Check for updates

Article

SAGE Open
October-December 2015: 1–11
© The Author(s) 2015
DOI: 10.1177/2158244015618235
sgo.sagepub.com
SAGE

Retention and Attrition of Umbundu in São Tomé and Príncipe

Gerardo Augusto Lorenzino[1]

Abstract

A major event in the modern history of São Tomé and Príncipe, located in the Gulf of Guinea, was the recruitment of contracted labor from Angola, Mozambique, and Cape Verde between 1870 and 1960. This article will address one linguistic effect of this large-scale migration, namely, the retention and attrition of the Umbundu language spoken by the Tongas, who are descendants of the indentured laborers who never repatriated back to their homes in Africa. As opposed to noncontact Umbundu, the Tonga Umbundu (TUm) of São Tomé has a simplified nominal class system and significant borrowings from the Portuguese lexicon, which became part of the TUm class system through incorporation. The sociolinguistic factors that will be considered in evaluating the nature of the Tongas' linguistic heritage are (a) the language situation that best explains the retention of Umbundu and (b) the typological distance between TUm, Afro-Portuguese Creoles, and Portuguese language.

Keywords

Portuguese, Creoles, African linguistics, language contact

The Gulf of Guinea islands of São Tomé and Príncipe (hereafter STP) offer an interesting case study for the social, historical, and linguistic factors underlying the full cycle of a language, from its emergence and development to linguistic changes, attrition, and death in a colonial setting. Since its discovery and settlement by the Portuguese in the late 15th century, STP has been shaped by two major historical events with lasting linguistic consequences:

1. First, the intensive sugar plantation–based economy and the concomitant need for renewable labor led to linguistic contact between Portuguese and African languages, first from the Benin region and then from the Congo. This language-contact situation took place within a social matrix that hindered normal language transmission across generations. The low White/slave population ratio favored the development of new patterns of communication which, among other things, led to the emergence of three Creole languages, namely, *Lungwa Santomé* or Santomense (ST), *Lungwa Ié* or Principense, and *Lunga Ngolá* or Angolar (Ferraz, 1979; Günther, 1973; Maurer, 1995, 2009).
2. Second, at the time of the abolition of slavery in 1869, cocoa and coffee were sources of lucrative foreign exchange for the plantation owners. In 1875, the freed slaves or *libertos* (in Portuguese) abandoned the plantations, which in turn forced the plantocracy to seek labor from other colonial territories to replace slave labor. As a result, a second and no less important demographic upheaval resulted with the arrival of indentured laborers from the colonies of Angola, Mozambique, and Cape Verde in STP. This article will address the relevant linguistic factors related to the second event, that is, the language situation resulting from the social marginalization that indentured workers and their offspring, or *Tongas*, experienced at the hands of the native population.

Sociohistorical Background

The recruitment of contracted labor from Angola, Mozambique, and Cape Verde lasted nearly a 100 years, from the 1870s until the late 1950s. Table 1 shows the distribution of indentured workers by country of origin in 1950, including the Tongas born in STP.

In 1827, the STP population of 7,612 inhabitants had remained stationary since the late 16th century as a result of Portugal's increasing involvement in Brazil during the two previous centuries, which relegated STP's principal economic role to the slave trade (Morgado, 1957; see also Figure 1). The population, which had reached approximately 12,000 in the 1850s, increased more than fivefold to 64,221 inhabitants by 1906, mostly due to the influx of laborers from Portuguese

[1]Temple University, Philadelphia, PA, USA

Corresponding Author:
Gerardo Augusto Lorenzino, TempleUniversity, 1114 Polett Walk, Philadelphia, PA19046, USA.
Email: galorenz@temple.edu

Creative Commons CC-BY: This article is distributed under the terms of the Creative Commons Attribution 3.0 License (http://www.creativecommons.org/licenses/by/3.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

Table 1. The Countries of Origin for Indentured Workers in STP in 1950.

Nationality	Population	%
Tongas	2,696	11.4%
Angolans	9,680	41.0%
Mozambicans	4,917	20.8%
CapeVerdeans	6,320	26.8%
Total	23,613	

*Source.*Tenreiro (1961, p. 191).
Note. STP = São Tomé and Príncipe.

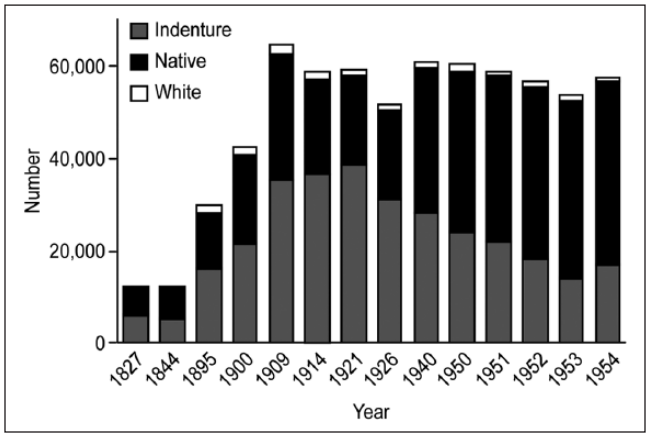


Figure 1. Population of STP (1827-1954).
Note. STP = São Tomé and Príncipe.

colonies in continental Africa, who were recruited to work on cocoa and coffee plantations (Morgado, 1957).

