

Cilene RODRIGUES

(PUC-Rio - Pontifical Catholic University of Rio de Janeiro, Brazil)
crodrigues@puc-rio.br

Migration within Tupi-Guarani Languages¹

Abstract

Around 2,000 years ago, the Guarani languages, a subgroup of the Tupi-Guarani family (Tupi Stock), migrated from the Amazon to the lowlands of South America forming the modern Guarani languages, which might be classified as pluricentric languages. The present paper aims at presenting an overview of some phonemic and grammatical aspects of these languages, focusing on possible effects of their expansion southwards. First, it is shown that Guarani languages display a founder effect, a reduction in phonemic variability. Second, it is hypothesized that the possible migratory routes explored by Xetá and Aché might have resulted in language contact and depopulation, causing internal grammatical changes. Third, intensive language contact with positive language attitude and practices might have favored Paraguayan Guarani, which has the biggest number of speakers and the largest phonemic inventory among Guarani languages.

1. Introduction

Formation of pluricentric languages can involve long distance migrations. A group of individuals from population A spreads across the territory carrying A's language, but reaching unconnected distances with respect to A. Also, speakers can cross international borders, thus being subject to different political regulations that can affect their language. With that in mind, in this paper, we assess some historical migrations within Tupi-Guarani (Tupi stock), and their grammatical consequences, focusing on varieties of Guarani, which are spoken in the lowlands of South America. The following topics will be addressed:

- (a) Consequences of migrations to languages
- (b) How language contact and depopulation can affect founder languages
- (c) Influences of language attitudes and practices on the grammar of founder languages

¹ The study was financially supported by CNPq – Brazilian National Council for Scientific and Technological Development (Grant # 311093/2019-1).

An often-ignored treasure of South America is linguistic diversity, and little is known about the history of South American native languages. Thus, most of them are out of the scope of current research on grammar and its relationship with historical issues, such as migrations, depopulation and language contact. In the present paper, however, we emphasize that studies on these languages can bring important insights to our understanding of language as a biosocial phenomenon.

South America houses one fourth of the world's languages (Campbell & Grondona 2012), the majority of which are spoken by minorities, communities with less than 1 million speakers. According to Lewis (*Ethnologue: languages of the world* 2009), there are currently 624 languages in South America, but 179 of them are dying. Brazil is the country with the largest number of native languages within South America, but Brazilian native languages are very small (Rodrigues 2013), and Brazil is listed as the third country in terms of language loss (Lewis 2009). Therefore, we need to draw attention to the importance of these languages.

The paper is organized as follows: section 2 presents the history of Tupi, focusing on the phonemic consequences of migration within Guarani. In section 3, we discuss how certain grammatical features of two modern Guarani languages, Aché and Xetá, might reflect issues related to language contact and depopulation during migration and settlement in the south. Section 4 is devoted to Paraguayan Guarani, a discussion on how, under intense contact with Spanish, Paraguayan Guarani's grammatical system was preserved, arguably due to positive local social attitudes and language practices. Our main conclusions are in section 5.

2. Founder effects: a consequence of migration

Within genetics, a *founder effect* is defined as a case of genetic drift related to territorial expansions (Mayr 1954, 1963), when a small subgroup of individuals (the founder) gets territorially disconnected from its origins. Carrying just a small fraction of the total genetic variants of the original population, it may experience a founder effect, a reduction in its genetic pool, with a cline in heterozygosity and allelic richness.

Examples of founder effects are abundant. For example, Marques & Renesto (2017) report that, due to alterations in their habitat (introduction of dams), *Hemiodus orthonops*, a small fish endemic from the Paraná-Paraguay basin river, migrated to the upper part of the Paraná River, and genetic analyses of the

genotype of two founder populations indicate that they display a low mean of alleles per locus and an excess of homozygotes. Another example is that the rate of blood-type “O” among native Americans is 98%, but these natives descend from Asians, among whom a balanced 50% rate of “O” type is attested. Thus, Cavalli-Sforza (2001) hypothesizes that the founders who crossed over the Bering strait towards North America were exclusive carriers of the “O” allele.

Arguably, the most stunning and impactful human geographic expansion was the exodus from Africa in the lower Paleolithic period (~ 60,000 to 100,000 B.P.). This territorial expansion gradually resulted in the occupation of all continents, with Oceania and the Americas being the last ones to be occupied. This expansion went hand in hand with a subsequent serial founder effect: the further a modern human population is from Africa the poorer it is in genetic diversity. Sub-Saharan native African populations have an exceptional number of genetic variants, whereas native populations of the Americas and Oceania display a significant reduction (Cavalli-Sforza 2001, Ramachandran et al. 2005), as shown in figure 1.

| | | |
|--------|------|---------|
| Africa | 16.6 | Europe |
| | 20.6 | Asia |
| | 22.6 | America |
| | 24.7 | Oceania |

Fig. 1: Genetic distance between Africa and other continents, based on Cavalli-Sforza (2001:52)

Another impact of this great migration is on language. The phonemic inventories of non-African modern languages display a founder effect.

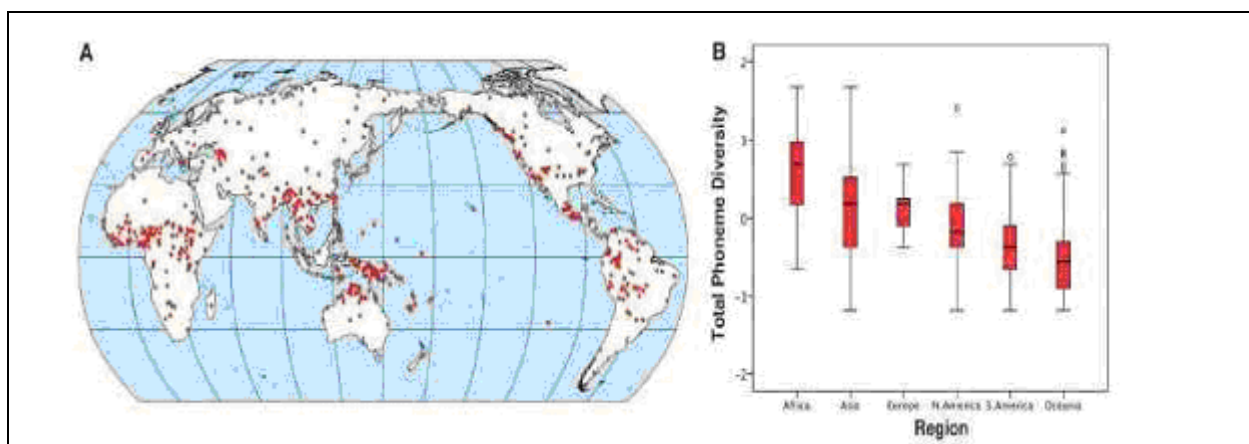


Fig. 2: A: location of analyzed languages, B: Phonemic diversity plotted against region (Atkinson 2011: 333).

Based on data from WALS - *The world Atlas of Languages Structures* (Dryer and Haspelmath 2005), Atkinson (2011) conducted a statistical analysis on a

sample of 504 modern languages from different families and continents. A correlation between phonemic reduction and physical distance from Africa was found, the further a language is from eastern Africa, the smaller its phoneme inventory is. (See Fig. 2).

Aktinson's findings correspond to some other important facts about phoneme inventory size. First, languages with too many neighbors present less phonemic variability (Creanza et al. 2015), but languages spoken by large populations have more phonemes than languages spoken by small populations (Trudgill 2005, Hay & Bauer 2007). Also, languages spoken by large populations increase their phonemic inventory over time at a rate between 0.26 and 0.38 per 1,000-year period (Perreault & Mathew 2012). Importantly, however, sonorant consonants and monophthongs seem to be less affected by this type of fluctuation. Spearman correlations between the means of language family, population and phonemic inventory size, were not significant for sonorant consonants and monophthongs (Hay & Bauer 2007:394)

2.1. Tupian languages: birthplace and migrations

Evidence from archeology, linguistics and genetics indicates that Proto-Tupi emerged around 5,000 years ago in the central-western part of the Amazon, in the Madeira-Guaporé Region (MGR), an area bounded by the rivers Amazon, Tocantins, Madeira and Guaporé, arguably where the current Brazilian state of Rondônia is located (Rodrigues 1964, Brochado 1984, Noelli 1998).

