ausbau* de Heinz Kloss. In: AMMON, p. 278-290, 1989.
- 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.
- GÖSCHEL, J.; NAIL, N.; VAN DER ELST, G. (eds.), **Zur Theorie des Dialekts: Aufsätze aus 100 Jahren Forschung. Zeitschrift für Dialektologie and Linguistik**, Beihefte, n. F., v. 16. Wiesbaden: F. Steiner, 1976.
- HARTT, C. F. Notes on the Lingoa Geral or Modern Tupi of the Amazonas. **Transactions of the American Philological Association**, 3, p. 58-76, 1875.
- HEMMING, J. **Amazon Frontier. The Defeat of the Brazilian Indians**. 2nd ed. London: Papermac, 1995.
- \_\_\_\_\_. **Red Gold. The Conquest of the Brazilian Indians**. 2nd ed. London: Papermac, 1987.
- HERMAN, J., Spoken and Written Latin in the Last Centuries of the Roman Empire. A Contribution to the Linguistic History of the Western Provinces. In: WRIGHT, p. 29-43, 1991.
- HOUAISS, A. **O português no Brasil**. Rio de Janeiro: Unibrade - Centro de Cultura, 1985.
- JENSEN, C. Tupi-Guarani. In: DIXON; AIKHENVALD, p. 125-163, 1999.
- JENSEN, T. Language change and metalinguistic change: Latin to Romance and other cases. In: WRIGHT, p. 19-28, 1996.
- KLOSS, H. Abstand languages and Ausbau languages. **Anthropological Linguistics**, v. 9, n. 7, p. 29-41, 1967.
- \_\_\_\_\_. Abstandsprachen und Ausbausprachen. In GÖSCHEL; NAIL; VAN DER ELST, p. 301-322, 1976.
- \_\_\_\_\_. **Die Entwicklung neuer germanischer Kultursprachen seit 1800**. 2<sup>nd</sup> ed., Schwann, Düsseldorf, 1978.
- KREFELD, T., Über ‚Dächer‘, ‚Schirme‘ und Diversität – Sprachsoziologie im kommunikativen Raum. **Korpus im Text**. Serie A, 2020 ([kit.gwi.uni-muenchen.de](http://kit.gwi.uni-muenchen.de)).
- LAGARES, X. C.; BAGNO, M. (org.). **Políticas de norma e conflito linguístico**. São Paulo: Contexto, 2010.
- LEDGEWAY, A.; SMITH, J. C.; VINCENT, N. (eds.). **Periphrasis and Inflexion in Diachrony: a View from Romance**. Oxford: Oxford University Press, 2022.
- LEE, M. K. **Conversing in Colony: The Brasília and the Vulgar in Portuguese America 1500-1759**. Ph.D. Dissertation, John Hopkins University, 2005.
- \_\_\_\_\_. Language and Conquest: Tupi-Guarani Expansion in the European Colonization of Brazil and Amazonia. In: MUFWENE, p. 143-167, 2014.
- LEITE, F. R. **A Língua Geral Paulista e o “Vocabulário elementar da Língua Geral Brasília”**. Dissertação de mestrado, UNICAMP 2013.
- LIMA-HERNANDEZ, M. C.; SANTOS, M. M. Soares; ADRIANO, P. Soma (orgs.), **Estudos de aquisição, Mudança e Ensino-aprendizagem de línguas**. VIIIº Simpósio Mundial de Língua Portuguesa (SIMELP 8) / IIIº Simpósio Interdisciplinar de Estudos de

- Linguagem (SINTEL 3), Brasil – Angola, 4 a 7 de outubro de 2022. vol. 6. São Paulo: Estige Editorial, 2023.
- LLOYD, P. M. On the names of languages (and other things). In: WRIGHT, p. 9-18, 1996.
- MADUREIRA, J.M. **A liberdade dos índios, a Companhia de Jesus, sua pedagogia e seus resultados**. Rio de Janeiro: Imprensa Nacional, 1977.
- MAGALHÃES, J. Vieira Couto de, **O selvagem**. Rio de Janeiro: Typ. da Reforma, 1876.
- MAIDEN M.; SMITH, J.C.; LEDGEWAY, A (eds.). **The Cambridge History of the Romance Languages**. Cambridge University Press, 2013.