With so many indentured workers arriving in STP in less than a century, one naturally asks what the linguistic consequences of such demographic changes were. Indeed, except for the Cape Verdeans, who were fluent in their own variety of one of the several Creole varieties that had emerged earlier in the Cape Verde Islands and a few with some access to Portuguese, many indentured workers only spoke their native languages, such as Kimbundu, Umbundu (Um), and Sele, upon their arrival to STP (Clarence-Smith, 1990). Even though plantation workers were initially under 5-year renewable contracts, it was not until 1910 that the first laborers began to be repatriated, thus nullifying the common practice of automatic recontracting (Clarence-Smith, 1990; Hodges & Newitt, 1988). The Kimbundu lyrics below drawn from *batukes* and *pwitas*—traditional songs and dances of Angolan indentured laborers—express the experience of what became for many a one-way ticket to São Tomé:

Ko San Tomé	In São Tomé
Kuri o'n bundi o ku nyingira	there is a door to enter
Ka kuri o'n bundi o kupita.	but there is no door to leave.

(Eyzaguirre, 1986, p. 187)

Ethnicity and space played an important role in the socialization practices of contracted laborers, as it often related to the identity of the group's language within the relatively fixed boundaries of a plantation setting. The distribution of workers in the large plantations of São Tomé reflected the practice at the time of hiring workers from the same geographical area in Africa and from related ethnolinguistic groups. On the plantations, the main languages spoken were "Monte Café" (the Um language),[1] "Bela Vista" (the Kimbundu language), and "Santa Margarida" (the language or languages of Mozambican indentured workers of unknown ethnic origin; Rougé, 1992, p. 172).[2] The children of indentured workers remained on the plantations where their parents had been hired to work upon arrival from Africa until the practice of automatic recontracting was disallowed at the turn of the 20th century (Hodges & Newitt, 1988).

Language Attrition and Tonga Umbundu (TUm) in the African Language Scenario

There is no agreement among language attrition researchers on what language attrition is. Some of the questions address fundamentally whether languages can be attrited, the factors (internal, external) that cause L1 attrition, the mechanisms that affect sublinguistic levels (lexicon, semantics, morphology, phonology, etc.), and methodological aspects for analyzing attrition data given the interdisciplinary approaches. It goes beyond this article's scope a critical review of the literature on language attrition and language death. However, it will be illustrative to refer to some of the works that have been more explicit in describing the structural changes affecting dying languages.

Nancy Dorian's research about the death of the East Sutherland dialect of Scottish Gaelic is an influential longitudinal study on the grammatical changes and the external considerations (e.g., sociolinguistic attitudes, language prestige) that favor the borrowing of features from English as the dominant language. Among the Gaelic speakers, she identified younger ones who spoke the language imperfectly ("semi-speakers") but "were very much more at home in English" (Dorian, 1973, p. 417). The notion of the semispeaker proved a major incentive to following research on the linguistic, social, and psychological effects of bilingualism.

Two edited books published in the 1990s, *First Language Attrition* (Seliger & Vago, 1991) and *Investigating Obsolescence: Studies in Language Contraction and Death* (Dorian, 1992), reawakened the research began by Dorian in the 1960s and 1970s.[3] There, two chapters, one by Campbell and Muntzel, and

the other by Hamp, address specifically the structural consequences of language death on all linguistic levels (phonology, morphology, syntax, semantics, and lexicon) with regard to intergenerational variables, language contact, and emigration. They claim that the changes in the attrited language will favor less over more marked features, though the latter are likely to be retained in the dominant language if they have a high functional load. This conservatism is manifested even when the endangered language ("dialect death without capitulation") is not passed to the next generation (Hamp, 1992, p. 205). The authors argue that the structure of the dominant language and L1 markedness are at play in the incidence of specific changes in the bilingual speaker. Moreover, the prestige of the dominant language will affect changes in the attrited language, especially when a feature is incompatible between the two variants (Campbell & Muntzel, 1992).