Around 3,000 years ago, Tupi started branching into different languages (Rodrigues 1964, Urban 1992, 1996) which corresponds to 10 modern language families (Rodrigues 1986, Rodrigues & Dietrich 1997). Currently, five of these language families are still located in MGR: *Arikém*, *Mondé*, *Ramaráma*, *Tuparí*, *Puruborá*, while others are located either to the north or to the east of MGR (*Awetí*, *Juruna*, *Mauwé*, *Munduruku*), as shown in figure 3. Tupi-Guarani is an exception.

Fig. 3: Approximate location of Tupi-Languages: I *Arikém* (a. Karitiana, b. †Arikén), II *Aweté* (Awetí), III **Juruna** (a. Juruna, b. Xipaya, c. † Minitawá), IV *Mawé* (Mawé), V *Mondé* (a. Mondé, Aruá, c. Gavião, d. Suruí, e. Cinta-Larga, f. Zoró), VI *Mundurukú* (a. Mundurukú, b. Kuruáya), VII *Puruborá* (Puruborá), VIII *Ramaráma* (a. Karo, b. Urumi), IX *Tuparí* (a. Tuparí, b. Wayoró, c. Menkéns, d. Makuráp, e. Sakirabiát, f. Kepkiriwát). (Rodrigues, 2007: 108)

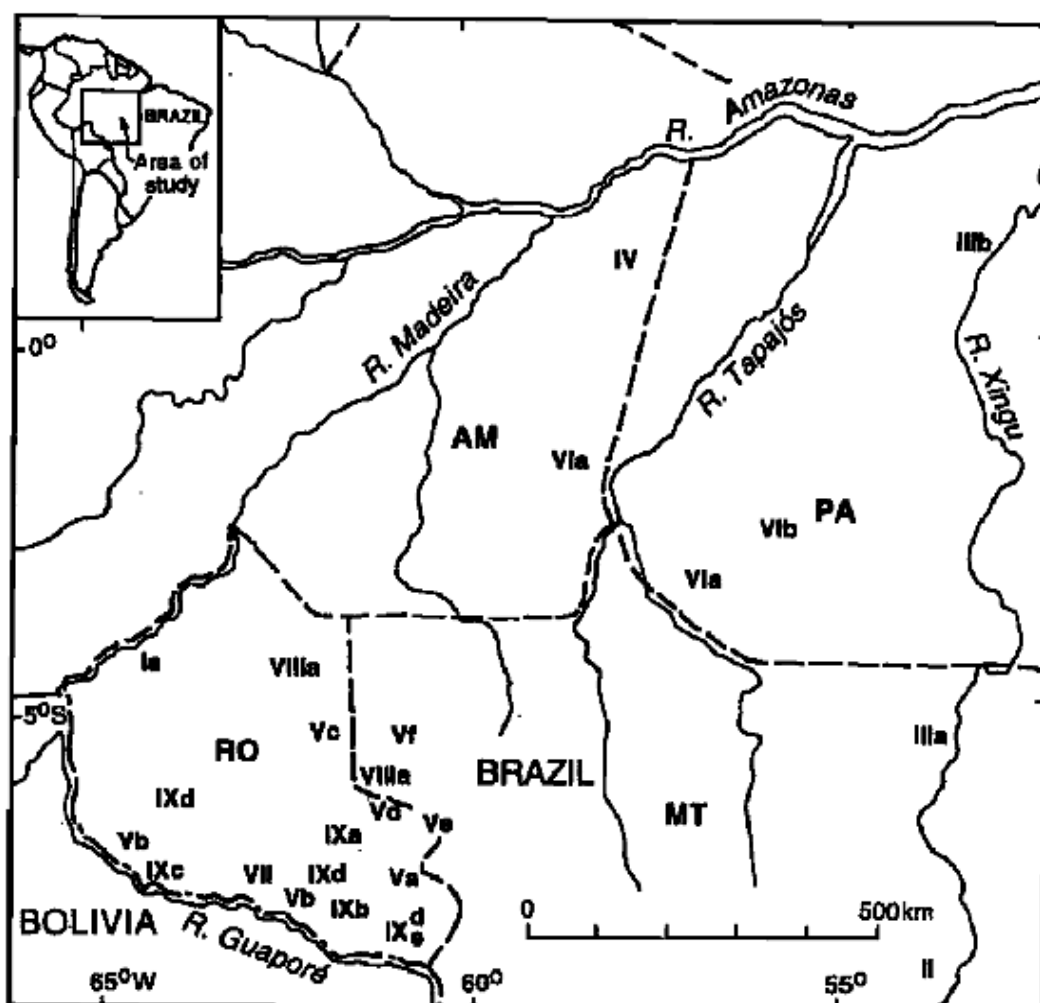


Fig. 3 Tupi languages with approximate locations

Tupi-Guarani (TG) is the biggest language family within the Tupi stock. It contains approximately 40 languages (Rodrigues 1985) which can be divided into 8 subgroups based on grammatical similarities (Dietrich 1990, Rodrigues & Cabral 2002, Michael et al. 2015). TG languages are spread out within the South American territory and are currently spoken in Brazil, Bolivia, Argentina, Paraguay, Peru and French Guiana. It is hypothesized that right after the formation of Proto-TG, its speakers fanned out in a radial fashion across South America (figure 4). The reason for this dispersal is unknown (see Noeli (2008) for a hypothesis based on demographic growth), but it was fast and took TG languages far from MGR. They expanded northward (Emerillon- Wayampí) and southward through routes in the east (Tupinambá) and in the west (Guarani). Around 2,000 B.P., some TG people (pre-Guaranis) were already in the lowlands of South America, occupying fertile fluvial zones in the Paraná-Paraguay River

basin, on lands that are now part of five countries, Bolivia, Brazil, Argentina, Paraguay and Uruguay (Mello 2000, Rodrigues 2007).