- MACKENZIE, D.; MICHAEL, I. (eds.), *Hispanic Linguistic Studies in Honour of F. W. Hodcroft*. Llangrannog: Dolphin Book, 1993.
- MAMIANI, L. V. **Arte de grammatica da lingua brazilica da naçam Kiriri**. Lisboa: Miguel Deslandes, 1699.
- \_\_\_\_\_. **Arte de grammatica da lingua brazilica da naçam Kiriri**. Lisboa: Miguel Deslandes, 1699., 2a. edição (com notas introdutórias de Baptista Caetano de Almeida Nogueira). Rio de Janeiro: Bibliotheca Nacional, 1877.
- \_\_\_\_\_. **Catecismo da doutrina christã na lingua brasilica da nação Kiriri**. Lisboa: Miguel Deslandes, 1698.
- MARTIUS, C. F. P. von. *Beiträge zum Ethnographie und Sprachkunde Amerikas zumal Brasiliens. Glossarium linguarum Brasiliensium*. Leipzig: Friedrich Fleicher, 1867.
- MATTOS E SILVA, R.V. **Ensaio para uma sócio-história do português brasileiro**. São Paulo: Parábola, 2004.
- 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, 2012.
- MELO, A. A. Souza; KNEIP, A. Novas evidências linguísticas (e algumas arqueológicas) que apontam para a origem dos povos tupi-guarani no leste amazônico. *Literatura y lingüística*, (36) Santiago, dic., p. 299-312, 2017.
- MÉTRAUX, A. **La religion des Tupinamba et ses rapports avec celle des autres tribus Tupi-Guarani**. Paris: Librairie Ernest Leroux, 1929.
- \_\_\_\_\_. The Tupinambá. In: STEWARD, J. H. (ed.), **Handbook of South American Indians, V. 3: The tropical forest tribes**. Smithsonian Institution, Bureau of American Ethnology, Bulletin 143. Washington: Government Publishing Office, p. 193-196, 1946.
- \_\_\_\_\_. **A religião dos tupinambás**. São Paulo: Companhia Editora Nacional; Editora da Universidade de São Paulo, 1979.
- MICHAEL, L. D.; CHOUSOU-POLYDOURI, N.; BARTOLOMEI, K.; DONNELLY, E.; WAUTERS, V.; MEIRA, S.; O'HAGAN, Z. A Bayesian Phylogenetic classification of Tupi-Guarani. **LIAMES**, 15.2, p. 1-36, 2015
- MONSERRAT, R. O Tupi do século XVIII (tupi médio). In: FREIRE; ROSA, 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.

- MONSERRAT, R.; BARROS, C. de. Fontes manuscritas sobre a língua geral da amazônia escritas por jesuítas “Tapuitinga” (século XVIII). **Confluência**, 49, p. 236-254, 2015.
- MONTEIRO, J. M. **Os negros da terra: índios e bandeirantes nas origens de São Paulo**. São Paulo: Companhia das Letras, 1994.
- MONTOYA, A. Ruiz de. **Arte de la lengua guarani**. 1640. Asunción: CEPAG, 2011.
- MUFWENE, S. S. **The Ecology of Language Evolution**. Cambridge: C.U.P, 2003.
- \_\_\_\_\_. **Language Evolution. Contact, Competition and Change**. London & New York: Continuum, 2008.
- \_\_\_\_\_. (ed.). **Iberian Imperialism and Language Evolution in Latin America**. Chicago: University of Chicago Press, 2014.
- MULJAČIĆ, Ž., Über den Begriff Dachsprache. In: AMMON, p. 256-277, 1989.
- MÜLLER, K.-L. Zum Fortleben von LATINU und seinen Verwandten in der Romania”. **Zeitschrift für romanischen Philologie**, v. 79, p. 38-73, 1963.
- NANTES, B. de. **Katecismo indico da lingua Kariris**. Lisboa: Officina de Valentim da Costa Deslandes, 1709.
- \_\_\_\_\_. **Katecismo indico da lingua Kariris**. Edição facsimilar, por Julio Platzmann, da primeira edição (1709). Leipzig: B. G. Teubner, 1896.
- NAVARRO, E. de Almeida. **Dicionário de tupi antigo. A língua indígena clássica do Brasil**. São Paulo: Editora Global, 2013.