A number of hypotheses have been formulated for explaining linguistic phenomena associated with attrition. Schmid enumerates five criteria commonly used in attrition studies—singly or in combination—pointing out that they do not always distinguish L1 attrition from other nonrelated attrition changes resulting from language interference, for example, analytical structures in L1 attrition as the outcome of contact with a synthetic language system (Schmid, 2002). Those criteria for evaluating L1 attrition may be summarized as follows: (a) Acquisitional processes such as Jakobson's regression hypothesis that states the sequence of acquisition might determine the sequence of attrition, that is, linguistic forms learned last will be lost first as the language attrite (Andersen, 1982), (b) interlanguage effects, (c) contact-induced language change, (d) Universal Grammar (UG), and (e) retrieval processes. The interlanguage hypothesis summarily states that in a contact situation, the L2 system will have an effect on the L1. Lexically, the interlanguage shows in codeswitching and borrowing from the L2 into the attrited L1, though it is less known how it impinges on the morphology and syntax, even though in cases of intense contact, codeswitching may lead to borrowing of function words and inflections, and eventually language shift and language death (Myers-Scotton, 1992). Internal simplification was often invoked as the main criterion to explain L1 attrition. However, studies by Andersen (1982), Maher (1991), and Seliger and Vago (1991; quoted in Schmid, 2002) suggested that attrition may exhibit the same pattern as contact-induced language change.[4] For example, both processes share a decrease in registers and less frequently used words as well as less complex morphological structures and the replacement of synthetic by the analytical constructions. The regression hypothesis and the UG explain language attrition on acquisition sequences. Unlike the regression hypothesis, UG focuses on grammar (parameter setting) rather than the observable sequences of acquisition, suggesting that marked linguistic features of L1 are unchanged (i.e., not attrited) as long as those features differ from L2 in a language-contact situation. The retrieval hypothesis is based on psychological considerations about language retrieval in bilinguals, that is, whether linguistic knowledge is permanently or temporarily lost, and it is informed by the competence and performance debate in language attrition (Schmid, 2002).

Multilingualism and multiculturalism are pervading among the more than 50 states in Africa, in which a European language (English, French, or Portuguese) became the official language after independence and one or more vernacular languages spoken by a larger population or enjoying a more prestigious or public status were elevated to national languages over minority languages. The latter are demographically inferior and, in several cases, have minimal public functions, mostly limited to communication in the family and the village. Their displacement by the more widespread national languages is of great concern for educators, literacy experts, and language policy makers in Africa.

According to Batibo (2005),

> A minority language can be identified horizontally by looking at its weak or non-dominant position in relation to other languages in the region or nation, and vertically on the basis of its low status and absence of use in public or official areas. Following this definition, most of the African languages would be designated as minority languages in view of their relative demographic, political and socio-economic inferiority. (p. 51)

Known as languages of wider communication, area importance or *lingua francas*, Swahili (Kenya, Tanzania); Kimbundu, Kikongo, and Um (Angola); Lingala (Democratic Republic of Congo); and Hausa, Yoruba, and Igbo (Nigeria) are responsible for language shift and death of some of the minority languages in Africa. Batibo (2005) and Sommer (1992) present a preliminary survey on language endangerment in Africa identifying minority or ethnic languages undergoing language shift and, in some cases, some already extinct.[5]

The linguistic changes documented below for TUm are not unique to the Tongas, but rather shared to a greater or lesser degree with languages (Bantu and non-Bantu) spoken in other Lusophone and non-Lusophone African countries. Furthermore, TUm attrition should be understood as a phenomenon in language endangerment resulting from the political, social, economic, and linguistic subordination that paved the way to the displacement of indigenous languages by Portuguese as L1, Portuguese growth as L2, and the emergence of new languages (pidgins, creoles).[6]

Four of the 39 languages spoken in Angola are "highly endangered languages" (Nyendo, Kung-Ekoka, Maligo, and !Kung), whereas two (Kwadi, Kwisi) are already extinct (Batibo, 2005). The political and ethnic conflicts that took place since Angolan independence until the 2002 cease-fire between the MPLA (Movimento Popular de Libertação de Angola [Popular Movement for the Liberation of Angola]) in power and the opposition UNITA (União Nacional para a Independência Total de Angola [National Union for the Total

Table 2. Umbundu's Noun Class System.

Class	Singular	Plural	Semantic field
I	⦰.Suku "God" ⦰.mume "brother"	**vo**.Suku "Gods" **va**.mume "brothers"	Relatives/proper names
II	**u**.kwenje "boy" **u**.Suku "divinity"	**a**.kwenje "boys" **ovo**.Suku "divinities"	Person/abstract nouns/ parts of the body
III	**e**.timba "body"	**a**.timba "bodies"	Mixed group
IV	**o**.mbwa "dog"	**olo**.mbwa "dogs"	Animals
V	**oku**.popya "speaking"	no plural form	Nominal verbs
VI	**oci**.ti "big tree"	**ovi**.ti "big trees"	Augmentatives
VII	**olu**.nhi "little fly"	**olo**.nhi "little flies"	Mixed group
VIII	**omu**.kanda "letter"	**amu**.kanda "letters"	Mixed group
IX	**ow**.ato "canoe"	**aw**.to "canoes"	Class IV vowel-initial nouns
X	**oko**.tima (<**oka**.u.tima) "little heart"	**otu**.tima "little hearts"	Diminutives

Note. Classification adapted from Chatelain (1888-1889), Valente (1964), and Schadeberg (1990).

Independence of Angola]) resulted in people's urban displacement as well as ethnic and unequal power sharing that, ultimately, aggravated linguistic inequalities in Angola (Batibo, 2005). Whether the civil war in Angola marginalized minority languages forcing people to shift to one of the national languages or acquire Portuguese as a first language is not known, as language distribution data from the 2013 Angolan census are not available yet. The languages' list from which Angolans during the census were asked to choose first or mother tongue ("língua materna") were Portuguese, Umbundu, Kikongo, Kimbundu, and Chokwe; if none of them, they could check "others" (*outras*). In combination with information on regional demographics, the census will be very useful in correlating language dynamics with population changes (Instituto Nacional de Estatística, 2013).