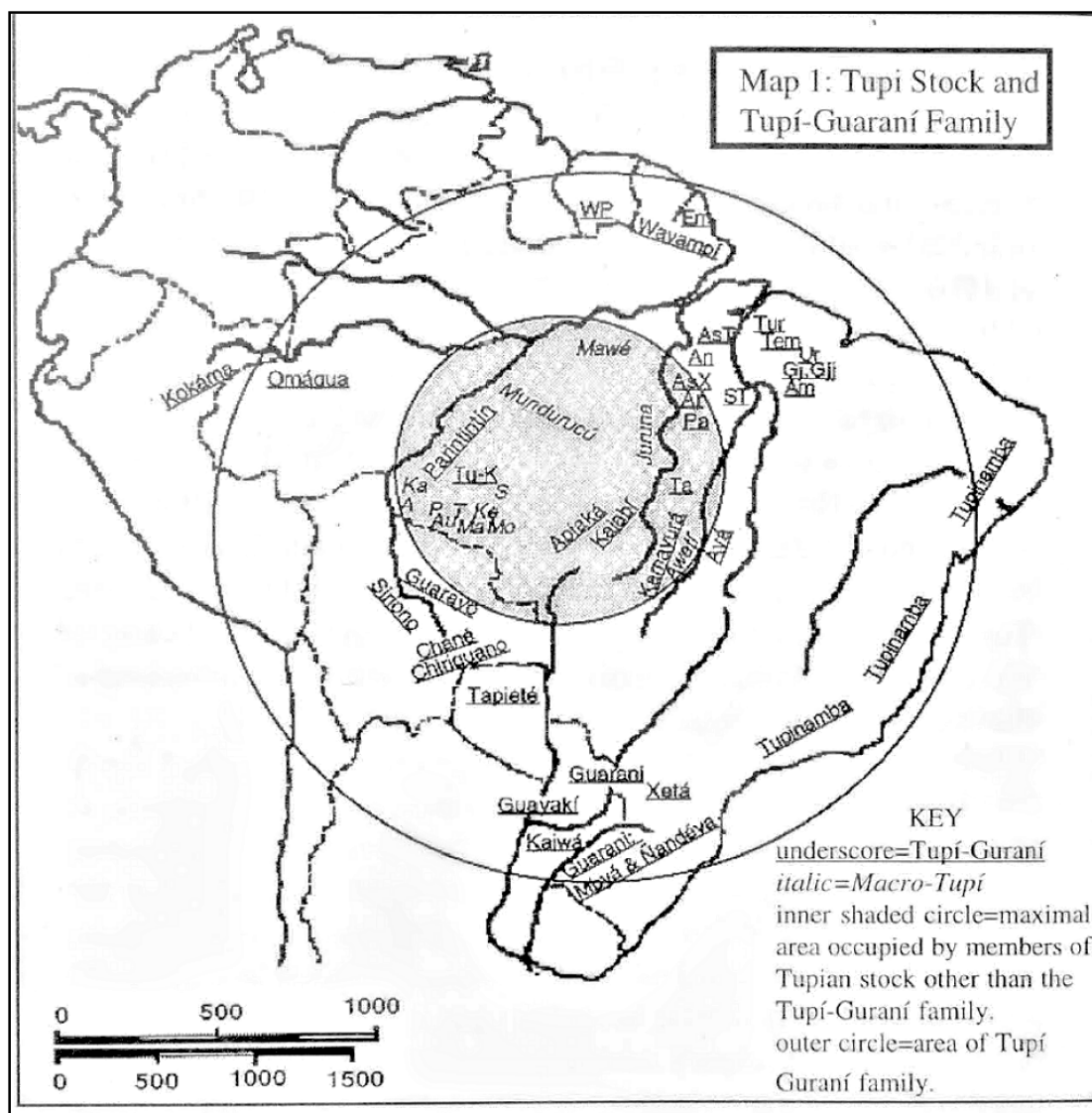
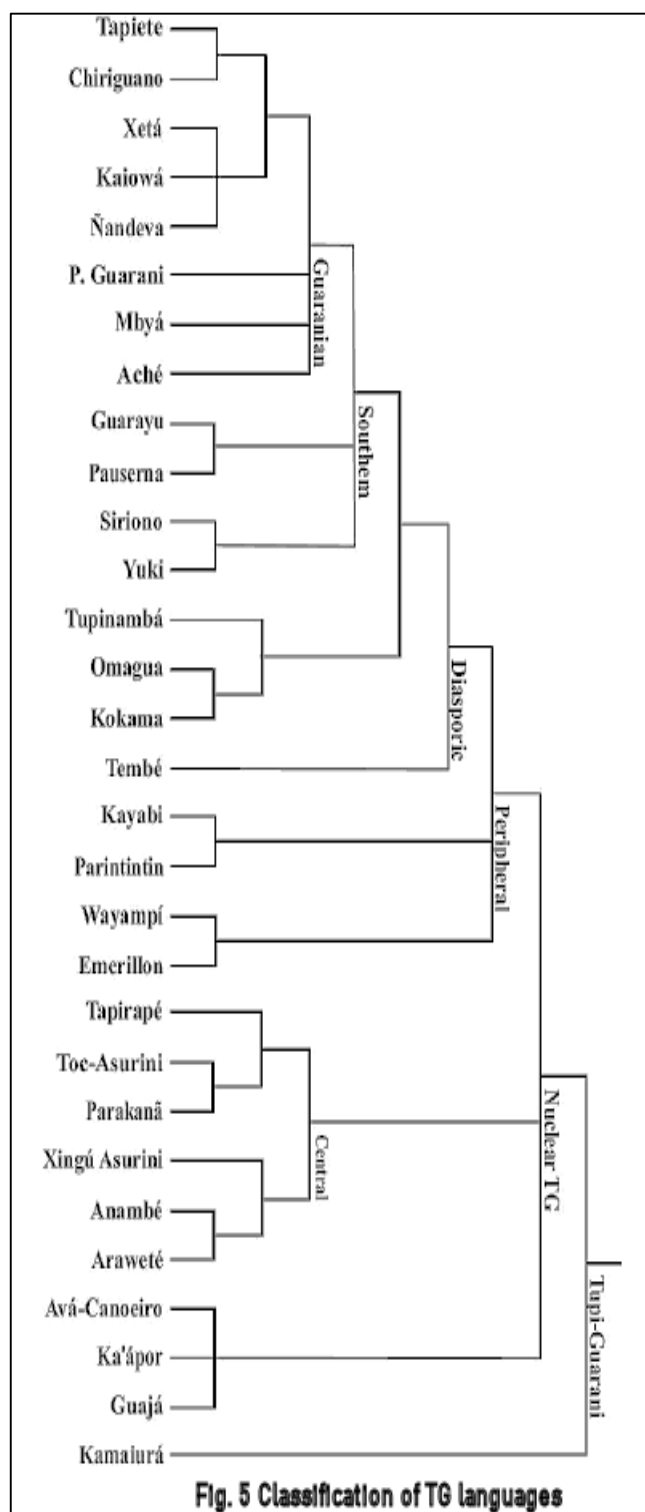


Fig. 4: Location of modern TG languages (Urban 1996:66)²

² Abbreviations for Tupi: Stock Languages: A=Arara; Au=Aruá; Ka=Kepkiriwat; Ma=Makurap; Mo=Mondé, P=Puruboná, S=Surui; T=Tuparí; Abbreviations for Tupi-Guaraní Family/Languages/dialects: Am=Amanajé; An=Anambé; Ar=Araweté; AsT=Assurini do Tocantins; AsX=Assurini do Xingú; Em=Einérillon; Gj=Guajá; Gjj=Guajára ST=Suruí do Tocantins. Not listed on map: Horá=southernmost dialect of Sirionó and adjacent to it. Kokamáya=almost identical to Kokama and adjacent to it on map.



Guarani languages correspond to Group I in Rodrigues' (1985) classification, according to whom are variants, dialects of one single language. Rodrigues' grouping is confirmed by recent studies such as Michael et al. (2015) (see figure 5), where the Guarani branch is a monophyletic group within southern TG.³

The term pluricentric languages (Clyne 1992) may apply to Guarani languages. There are plenty of linguistic interactions among their speakers and the term Guarani provides a sense of belonging and identity to all Guaranis⁴ while to its speakers, each Guarani language represents a separated, distinct language, thus playing an important role in community cohesion and identity. Particularly, Guarani languages might be classified as migrant PLC languages, type 10 in Muhr's (2016) types of PLC languages. But a proviso is in order: there is no dominant relationship or shared standardized form among Guarani varieties.⁵ Also, there is no synchronic center of dispersal among them.

³ Nowadays, due to recent migrations, Guarani languages are currently occupying areas in the South (Rio Grande do Sul, Santa Catarina e Paraná), Southeast (São Paulo e Rio de Janeiro, Espírito Santo), and Midwest (Mato Grosso do Sul) of Brazil.

⁴ Also known as Carijós in some Portuguese sources.

⁵ In the 16th-18th century period, the Jesuit reductions (or missions) used a variety of Guarani as the *lingua franca* of the south. This variety was considered by non-natives to be the standard variety.