- \_\_\_\_\_. **Método novo de tupi antigo**. 3. ed. São Paulo: Editora Global, 2008.
- NOLL, V. **Das brasilianische Portugiesisch**. Heidelberg: Universitätsverlag Winter, 1999.
- \_\_\_\_\_. **O português brasileiro: formação e contrastes**. São Paulo: Globo. Tradução: Mário Eduardo Viaro, 2008.
- NOLL, V.; DIETRICH, W. (org.). **O português e o tupi no Brasil**. São Paulo: Contexto, 2010.
- 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.
- 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.
- PRAZERES, Fr. F. de N.S. dos. Poranduba maranhense, ou relação historica da provincia do Maranhão [...] com [...] um dicionario abreviado da lingua geral do Brazil. In: **Revista Trimensal do Instituto Historico e Geographico Brasileiro**, Rio de Janeiro, v. 54, pt. 1, pp. [4]-277, 1891.
- QUEIROZ, J. M. Correia de. **Aspectos da fonologia Dzubukuá**. Dissertação de mestrado. Universidade Federal da Paraíba, 2008.
- \_\_\_\_\_. **Um estudo gramatical da língua Dzubukuá, família Karirí**. Tese de doutorado. Universidade Federal do Pernambuco, 20012.
- RAIOL, D. A. **Catechese de índios do Pará**. Belém: Annaes da Bibliotheca e Archivo Público do Pará, v.2, 1900.
- RAMOS PÉREZ, D. La crisis Indiana y la junta magna de 1568. **Jahrbuch für Geschichte Lateinamerikas/Anuario de historia de América Latina**, v. 23, p. 1-61, 1986.
- REICH, U. Mudança sintática e pragmática na Língua Geral Amazônica (LGA): marcação de caso e sistema pronominal. In: FREIRE; ROSA, p. 167-184, 2003.

- RODRIGUES, A. Dall'Igna. Tupi, tupinambá, línguas gerais e português do Brasil. In: NOLL; DIETRICH, p. 27-47, 2010a.
- \_\_\_\_\_. Some cases of regrammaticalization in Tupí-Guaraní languages. **Revista Brasileira de Linguística Antropológica**, vol. 2.2, p. 65-74, Dez. 2010b.
- \_\_\_\_\_. As línguas gerais sul-americanas. **Papia: Revista Brasileira de Estudos Crioulos e Similares** v. 4, p. 6-18, 1996.
- \_\_\_\_\_. **As línguas brasileiras. Para o conhecimento das línguas indígenas**. São Paulo, Loyola, 1986.
- \_\_\_\_\_. Relações internas na família tupi-guarani. **Revista de Antropologia** v. 27/28, p. 35-54, 1984/1985.
- \_\_\_\_\_. Contribuição para a etimologia dos brasileirismos. **Revista Portuguesa de Filologia**, vol. 9, p. 1-54, 1958/59.
- \_\_\_\_\_. Classification of Tupi-Guarani. **International Journal of American Linguistics**, v. 24, n. 3, p. 231-234, 1958.
- RODRIGUES, A. Dall'igna; CABRAL, A. S. Arruda Camara. Revendo a classificação interna da família Tupí-Guaraní. In: CABRAL; RODRIGUES, p. 327-337, 2002.
- RODRIGUES A. Dall'Igna; CABRAL, A. S. Arruda Camara. A contribution to the linguistic history of the Língua Geral Amazônica. **A.L.F.A**, 55.2, p. 613-639, 2010.
- ROSA, M. C. The 16th and 17th centuries: Tupi or Língua Geral? The grammars of Anchieta and Figueira. **I: Actes: La "découverte" des langues et des écritures d'Amérique**, 1990.
- \_\_\_\_\_. Descrições missionárias de língua geral nos séculos XVI-XVII: que descreveram? **Papia**, v. 1, p. 85-98, 1992.
- SANTOS, B. Machado 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 on 24/10/2023 at: <https://ptdocz.com/doc/424876/as-cartas-do-padre-manuel-gomes-e-suas-representa%C3%A7%C3%B5es>
- SARAGOÇA, L. **De Feliz Lusitânia aos Confins da Amazônia 1616-62**. São Paulo: Cosmos, 2000.
- SILVA NETO, S. da. **Introdução ao estudo da língua portuguesa no Brasil**. Rio de Janeiro: Presença, 1951.