For Mozambique, the last three censuses (1980, 1987, and 2007) show an increase of Portuguese as L1 at the expense of the ethnic languages. Thus, the percentages of Mozambicans reporting Portuguese as their first language ("língua maternal") were 1.2 (1980), 6.5 (1987), and 10.7 (2007), whereas those who claimed instead a Bantu language decreased for the same 27-year period: 98.8% (1980), 93.5%% (1987), and 85.2 (2007; Gonçalves, 2012). The data also report more Portuguese–Bantu bilinguals, from 27.2% (1980) to 39.7% (2007). Portuguese monolingualism and bilingualism will amount to approximately half of the population in Mozambique, which is attributed to improvements in education (literacy, graduation rates) and a growing urban population that attaches a higher social prestige and more practical value to Portuguese as the vehicle of wider communication (Gonçalves, 2012). Nonetheless, it is important to be cautious regarding census data because questions like "Em que lingua aprendeu a falar?" ("In which language did you learn to speak?") or "Sabe falar Português?" ("Do you speak Portuguese?") may be misleading in relation to the speaker's competence/performance and the state of language attrition and loss due to its self-reported nature. In this case, census information together with qualitative and ethnographic evidence will offer a more realistic picture of the patterns of language use in Mozambique (Gonçalves, 2012).

Linguistic Description of TUm[7]

Although a growing number of studies on the Afro-Portuguese Creoles of the Gulf of Guinea have appeared in the last two decades (Hagemeijer, 2007, 2009, 2013; Lorenzino, 1998, 2006, 2007; Maurer, 1995, 2009; Rougé, 2004), the languages of the Tongas are relatively unknown, except for Jean-Louis Rougé's pioneering article published in 1992.[8] Thus, filling this gap in our knowledge of the Tongas' communication practices, this study intends, moreover, to recognize their presence in modern STP society.

The internal processes underlying the attrition of TUm result from its increasing use as a second language by second and third generation descendants of Um-speaking indentured workers due to the descendants' language shift to either Portuguese or ST. Attrition processes in TUm with morphological and syntactic results are as follows:

1. The simplification of the noun class system found in the typical Bantu language.
2. The borrowing and incorporation of Portuguese words into the TUm simplified noun class system through the lexicalization, morphologization, and overgeneralization rules of first language attrition.
3. The differential borrowing of ST words, which, unlike the Portuguese-derived lexicon, did not undergo affixation.

Linguistic Simplification

Table 2 summarizes a list of prefixes for the 10 different classes in which a noun can be declined in Um.

The system of class affixes in TUm was significantly reduced due to competing forms yielding to the least complex variants. Thus, attrited nouns in TUm generally no longer share the agglutinative morphology of Um and other Bantu languages. Out of the over 20 Um noun prefixes, only two pairs have survived in TUm, here abbreviated as Pair I and Pair II, respectively (Rougé, 1992).

Pair I is formed by the prefixes o- and a-; it is the only one still productive in number assignment (singular and plural) to nouns borrowed from Portuguese.

Pair II is formed by prefixes o- and olo-; it is no longer productive.

There may be semantic, phonological, and syntactic constraints for explaining the retention of Um o-/a- as nominal prefixes or augments in TUm (Schadeberg, 1986). These prefixes can be seen as having article-like meanings in Um mapping the definiteness/indefiniteness contrasts commonly found in Portuguese definite articles "o" and "a," respectively. Furthermore, the usage of Um augments in all syntactic environments might have favored the encompassing prefixation of TUm o-/a- to all Um nouns regardless of class origin.

Prior to explaining the internal constraints affecting the retention of Pairs I and II prefixes, examples of TUm nouns that experienced full, partial or no change at all are given below.

Full change. The singular and plural forms of the noun are changed by incorporating the singular o- and the plural a- independent of the original class group in Um. Occasionally, one may find a singular noun with the original Um prefix, but it is now lexicalized as part of the TUm word stem. For example, the prefix oci- in *ocina*/*ovina* (1a) was partially lexicalized, and morphological reanalysis took place via o-/a- prefixation. Likewise, Um ***e***.*yo*/***ova***.*yo* "eye/s" became TUm ***o***.*yo*/***a***.*yo*.

(1a) Um ***oci***.*na*/***ovi***.*na*> TUm ***o***.*cina*/***a***.*cina* "thing/s"

(1b) Um ***oci***.*banda*/***ovi***.*banda*> TUm***o***.*cibanda*/***a***.*cibanda* "healer/s" (Rougé, 1992, p. 175)

Partial change. Either the singular or the plural TUm form maintains the original o- or a- Um prefix; the singular form in (2a) already has the prefix o-. Pluralization with the prefix a- is still productive and supported by a sandhi rule, which reduces the diphthong /ao/ to /a/, as in (2b).