2.2. Founder effects within Tupi

A genetic study conducted by Santos et al. (2015) indicates that among Tupi people, populations located outside MGR present less genetic variation than populations within MGR. Tupian populations were screened for autosomal and uniparental (Y- chromosomes and MtDNA) markers. The following Tupian populations were considered: *MGR: Zoró, Gavião, Suruí, Cinta Larga, Karitiana; Outside MGR (NonMGR): Mundurukú, Asuriní do Trocará, Asuriní do Coatinemo, Araweté, Urubú-Ka'apor, Parakanã, Awá-Guajá, Wayampi, Zoé, Aché and Guarani.* Alongside a decay in genetic variability in NonMGR (figure 6), results also suggest recent depopulation in both MGR and NonMGR, which might be a result of the colonization process, but in ancient times MGR experienced a population growth while NonMGR experienced demographic bottlenecks.

| Markers | MGR | No-MGR |
|-----------------|-------|--------|
| Classics1 | 0.193 | 0.172 |
| STR2 | 0.683 | 0.632 |
| Y-STR3 | 0.253 | 0.164 |
| Alu insertions4 | 0.235 | 0.213 |
| mtDNA sequence5 | 0,865 | 0.845 |

Fig. 6. Variation in genetic markers in Tupi populations (Santos et al. 2015: 4)

Parallel results are presented by Walker et al. (2012), which focused on loss of cultural traits. Based on a linguistic phylogenetic analysis, the authors concluded that NonMGR populations lost some cultural traits present in Tupi, such as uxori-locality (transference of males to live with in-laws after marriage), aggressive warfare, corporate structure (e.g., clans and lineages and moieties), canoe, tattooing, shamanism and lip plugs. This loss of cultural traits was particularly strong in four TG groups, Guajá (northeast of Brazil), Sirinó (Southern - Bolivia), Xetá (Guarani - south of Brazil) and Aché (Guarani - Paraguay).

Rodrigues (2020) investigated linguistic founder effects within Tupi, considering the phonemic inventories of two groups of languages: Southern TG languages (SouthTG) and Tupi within MGR (MGR). Her samples contained the following languages: *MGR: Karitiana (Arikém), Gaivão (Mondé), Suruí Paitér (Mondé), Puruborá (Poruborá), Káro (Ramaráma), Makuráp (Tuparí), Akuntsú (Tuparí), Tuparí (Tuapri), Sakirabiát (Tuparí), Wayoó (Tuparí); SouthTG: Kayowá (Guarani), Nhadeva (Guarani), Xetá (Guarani), Guarani Mbyá (Guarani), Chiriaguano Izoceño (Guarani),*

Chiriguano Chané (Guarani), Guarayu (Southern), Sirinó (Southern), Paraguayan Guarani (Guarani) and Tapieté (Guarani).

Between-group statistical analyses were conducted, taking into consideration the following factors: number of consonants and number of vowels (oral & nasal) and (b) population size. T-tests and variance tests were not significant for consonants, although a tendency towards a smaller consonant inventory was observed in SouthTG. A significant difference was found for vowel inventory, with MGR languages presenting, among themselves, more variance in vowel inventory size than SouthTG languages (Nasals: $P < 0.00001$, Oral: $P < 0.00001$). No effect on population size was found, which is expected because Tupi languages have a small number of speakers, except for Paraguayan Guarani.

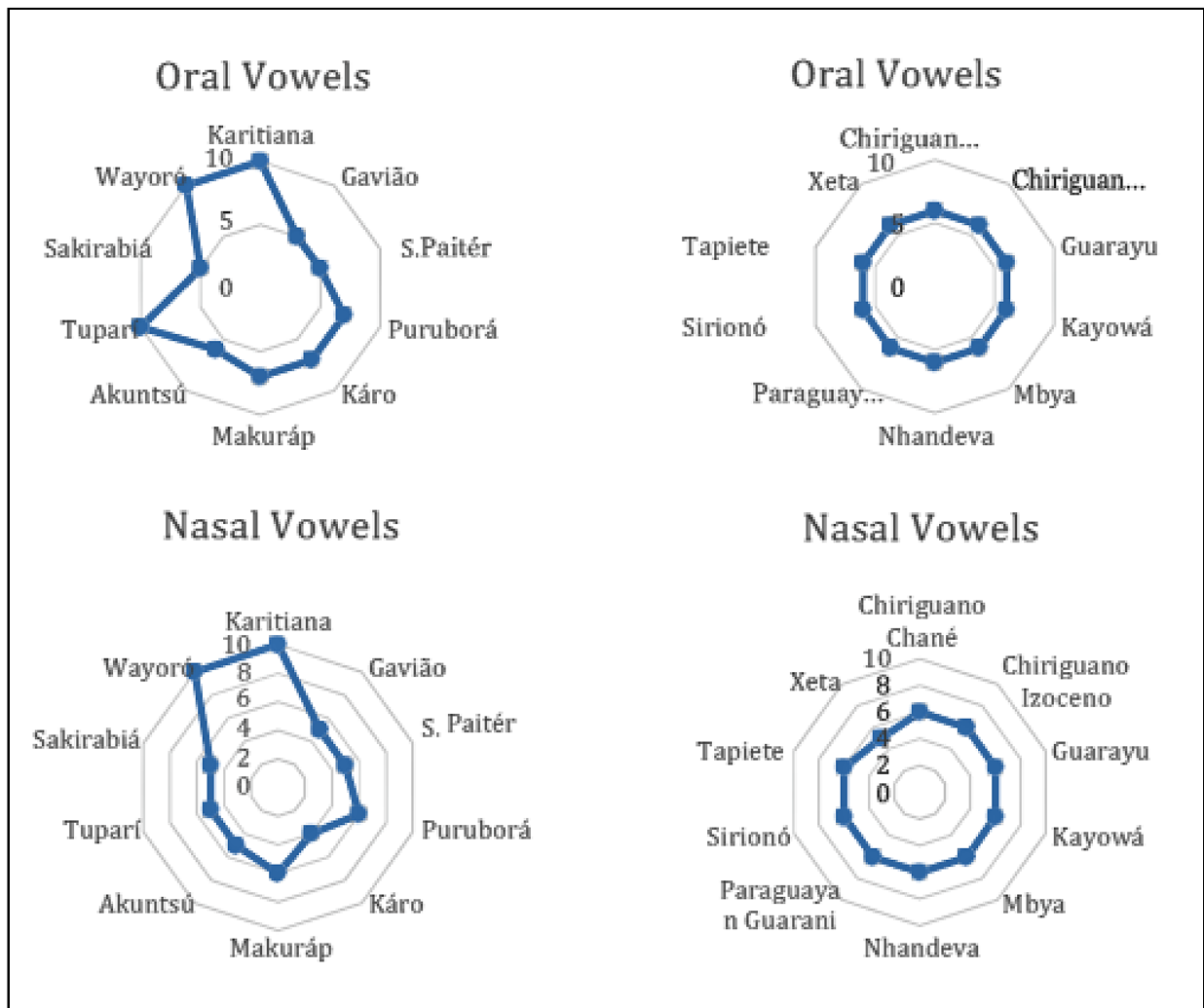


Figure 7. Variation in Vowel inventory size among MGR and SouthTG languages (Rodrigues 2020: 88).

Rodrigues' results indicate no inter-language variation in vocalic phonemes among Guarani languages. All Guarani languages considered in the

study, except for Xetá, present a harmonious 2x6 vowel system (6 oral, 6 nasal). Contrastively, MGR languages exhibit variance among themselves. This therefore suggests a phonemic founder effect in Guarani, with significant reduction in vowel variance.

This conclusion seems to challenge Hay & Bauer's (2007) observation that vowels are more resistant to fluctuations in population size. This contradiction is superficial, however. All reconstruction analyses of Proto-TG suggest that it had the same harmonious 2x6 vowel system observed in Guarani (Lemle 1971, Rodrigues & Dietrich 1997, Schleicher 1998, Meira & Drude 2015). If so, Guarani is highly conservative with respect to vowels, preserving the system inherited from Proto-TG. MGR languages on the other hand, seem to be less conservative. Languages of the Tuparí family for instance, present great variability among themselves (Rodrigues, 2021).