- SOUSA, G. Soares de. **Tratado descritivo do Brasil em 1587**. São Paulo: Editora Hedra, 2010.
- SCHMIDT-RIESE, R. Condições da mudança em nheengatu: pragmática e contatos linguísticos. In FREIRE; ROSA, p. 147-166, 2003.
- TEYSSIER, P. **História da língua portuguesa**. 3. ed. São Paulo: Martins Fontes. Tradução: Celso Cunha, 2007.
- TEJEDO-HERRERO, F. Algunas reflexiones en torno al término *latín* en la documentación alfonsí. **Romance Quarterly** 56.1, p. 4-12, 2009.
- VIEIRA, M. M. Damaso. A Manifestação de Tópico e Foco em Línguas da Família Tupi-Guarani. **D.E.L.T.A.**, 30 especial, p. 659-683, 2014.
- WOOLARD, K. A.; GENOVESE, E. N. Strategic bivalency in Latin and Spanish in early modern Spain. **Language in society**, v. 36, n. 4, p. 487-509, 2007.

- WRIGHT, R. Speaking, reading and writing Late Latin and Early Romance. **Neophilologus** v. 60, p. 178-89, 1976.
- \_\_\_\_\_. **Late Latin and Early Romance in Spain and Carolingian France**. Liverpool: F. Cairns, 1982.
- \_\_\_\_\_. Versatility and vagueness in Early Medieval Spain. In: MACKENZIE; MICHAEL, p. 207-223, 1993.
- \_\_\_\_\_. (ed.) **Latin and the Romance Languages in the Early Middle Ages**. Reprint 1996: University Park, PA: Pennsylvania State University Press, 1996.
- \_\_\_\_\_. **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.
- \_\_\_\_\_. Reading a will in twelfth-century Salamanca. **Latin vulgaire – latin tardif**, 5, p. 505-516, 1999.
- \_\_\_\_\_. The assertion of Ibero-Romance', **Forum for Modern Language Studies**, v. 36, p. 231-240, 2000a.
- \_\_\_\_\_. **El Tratado de Cabrerros (1206): estudio sociofilológico de una reforma ortográfica**. London: Department of Hispanic Studies Queen Mary and Westfield College, 2000b.
- \_\_\_\_\_. **A Sociophilological Study of Late Latin**. Turnhout: Brepols, 2002.
- \_\_\_\_\_. Sociophilology and twelfth-century Spain. **Medioevo romanzo**. mag./ago., p. 1000-1018, 2003.
- \_\_\_\_\_. The Monolingual Latin Glossaries of the Iberian Peninsula: Can They Help the Romanist??. **Latin écrit-Roman oral**, 2008.
- ZAVALA, S. **El castellano, lengua obligatoria?** SEP, Coordinacion de publicaciones, Promoción Editorial y Bibliotecas, 1977.

This preprint was submitted under the following conditions:

- The authors declare that they are aware that they are solely responsible for the content of the preprint and that the deposit in SciELO Preprints does not mean any commitment on the part of SciELO, except its preservation and dissemination.
- The authors declare that the necessary Terms of Free and Informed Consent of participants or patients in the research were obtained and are described in the manuscript, when applicable.
- The authors declare that the preparation of the manuscript followed the ethical norms of scientific communication.
- The authors declare that the data, applications, and other content underlying the manuscript are referenced.
- The deposited manuscript is in PDF format.
- The authors declare that the research that originated the manuscript followed good ethical practices and that the necessary approvals from research ethics committees, when applicable, are described in the manuscript.
- The authors declare that once a manuscript is posted on the SciELO Preprints server, it can only be taken down on request to the SciELO Preprints server Editorial Secretariat, who will post a retraction notice in its place.
- The authors agree that the approved manuscript will be made available under a [Creative Commons CC-BY](#) license.
- The submitting author declares that the contributions of all authors and conflict of interest statement are included explicitly and in specific sections of the manuscript.
- The authors declare that the manuscript was not deposited and/or previously made available on another preprint server or published by a journal.
- If the manuscript is being reviewed or being prepared for publishing but not yet published by a journal, the authors declare that they have received authorization from the journal to make this deposit.
- The submitting author declares that all authors of the manuscript agree with the submission to SciELO Preprints.