(2a) Um ***o***.*mbesi*/***olo***.*mbesi*> TUm ***o***.*mbesi*/***a***.*mbesi* "fish"

(2b) *Sandhi Rule*: Um *wakol**ao**po>wakol**a**po* "of health" (Valente, 1964, p. 29)

Thus, Um ***o***.*mbesi*> * TUm ***ao***.*mbesi*>***a***.*mbesi*

Unlike (2a), (2c-e) show the singular form e- undergoing a prefix change because the pluralized word already has the prefix a-.

(2c) Um ***e***.*nhulu*/***a***.*nhulu*> TUm ***o***.*nhulu*/***a***.*nhulu* "nose/s"

(2d) Um ***e***.*yaki*/***a***.*yaki*> TUm ***o***.*yaki*/***a***.*yaki* "egg/s"

(2e) Um ***e***.*singa*/***a***.*singa*> TUm ***o***.*singa*/***a***.*singa* "hair/s"

No change. Some TUm words kept the etymological Um pair of prefixes. In addition to the Um o-/a- pair, one finds words with o-/olo-.

(3a) Um and TUm ***u***.*kayi*/***a.***.*kayi* "woman/women"

(3b) Um and TUm ***o***.*njo*/***olo***.*njo* "house/es"

(3c) Um and TUm ***o***.*sanji*/***olo***.*sanji* "chicken/s"

Language attrition manifests a high degree of variation extending to even Pair II, where plural olo- shows partial lexicalization and morpheme reanalysis of vowel-initial words in the singular:

(3d) Um ***olo***.*mbwa*> *TUm ***o***.*lombwa*> TUm ***a***.*lombwa* "dogs"

↓↓

partial morpheme

lexicalization reanalysis

Morphological displacement by the more productive a- is shown below:

(3e) Um ***olo***.*mbesi*> TUm ***a***.*mbesi* "fish (pl.)"

TUm Borrowing From Portuguese

As stated above, o-/a- are the only productive prefixes in TUm that are observed in the morphological incorporation of Portuguese nouns (plurals with a- are omitted):

(4a) PT *folha* "leaf" > TUm *ofoya* (Um *olumbi*, ST *fia*)

PT *grão* "seed" > TUm *ogrão* (Um *ochipoke*, ST *ukwe*)

PT *sangue* "blood" > TUm *osangue* (Um *osonde*, ST *sangi*)

PT *figado* "liver" > TUm *ofigado* (Um *omuma*, ST *figadu*)

PT *fumo* "smoke" > TUm *ofumo* (Um *evonge*, ST *irigo*)

TUm Borrowing From ST

Interestingly, the language of the Tongas did not borrow extensively from ST Creole, which may be explained in terms of the social reality separating Tongas and STs. On one hand, the Portuguese preferred to keep indentured workers away from the Creoles; on the other hand, the latter refused

to work on the plantations, whose presence served as a painful reminder of slavery. Stigmatization of indentured life and the need to adapt to new economic conditions after independence might exert pressure on the Tongas of Monte Café to shift from Um to ST, especially among the younger Tongas.[9]

The TUm examples in (4a) were borrowed directly from Portuguese because ST words either have a different origin (TUm *ofumo*, PT *fumo*, but ST *irigo*) or are phonologically closer to Portuguese than ST (TUm *ofoya*, PT *folha*, and ST *fia*). Conversely, TUm words derived from ST have none of the characteristic prefixes attached to Portuguese-derived words (o- and a-) as shown in (4b).

(4b) ST *glaxa* (also PT *graxa*) "grease" > TUm *glaxa* (Um *ulela*)

ST *alea* "sand" > TUm *alea* (Um *eseke*)

ST *leji* "root" > TUm *lezi* (Um *olumbombo*)

The absence of class markers in ST-derived TUm words that are not derived from Portuguese may be sought in the convergence of structural and sociolinguistic factors:

- First, ST nouns are morphologically invariable for gender and number. Thus, the ST plural marker *inen* precedes the pluralized noun that has a definite and human reference: *mina*/*inen mina* "child / the children" (Alexandre & Hagemeijer, 2007).
- Second, most Portuguese nouns and adjectives are marked for gender and number, including determiners. The definite article for masculine singular nouns is o (*o homem* "the man"), which is homophonous and functionally equivalent to the noun class prefix o- (pronounced /u/). Thus, it is quite possible that the Portuguese definite article was reinterpreted by Um speakers as a noun class prefix, a phenomenon favored by the preponderance of Class IV nouns in Um (Baxter, 2002; Rougé, 1992; Valente, 1964).
- Third, olo- is the complementary pluralizer for o- in Um (see Table 2). Instead, it was a- that became the productive pluralizer in TUm. Here, the merger of Classes II and IV singular prefixes (o-, u-) to yield o- on phonological grounds (and Portuguese reinforcement), along with the choice of the shorter a- to the exclusion of olo-, may have favored the overgeneralization of the TUm a-. Alternatively, it may be argued that TUm pluralization was motivated by the reinterpretation of the Portuguese definite article a as the nominal prefix (*a mulher* "the woman" >**a**.mulher "the women"), based on the homophonous Class II noun plural prefix, for example, Um **u**.kwenje '"boy" / **a**.kwenje "boys" (Table 2). This type of lexicalization is also attested in Portuguese loanwords across the noun class system of Kimbundu, another Bantu language spoken in Angola: **di**.xalu / **ma**.xalu (Class 5/6) < PT *machado* "ax" and **m**.bolo / **ji**.mbolo (Class 9/10) < PT *pão* "bread" (Huth, 1991, pp. 29-30). At any rate, language contact with Portuguese and, to a lesser degree, with ST among the younger Tonga generations (see below) along with the decline of competent Um speakers among second-generation speakers or Tongas (Figure 2) might have contributed to the attrition of TUm.
- Fourth, additional linguistic support for the reinterpretation of the PT definite article o as a nominal class prefix is the restructuring of PT borrowed words in the Um spoken in Angola: *opato* "duck" (PT *o pato* "the duck"), *ogato* "cat" (PT *gato* "the cat"; Valente, 1964, p. 74) and *oPutu* "Portugal" (Schadeberg, 1990, p. 13).[10]
- Fifth, sociolinguistic considerations minimized the Tonga influence on ST and vice versa due to attitudinal, cultural, and spatial separation between the two speech communities (see below).[11]

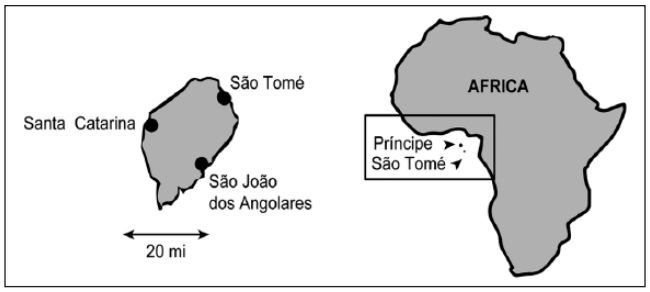


Figure 2. Map of STP.
Note. STP = São Tomé and Príncipe.

The Simplification of Umbundu's Concordance System

Umbundu has its own set of grammatical prefixes, as shown in (5):

(5a) Um ***u***.*kwenje* ***wa***.*ove*

2.child 2.2sg

"your child" (Valente, 1964, p. 132)

(5b) Um ***o***.*njo* ***ya***.*nge eyi*

4.house 4.1sg cl 4.this

"This is my house" (Schadeberg, 1990, p. 20)

According to Wurm (1991),

When the traditional culture of the speakers is largely or entirely lost as a result of culture clash, and their traditional

> world-view replaced by that of culturally more aggressive people, or more usually by a rudimentary and modified form of it, the basis for the noun class system disappears, and the noun class system itself with its accompanying grammatical features (such as concordance systems, inflections, etc.) largely or completely ceases to be used and is soon forgotten. (p. 10)

Such a case of linguistic reduction may be argued to have taken place in TUm because the language has a single concordance marker, *ya*, which serves for all noun classes and noun phrase constructions: noun + adjective/noun/adjectival verb. However, Um *cya* and *vya* are subject concord prefixes lost in TUm (6b, 7b):

Noun + adjective.

(6a) TUm *ocibanda* ***ya*** *mbote*

warrior (sg.) CONC good

"The good warrior" (Rougé, 1992, p. 175)

(6b) Um ***oci.****banda* ***cya****mbote*

6.warrior 6.good

Idem.

Noun + noun.

(7a) TUm *a.cina* ***ya*** *tata* ***ya*** *nge*

things CONC father CONC 1s

"My father's things"

(Rougé, 1992, p. 175)

(7b) Um *ovina* ***vya****tate* ***ya****nge*

6.things 6.father 6.1 s

Idem.

Other Types of Linguistic Restructuring in TUm

In addition to the reduction of the noun and agreement class system, the following phenomenon is also evidence of linguistic restructuring in TUm:

- lexical periphrases: comparative adjectives

(8a) Um *wa mbote* "good," *hise* "better"

TUm *wa mbote* "good," *wa mbote muntu* (literally "good much")

(< PT *muito* or ST *muntu* "many, much," but Umbundu *lwa*)

- code mixing:

(8b) TUm: *nd.a.kalala ko* ***losa dez anos***[12]

1sg.PAST.work LOCATIVE plantation ten years

"I worked on the plantation for ten years."

(Um *ocitaka* "plantation," *ekwi* "ten," *elima* "year")

- loss of contrast between tense and aspect affixes (negative 1sg. si-):

(8c) TUm: ***ka****.ndi.sena kalala ko losa*

NEG.1sg.PRES.want work LOC plantation

"I don' want to work on the plantation."