Following Santos et al.'s (2015) conclusion that NonMGR languages suffered a doubled depopulation process before and after Portuguese colonization, in accordance with Hay & Baur, Rodrigues (2020) suggests that the Guarani vowel system was not affected by demographic reduction (but see section 3). In addition, as opposed to Amazonian Tupi, given that most South TG is in constant contact with other languages (e.g., languages of the colonizers (Portuguese and Spanish) and Jê languages), Rodrigues' findings align with Creanza et al.'s (2015) conclusion that languages with too many neighbors present less phonemic variability (we will return to this matter in section 3). It is also compatible with Trudgill (2005) and Hay & Bauer's (2007) observation that small populations have fewer phonemes.

This discussion raises important issues about the role played by vowels and consonants in human language. Why should vowels be more resilient than consonants to external causers of phonemic changes? Investigations on language acquisition point towards a division of labor between vowels and consonants. While consonants are used to build the lexicon, identifying lexical items and highlighting differences among them, vowels signal structural relations and prosodic boundaries. Vowels, not consonants, indicate syllabic structure, boundaries between syntactic constituent and word order, and morphosyntactic process (Nespor et al. 2003, Hochmann et al. 2011, Nespor et al. 2011, among others). This arguably provides a rationale for why languages, with very few exceptions, have more consonants than vowels, the most common being a 4 (20C/5V) ratio (Maddieson & Precoda 2008). In general, vowels are used as structural signalers in Tupi languages. In Guarani Mbyá for example, the

nominalizer morpheme is expressed by -a, as shown by examples in (1), and vowels are used as agreement markers. As shown in fig. 8, contrasts in person agreement are realized by vowels.

(1) Nominalization in Guaraní Mbyá (Martins 2003: 123)

ayvu -jopy -a
speech-grab-NMLZer
 ‘recorder’

Karu-a
 To-eat-NMLZer
 ‘place to eat’

mba’e-xo-a
 thing-pound-NMLZer
 ‘plastle’

| Person | Active Pronouns | Inactive Pronouns |
|------------------|-----------------|-------------------|
| 1PSg | a- | Xe |
| 1PPI (Inclusive) | ja- | nhande |
| 1PPI (exclusive) | oro- (ro) | Ore |
| 2PSg | ere – (re) | nde (ne) |
| 2PPI | pe- | pende (pene) |
| 3PSg/Pl | o- | |

Fig. 8. Person agreement markers in Guaraní Mbyá, adapted from Martins (2003: 36)

The fact that Guaraní languages share the same vowel system, thus allow us to form the hypothesis that these languages are grammatically similar, presenting few parametric variations among them. This hypothesis is aligned with Rodrigues’ (1985) observation that Guaraní languages are dialects of a single language.

3. The effects of language contact on founder languages

An important issue in understanding ancient founder languages is the interaction between ecological factors, population size, language contact and migration. The underlying question is whether human migrations preserve familiar habitats or not. This is an important migratory topic as it converses with questions related to language contact, resulting language changes and linguistic diversity. In this section, we will discuss this issue, focusing on two Guaraní languages, Aché and Xetá. In the next section, we will come back to language contact in Paraguayan Guaraní.

Based on phylogenetic data from 400 modern Bantu languages, Grollemund et al. (2015) concluded that during their territorial expansion within the African territory (~5.000 B.P.), populations of Bantu speakers explored savannah corridors, avoiding rainforests. Populations that moved into rainforests presented a slow migration rate, delaying the occupation of the new area. As O'Hagan (2014) puts it, "remaining within ecologically similar zones allows for continuity in subsistence practices during migrations".

Hence, ecological factors might interact with the formation of founder languages in that, migrating groups tend to explore routes that are ecologically similar to their original habitat, and this can shape their languages in different ways. For example, different routes may favor or impede language contact, one of the external driving forces behind language change. It can also result in demographic bottlenecks that might impair language transmission.

The routes taken southwards by pre-Guaranis are rather unclear and it is arguably the case that their territorial expansions occurred in three temporally different migratory waves. First, in Wave 1, Aché went down to Paraguay and Sirinó to Bolivia.⁶ In wave 2, Xetá went to Serra dos Dourados in Paraná/Brazil. Wave 3 consisted of Chiriguano moving to Bolivia, Tapiete and Guaraní to Paraguay, and Kaiowa into the region between Argentina, Brazil and Paraguay (Urban 1992, Noelli 2008). Despite this, given the Tupian preference for headwater regions, it is very likely that all pre-Guaranis explored ravines during migration and followed water courses (O'Hagan 2014, O'Hagan et al. 2019).

The location of the center of dispersal of TG languages is also very obscure. Urban (1992, 1996) argues that it is somewhere between the rivers Madeira and Xingu. Recently, using Linguistic Migration Theory (Sapir 1916, Dyen 1956), and based on Michael et al.'s (2015) classification of TG languages, O'Hagan (2014) and O'Hagan et al. (2019) concluded that Proto-TG emerged in the lower part of the Xingu River. First, a branch of Tupi (Proto-Mundurucic-Maweti-TG) migrated to northeast Amazon. Then, in the lower Tapajós River, Proto-Mawetí-TG split off, migrating latter to lower Xingu, where Proto-TG emerged (Fig. 10).

From lower Xingu, pre-Guaranis might have followed different routes southwards (figure 10) (O'Hagan 2014). They might have followed Tocantins and Araguaia rivers towards Paraná River, which would be the shortest route, but, moving west and then south, they could have explored the watercourses of the Tapajós River or the Madeira River.

⁶ Aché is spoken in Paraguay. It was not included in Rodrigues' (2021) sample of SouthTG languages.

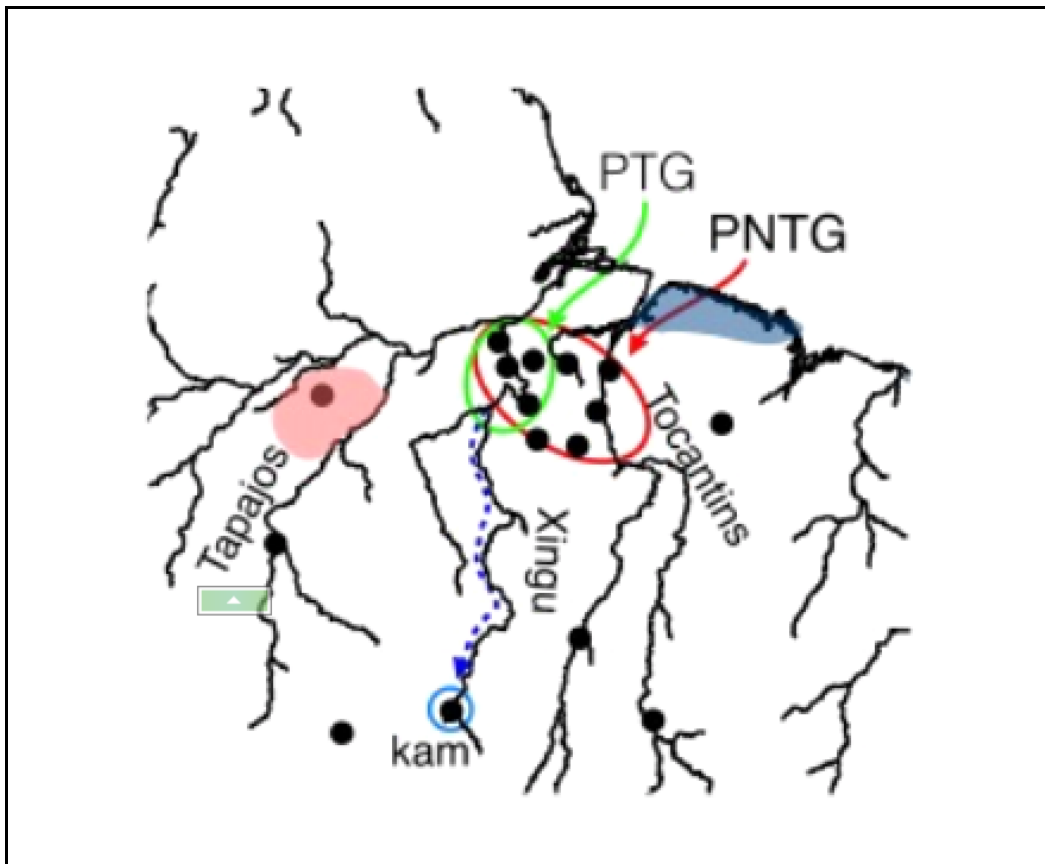


Fig 9. Possible Proto-TG (PTG) birthplace. (O'Hagan et al. 2019: 9)