(Umbundu **si**ndisena[13])

Sociolinguistic Characterization of the TUm Speech Community

The Portuguese variety used by the Tongas shares a number of features, such as the reduction of consonant clusters, rhotacism, and noun and verb phrase simplification, with the Portuguese spoken by the general population of STP. However, it is the Um language of the Tongas that sets them apart from the Creole-speaking population. The latter are bilingual in any of the three Creole varieties and Portuguese, but they do not speak a continental African language (Um, Kimbundu). Based on ethnographic observations on language use, Um among the older speakers remains a viable home language. Young Tongas do not speak their parents' language, a language shift that has intensified through intermarriage with other groups, especially ST Creoles, and work outside the plantations (Rougé, 1992).[14]

The sociolinguistic history of the STP plantations, beginning with the hiring of indentured labor from Africa in the 1870s, correlates with the retention of Um, followed by the attrition of TUm and its final disappearance (language death). This situation imposed, in addition, constraints on the degree of linguistic interference and transferred features between Tongas and STs.[15] In this respect, Um's retention by first-generation indentured workers was favored by the rigid plantation system with its demarcation of language domains and group isolation, as well as a colonial policy that did not allow ST workforce on the plantations. Another contributing factor to its continuity was the typological distance between Um and ST. It is noteworthy that while speakers of Creole languages in STP view their languages as dialects (ST *dialetos*), African languages are thought of as *línguas pesadas*

("heavy languages") that are difficult to learn. The negative attitude held by STs toward their language as a corrupted version of Portuguese (Morais-Barbosa, 1967) has been largely eliminated after independence, as there has been a growing awareness, at least among the intellectual elite, of the African contributions to the genesis and development of Creole languages and their presence in ST literature (de Campos, 2010; Mata, 2001). This awareness resulted, to a large extent, from STP's relatively new nationhood since independence from Portugal in 1975 and its subsequent renewed ties with Africa (Espírito Santo, 1985; Mata, 1998).

Finally, Tongas' limited influence on ST stemmed from their shared experience and ethnic and geographical roots with their ancestors in Africa. In turn, this influence helped Tongas occupy intermediate roles in the plantation social structure, between a small European managerial population and a considerable migrant population as shown in Table 1 (Clarence-Smith, 1990). In fact, among the native population (STs), this hierarchy produced a scornful attitude toward plantation life and segregation, which was seen as subservient to colonial power (Hodges & Newitt, 1988). Indeed, Tongas who were associated with colonial power seem to have been unlikely agents for changes in ST. Nonetheless, independence propagated better intergroup relationships through government policies that set out to eradicate social and ethnic divisions among all communities in STP (Nazare Ceita, 1991). Literacy and the media have played an important role in extending the use of Portuguese as a lingua franca. In addition, limited urbanization and some social mobility—two influential factors preventing competition for the same language domains—have eased, but social compartmentalization has not completely been dismantled in STP.

Conclusion

Language retention and attrition for the Um-speaking Tongas of STP were characterized in structural and sociolinguistic terms. Their language shows a reduced nominal class system with all nonfunctional markers, except for o-/a- prefixation. In addition, lexical borrowing is salient in TUm, where Portuguese borrowed words underwent noun affixation with o-/a- prefixes. Their retention in TUm and the lack any of these morphological features in ST were explained as an outcome of lexicalization and overgeneralization.

The sociohistorical basis for Tongas' linguistic heritage is rooted in the colonial Portuguese plantocracy, which remained in place until the mid-20th century. This labor practice served as a social barrier that separated the Tongas from their surrounding Creole-speaking population. The linguistic differences between their Um language and the Afro-Portuguese Creole languages underlie this type of language situation (retention and attrition), which describes the Tongas' speech community.

Acknowledgments

I would like to thank two anonymous reviewers for their constructive comments and helpful suggestions for improving the final version of this article. I would also like to thank Dr. Tjerk Hagemeijer, the editor of the article, for his generous comments and support during the review process.

Author's Note

All remaining errors are the sole responsibility of the author.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research and/or authorship of this article.

Notes

1. Umbundu, also known as M'bundo, Mbundu Benguela, and Ovimbundu, is spoken in Western Angola, mostly in Benguela and Huambo. It enjoys wider use, as it is spoken by 38% of the population as the first or second language and also serves as a trade language; Portuguese is the official language of the country (Lewis, Simons, & Fennig, 2014).
2. *Ethnologue* (Lewis et al., 2014) lists 43 indigenous languages spoken in Mozambique. A 1989 report on language and education in Mozambique languages prepared by *Núcleo de Estudo das Línguas Moçambicanas* (NELIMO; Unit for the Study of Mozambican Languages) includes 20 Bantu languages (quoted in A. Lopes, 2004). This source discrepancy on Mozambique's current language situation points to the difficulty in assessing the linguistic makeup of indentured laborers from Mozambique in the Santa Margarida plantation.
3. Lambert and Freed (1982) and Thomason and Kaufman (1988) were two important precursors to Dorian (1992) by contributing programmatic approaches that included assessment tools and concepts for the structural and social constraints underlying language attrition and loss. For a comprehensive review of the literature on language attrition, see Köpke and Schmid (2007).
4. Sasse (1992) discusses the similarities and differences between language attrition ("decay") and normal language transfer in contact situations in the context of Africa.
5. Sommer's (1992) and Batibo's (2005) studies are general language surveys and, thus, do not attempt to present a case-by-case analysis of the causes underlying language endangerment in Africa; thus, it is unclear whether the European language, a regional language, or both are responsible for language death.
6. The social settings in Angola and Mozambique (and non-Lusophone African countries) during colonial times may not be similar to the type of residential segregation that laborers had to endure in São Tomé's plantations. Therefore, a comparison between attrition in Tonga Umbundu (TUm) and minority languages in the continent coming under pressure from Portuguese and national languages (including Santomense [ST]) must identify linguistic and nonlinguistic constraints in each particular case.