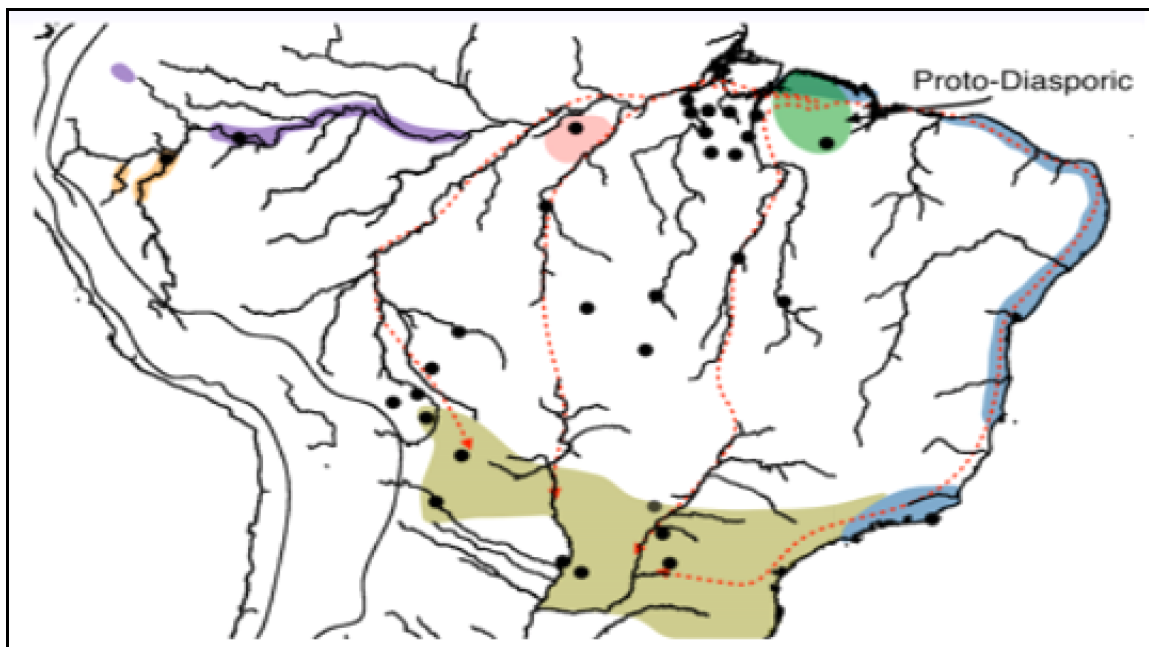


Fig. 10. Possible southward routes taken by pre-Guaranis. (O'Hagan 2014).

Given our current knowledge of this territorial expansion, it is not possible to pinpoint the chosen routes with certainty, although some hypotheses can be raised. One such hypothesis is related to the fact that Proto-Guarani

languages migrated in waves. This is particularly revealing for Xetá and Aché.

Modern Xetá and Aché differ from other Guaraní languages in lexicon, phonology and morphosyntax. According to Rodrigues (1978), taboo nouns and simple nouns denoting mythic factors from Guaraní were substituted in Xetá by nominal phrases. The stress pattern observed in Guaraní, where main stress falls mainly in the last syllable, was altered, being placed in the penultimate syllables. Cabral et al. (2005) observe that this shift in main stress caused weakening and reduction of the last syllables. As a result, modal suffixes disappeared in Xetá, affecting the person agreement system, which displays different forms for indicatives from class II and gerunds in Guaraní, but not in Xetá, where the gerund verbal form is morphologically identical to indicatives from class I. Importantly, Xetá has a reduced vowel inventory. It only has 5 nasal vowels (Rodrigues 2021).

Aché also differs from Guaraní. Röβler (2008) presents a detailed description of Aché, showing that it has around 11 consonants and a 6x2 vowel system, but free variation between [e] and [ɛ] and [o] and [ɔ] (oral and nasal). Aché also permits the use of complex syllabic onsets. In contrast with other Guaraní languages, the main stress has shifted to the penultimate syllable, and long nasal harmony, a feature of Guaraní languages, does not occur. In addition, all prefixes found in Guaraní languages (and in TG in general) disappeared in Aché, including those expressing person and number agreement. Changes are also observed in aspect and temporal markers (Röβler 2008).

Cultural aspects of these groups are also very telling with respect to their migratory history. TG people are known for cultivating their land. In contrast Xetá and Aché seem to have adopted a hunter-gather lifestyle (Hill & Hurtado 1996, Röβler 2008, Rodrigues 1978). Walker et al. (2012) also observed severe loss of other Tupi cultural traits in Xetá and Aché (section 2.2. above). Hence, despite their TG origins, Aché and Xetá exhibit some cultural dissociations from TG.

All in all, our question is: grammatically speaking, why did Xetá and Aché stray away from other Guaraní languages? The facts presented above suggest that the grammatical changes observed are correlated with changes in the vowel systems. This is a possibility that can be formally explored, but there must still be a diachronic event (or a series of events) that caused the observable phonemic changes and consequent grammatical changes.

It has been pointed out that both Aché and Xetá had intense contact with non-Tupi languages spoken in the south, particularly Jê languages, sometimes being mistaken by Jê groups secluded in the forest (Hill & Hurtado 1996, Röβler

2008, Rodrigues 1978). Thus, one hypothesis to be considered (see Röbler 2008), is that language contact is responsible for the changes observed, and it is associated with the fact that Xetá and Aché migrated first (waves I and II).

The migratory routes chosen by Xetá and Aché might have favored language contacts with Jê languages. Also, their route choices might have caused severe demographic bottlenecks forcing them to maintain contact with Jê groups in order to survive. Aché and Xetá are very small languages (see Hill & Hurtado (1996) and Röbler (2008) about Aché, and Rodrigues 2013 for Xetá).

A genetic study conducted on the Aché gene pool indicates descentance from Tupi with considerable transference of genetic material from Jê (Callegari-Jacques et al. 2008).

Although the observations above are mostly based on hypothetical scenarios, they are aligned with what we know about the history of TG groups and their languages.

4. The effects of language attitudes and practices on founder languages

Another important theme on the study of founder languages is related to their social and political use. Take Paraguayan Guarani (PG) as an example. Language contact added to the social and political status of PG set it apart from other Guarani varieties.

In distinct contrast with other Guarani languages, PG has around 5,000,000 speakers. It is the largest native language of South America, with emphasis on the fact that it is widely spoken by non-natives. 46.3% of Paraguayan homes use both Guarani and Spanish and 34%, only Guarani, with bilingualism concentrated in urban centers. Monolinguals (Guarani only) represent most of the rural population (Estigarribia 2020). Hence, PG is the only Guarani language that is not endangered.

However, there is no uniform, single grammatical system that can be called PG. The term refers to a set of linguistic varieties with different degrees of mixture with Spanish. Speakers use one of these varieties depending on their knowledge of, and fluency of Guarani. Jopara is the general term locals use to refer to mixtures between Guarani and Spanish, but according to Estigarribia, locals differentiate these varieties as follows: *Guaraniete*: “true” guarani, *Guarañol*: Jopara mostly based on Guarani and *Castení*: Jopara mostly based on Spanish. As examples of these varieties, consider the data below, extracted from

Estigarribia (2020:17).⁷

- (1) *Guaraniete*
 Chepytyvõmína
 che-pytyvõ-mi-na
1sg.inact-help-plead-req
 ‘Help me, please.’
Guarañol
 Cheayudamína
 che-ayuda-mi-na
1sg.inact-help-plead-req
 ‘Help me, please.’
Castení
 Ayudámena
 ayuda=me-na
help=me-req
 ‘Help me, please.’

The lexicon of these varieties contains words from Spanish and Guarani, and the amount of Spanish-based lexical influence depends on the speaker’s competence in Guarani.

With regard to phonemes, PG has the largest inventory among the Guarani languages. While there is a tendency in Guarani, as compared to MGR languages, towards a reduced consonant inventory, PG exhibits the opposite pattern. It has the Guarani vowel system (6x2) and 18 consonants. The mean size of the consonant inventory within Guarani is 14.11. Hence, the size of the consonant inventory of PG is well above the mean size of Guarani.

Reconstructions of Proto-TG have posited different numbers of consonants: 12 in Lemle (1971), 13 in Schleicher (1998) and 19 in Rodrigues and Dietrich (1997). Therefore, either PG gained consonants, or maintained the consonants it inherited from TG, whereas the other Guarani languages reduced theirs. Based on Perreault & Mathew’s (2012) conclusion that phonemes accumulate at a rate between 0.26 and 0.38 per 1,000-year period, and on Creanza et al.’s (2015) observation that languages in contact tend to be more conservative with respect to phonemic inventory size, we may conclude that PG is the most conservative Guarani language with respect to phonemes, preserving

⁷ See Estigarribia (2015), where it is argued that Jopara is a code-switching system.

the vowel and consonant inventories inherited from TG (see Rodrigues 2020). In accordance with this conclusion, Estigarribia (2020) shows that PG is phonologically faithful to Guarani, blocking complex onsets within syllables and preferring stress on the final syllable. Also, as to nasal harmony, PG follows the Guarani pattern (Estigarribia 2017, 2020). As for morphosyntax, although detailed comparative studies on Guarani languages are to be done, there seems to be no major morphosyntactic differences between PG and other Guarani Languages, with the exception of Xetá and Aché (section 3).

Given the intense contact between PG and Spanish, comparisons between PG and Aché and Xetá, lead to the conclusion that under language contact, a founder language can either preserve its grammar or have it altered. Thus, the question is why language contact with Spanish did not alter the grammar of PG, which I take to be well represented by the *Guaraniete* variety.

There are two relevant facts here. First, it is probably the case that Xetá and Aché suffered severe depopulations during migration to and settlement in the south, as discussed in section 3. Although the colonization of Paraguay caused an important dissemination of natives, the so-called Paraguayan Guaranis did not experience a depopulation as severe as Aché and Xetá did.⁸ Thus, while intergenerational language transmission might have been impaired in Xetá and Aché, it was not in PG. Second, more recently, positive language policies established by the Paraguayan government and positive national attitudes towards Guarani, were important factors in preserving the language of the Guarani founders in Paraguay.

In 1992, PG was declared an official language of Paraguay and education on Guarani became mandatory. In 2019, it was also listed as an official language of MERCOSUL. This formal status is not observed in other Guarani varieties, even though many speech communities of Guarani present a strong ethno-linguistic and cultural awareness. Hence, although PG doesn't have the same prestige that Spanish has in Paraguay, it is taken to be part of the Paraguayan ethno-identity. The existence of a language mixture, Jopará, is an indication of a positive local attitude and practice towards PG.

Bilingualism is very common among speakers of other varieties of Guarani. For instance, most Brazilian Guaranis are fluent speakers of Brazilian Portuguese as well, but no systematic language mixture with Brazilian Portuguese is

8 According to Röβler (2015), nowadays, Aché has only 250 fluent speakers. Xetá has only 3 speakers (Rodrigues 2013). On the history of the Xetá group of Serra dos Dourados, see the documentary *Xetá*, available at https://www.youtube.com/watch?v=_aUyib-tAlo (last accessed on 02/10/2022).

observed. This is largely due to the fact that the vast majority of non-native Brazilians have no desire to learn a native language.

5. Conclusion

In this paper, we tried to draw together some observations about migrant pluricentric languages, taking Guarani, a group of Tupi-Guarani languages spoken in the lowlands of South America, as an example. These languages exhibit a linguistic founder effect, showing less variation in vowel inventory than languages spoken in MGR, the birthplace of Proto-Tupi. However, given that Proto-Tupi-Guarani had the same vowel system that modern Guarani has, most Guarani languages are rather grammatically conservative. This relates to the role played by vowels in human grammar. Vowels are signalers of grammatical structure and prosodic boundaries. All Guarani languages that preserved the Tupi-Guarani vowel inventory are similar in phonology and morphosyntax. Xetá and Aché differ from Guarani, exhibiting changes in their vowel inventories and phonological and morphosyntactic alterations. We hypothesized that these languages result from language contact (arguably with Jê languages), and severe depopulation events during migration to, and settlement in the south. In contrast to this hypothesis, Paraguayan Guarani did not have its phonemic inventory and grammar altered, despite intense contact with Spanish. This might reflect the following facts: Paraguayan Guaranis did not experience a severe depopulation, positive language policies were established by the Paraguayan government, and among Paraguayans (natives and non-natives), there is a positive national attitude towards Guarani.

To conclude, investigations on migrant languages are an important field of research on language diversity, but in itself, it is an interdisciplinary field, involving knowledge of theoretical linguistics, language as a cognitive system, language typology, language contact, historical linguistics and knowledge about educational and regulatory language policies.

References

- Atkinson, Q. (2011): Phonemic Diversity supports a founder effect model of language expansion from Africa. IN: *Science*, 332. P. 346-349
- Brochado, J. (1984): An ecological model of the spread of pottery and agriculture into eastern South America. Ph.D. dissertation.
- Cabral, A. S. / Rodrigues, A. D. / Vasconcelos, E. A. (2005): Sobre o sistema pessoal da língua Xetá. In: L. Lobato et al. (eds.), *Anais IV Congresso internacional da ABRALIN*. Universidade de Brasília. P. 57-64.

- Callegari-Jacques, S. / Him, K. / Hurtado, A. / Rodrigues, L. (2008): Genetic clues about the origins of the Aché hunters-gatherers of Paraguay. In: *American Journal of Human Biology*, 20/6. P.735-737.
- Campbell, L./ Grondona, V. (eds.) (2012): *The indigenous languages of South America*. Berlin/New York. Mouton/de Gruyter.
- Cavalli-Sforza, L. L. (2001): *Genes, people and language*. University of California Press.
- Champaign, IL & Urbana, IL: University of Illinois.
- Clyne, Micheal (ed.) (1992): *Pluricentric Languages: Differing Norms in Different Nations*. Berlin/New York. Mouton/de Gruyter.
- Creanza, N./ Ruhlen, M. / Pemberton, J. T. / Rosenberg, N. A. / Feldman, M.W. / Ramachandran, S. (2015): A comparison of worldwide phonemic and genetic variation in human populations. *PNAS*, 112. P. 1265-1272.
- Dryer, M. / Haspelmath, M. (eds.) (2005): *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info>, Accessed on 2019-10-26.)
- Dyen, I. (1956): Language distribution and migration theory. In: *Language* 32/4. P. 611-626.
- Estigarribia, B. (2015): Guaraní-Spanish Jopara in a Paraguayan novel. In: *Journal of Language Contact*, 8. P. 183-222.
- Estigarribia, B. (2017): A Grammar sketch of Paraguayan Guaraní. In: B. Estigarribia / J. Pinta (eds), *Guarani Linguistics in the 21st Century*.. Leiden: Brill.
- Estigarribia, B. (2020): *A grammar of Paraguayan Guaraní*. London: UCL Press.
- Grollemund, R. S. / Brandfor, S. / Bostonen, K. / Meade, A. Venditti, C. / Pagel, M. (2015): Bantu expansion shows that habitat alters the route of human dispersals. In: *Proceedings of the National Academy of Science*, 112. P. 13296-13301.
- Hay, J. / Bauer, L. (2007): Phoneme inventory size and population size. In: *Language*, 83. P. 388- 400.
- Hill, K. / Hurtado, M. (1996): *Aché life history: The ecology and demography of a foraging people*. New York, Routledge.
- Hochmann, J. R. / Benavides-Varela, S. / Nespors, M. / Mehler, J. (2011): Consonants and vowels: different roles in early language acquisition. In: *Developmental Science*, 14. P. 1467-1687.
- Lemle, L. (1971): Internal classification of the Tupi-Guarani linguistic family. In: Bendor- Samuel D. (ed.) *Tupi Studies*. Norman: Oklahoma: Summer Institute of Linguistics, 107-129.

- Lewis, M. P. (2009): *Ethnologue: Languages of the World*. Sixteenth edition. Dallas, Texas: SIL International. <http://www.ethnologue.com/16>.
- Maddieson, I./ Precoda, K. (1990): The UCLA Phonological Segment Inventory Database, 1990. http://web.phonetik.uni-frankfurt.de/upsid_info.html.
- Marques, A. M. / Renesto, E. (2017): Genetic variability and evidence of founder effect in *Hemiodus orthonops* (Characiformes: Hemiodontidae) from the upper Paraná River Basin, Brazil. IN: *Acta Scientiarum, Biological Sciences*, 39. P. 53-58.
- Martins, M. F. (2003): *Descrição e análise de aspectos da gramática do Guarani Mbyá*. PhD. Dissertation, Universidade Estadual de Campinas.
- Mayr, E. (1954): Change of genetic environment and evolution. In.: Huxley J. et al. (eds)., *Evolution as a Process*. London: Allen & Unwin, 157-180.
- Mayr, E. (1963): *Animal species and evolution*. Cambridge, Massachusetts: Harvard University Press.
- Meira, S. / Drude, S. (2015): A summary reconstruction of Proto-mawetí-guarani segmental phonology. In: *Boletim do Museu Paranaense Emílio Goeldi. Ciências Humanas*, 10. P. 275- 69.
- Mello, A. (2000): *Estudo histórico da família linguística tupí-guaraní: Aspectos fonológicos e lexicais*. Ph.D. dissertation, Universidade Federal de Santa Catarina.
- Michael, L., N. / Chousou-Polydouri, B. / Donnelly, E. / Wauters, V. / Meira, S./ O'Hagan, Z. (2015): A Bayesian phylogenetic classification of Tupí-Guaraní. In: *LIAMES*, 15. P. 193-221.
- Muhr, R. (2016): *Pluricentric languages and non-dominant varieties worldwide*. Berlin: Peter Lang. In: Rudolf Muhr et al. (eds.). *Exploring Linguistic Standards in Non-Dominant Varieties of Pluricentric Languages*. Frankfurt am Main, etc. Peter Lang Verlag. 343-363.
- Nespor, M. / Peña. M. / Mehler. J. (2003): On the different roles of vowels and consonants in speech processing and language acquisition. In: *Lingue e Linguaggio*, 2. P. 2003-2229
- Noelli, F. S. (1998): The Tupi: Explaining origin and expansion in terms of archaeology and of historical linguistics. In: *Antiquity*, 72. P. 648-663.
- Noelli, F. S. (2008): The Tupi expansion. In H. Silverman and H. William (eds.) *Handbook of South American archaeology*. New York: Springer. 659-670.
- O'Hagan, Z. (2014): *A Computational-phylogenetic Classification of Tup'í-Guaraní and its Geographical Spread*. Paper presented at Language Variation and Change, University of Chicago. (Slides available at [http://linguistics.berkeley.edu/~zjohagan/pdflinks/ohagan_lvc_classifica](http://linguistics.berkeley.edu/~zjohagan/pdflinks/ohagan_lvc_tg_classifica)

- tion_spread_2014.pdf. (Last accessed: 08/02/2022).
- O'Hagan, Z. (2015). A Bayesian phylogenetic classification of Tupí-Guaraní. In: LIAMES, 15. P. 193-221.
- O'Hagan, Z. / Chousou-Polydouri, N. Micheal, L. (2019): Phylogenetic classification a northeastern Amazonian Proto Tupí-Guaraní homeland. In: LIAMES, 19. P. 1-29.
- Perreault, C. / Mathew, S. (2012): Dating the Origin of Language Using Phonemic Diversity. In: PLoS One, 7. e352-89.
- Rößler, E. M. (2008): Aspectos da gramática achê: Descrição e reflexão sobre uma hipótese de contato. MA Thesis, Universidade Estadual de Campinas.
- Rodrigues, A. D. (1964): Classificação do tronco linguístico Tupi. In: Revista de Antropologia, 12: 99-104.
- Rodrigues, A. D. (1978): A língua dos índios Xetá como dialeto Guaraní. In: Cadernos de Estudos Linguísticos, 1. P. 7-11.
- Rodrigues, A. D. (1985): Relações internas na família linguística Tupí-Guaraní. In: Revista de Antropologia, 27. P. 33-53.
- Rodrigues, A. D. (1986): Línguas brasileiras: para o conhecimento das línguas indígenas. São Paulo: Loyola.
- Rodrigues, A. D. (2005): As vogais orais do Proto Tupí. In.: Rodrigues, A. D. / Cabral, A. S. A.C. (eds.) Novos estudos sobre línguas indígenas. Brasília: Editora da UnB, 35-46.
- Rodrigues, A. D. (2007): Tupí languages in Rondônia and in eastern Bolivia. In: Wetzels, Leo (ed.): Language endangerment and endangered languages. Leiden: CNWS Publications, 355-363.
- Rodrigues, A. D. / Cabral, A. (2002): Revendo a classificação interna da família linguística Tupí- Guaraní. In.: A. Cabral / A. D. Rodrigues (eds.), Línguas indígenas brasileiras: Fonologia, gramática e história. Belém: UFPA, 327-337.
- Rodrigues, A. D. / Cabral. A. (2012): Tupian. The indigenous languages of South America: A comprehensive guide. Berlin/New York. Mouton/de Gruyter, 495-574.
- Rodrigues, A. D. / Dietrich, W. (1997): On the linguistic relationship between Mawé and Tupi- Guaraní. In: Diachronica, 14. P. 265-304.
- Rodrigues, C. (2020): Founder effect in Tupian languages. In: Revista Diadorim, 22. P. 65-97.
- Santos, E. / Silva, A. / Ewerton, P. / Takeshita, L. / Maia, M. (2015): Origins and demographic dynamics of Tupí expansion: a genetic tale. IN: Boletim do Museu Paranaense Emílio Goeldi. Ciências Humanas, 10. P. 217-28.

- Sapir, E. (1916): Time perspective in aboriginal American culture, a study in method. Ottawa: Government Printing Office.
- Schleicher, C. (1998): Comparative and internal reconstruction of the Tupi-Guarani language family. PhD Dissertation. University of Wisconsin, Madison.
- Trudgill, P. 2011. Social Structure and Phoneme Inventories. *Linguistic Typology*, v. 15, p. 155-160
- Ramachandran, S. / Deshpande, O. / Roseman, C. / Rosenberg, N. / Feldman, M./ Cavalli-Sforza, L. (2015): Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. In: *Proceedings of the National Academy of Sciences of the United States of America*, 102. P. 15 947.
- Urban, G. (1992): A história da cultura brasileira segundo as línguas nativas. In Cunha, M. C. (ed.) *História dos Índios no Brasil*. São Paulo, FAPESP/SMC/Cia das Letras. 87-102.
- Urban, G. (1996): On the geographical origins and dispersions of Tupian languages. In: *Revista de Antropologia*, 39. P. 61-104.
- Walker, R. S. S. /Wichmann, T./ Mailund, C. J./ Atkisson, Q. (2012): Cultural phylogenetics of the Tupi language family in lowland South America. In: *PLoS ONE* 7/4. e35,025.