7. My interest on TUm came about during fieldwork in the 1990s while I was collecting data on Angolar, one of the three Afro-Portuguese Creoles spoken in São Tomé and Príncipe (STP; Lorenzino, 1998). The analysis presented here results from a 2-hr recording made of natural speech, sentences, and word lists elicited from two TUm speakers, both male and in their 60s. The conclusions are necessarily tentative on the basis of this small corpus. Further research will be needed to better understand the linguistic situation of the Tongas. Language examples are mine unless otherwise indicated.
8. See also Baxter (2004), N. Lopes and Baxter (2011), Da Silva and Baxter (2011), Da Silva (2012), and N. Lopes (2009) for recent studies on the Portuguese variety spoken by the Tongas.
9. Future research should be able to show whether Tongas are actually shifting from Umbundu to ST or a local variety of Portuguese. Language shift and loss among the Tongas are among the several examples of the linguistic consequences related to indentured labor during postcolonial times in Africa, Asia, and the Caribbean. The case of Trinidad and Tobago, though not fully identical to the Tongas', shows how Indian, Portuguese, and Yoruba indentured laborers shifted to the Creole spoken by former slaves (Mufwene, 2004).
10. Like Umbundu, Angolan Kimbundu went through a similar process of Portuguese lexical borrowing and Kimbundu morphological incorporation: PT *televisor*> Class 9/10 *televisole* / *jitelevisole* (Huth, 1991).
11. However, the lexicon of flora and fauna in STP contains a considerable number of words that derive from the languages spoken by the indentured laborers (Silva, 1959).
12. One reviewer rightly pointed out that the TUm sentence in (8b) involving Portuguese numerals in a Umbundu matrix sentence may not be code mixing because speakers of Bantu languages in continental Africa in contact situations unlike the one encountered in STs (i.e., attrition) often use European numerals (Portuguese, English, French) when counting. Here, a distinction between internally from externally motivated language change will be crucial in determining whether code mixing and other grammatical changes (simplification, reduction) are indicative of linguistic attrition or normal transmission.
13. In some Umbundu dialects, the negative prefixes ka- (1s) and kú- (2s) have the variants ha- and hú-, respectively (Schadeberg, 1990; Sommer, 2003).
14. As rightfully pointed out by a reviewer, there is an urgency for more research to find out whether young Tongas in Monte Café have been acquiring ST since I collected the data in the 1990s.
15. This view is also adopted by Brenzinger, Heine, and Sommer (1991), and Thomason and Kaufman (1988) with regard to the types of processes affecting language shifts. See also Köpke, Schmid, Keijzer, and Dostert (2007) for a general perspective on first language attrition processes.

References

Alexandre, N., & Hagemeijer, T. (2007). Bare nouns and the nominal domain in Santome. In M. Baptista & J. Guerón (Eds.), *Noun phrases in Creole languages: A multi-faceted approach* (pp. 37-60). Amsterdam, The Netherlands: Benjamins.

Andersen, R. (1982). Determining the linguistic attributes of language attrition. In R. Lambert & B. Freed (Eds.), *The loss of language skills* (pp. 83-118). Rowley, MA: Newbury House.

Batibo, H. (2005). *Language decline and death in Africa: Causes, consequences and challenges*. Buffalo, NY: Multilingual Matters.

Baxter, A. (2002). Semicreolization?—The restructured Portuguese of the Tongas of São Tomé, a consequence of L1 acquisition in a special contact situation. *Journal of Portuguese Linguistics*, *1*, 7-39.

Baxter, A. (2004). The development of variable NP plural agreement in a restructured African variety of Portuguese. In G. Escure & A. Schwegler (Eds.), *Creoles, contact, and language change: Linguistic and Social implications* (pp. 97-126). Amsterdam, The Netherlands: Benjamins.

Brenzinger, M., Heine, B., & Sommer, G. (1991). Language death in Africa. In R. H. Robins & E. M. Uhlenbeck (Eds.), *Endangered languages* (pp. 19-44). Oxford, UK: Berg.

Campbell, L., & Muntzel, M. C. (1992). The structural consequences of language death. In N. Dorian (Ed.), *Investigating obsolescence: Studies in language contraction and death* (pp. 181-196). Cambridge, UK: CambridgeUniversity Press.

Chatelain, H. (1888-1889). *Grammatica elementar de Kimbundu ou Lingua de Angola* [Elementary grammar of Kimbundu, the language of Angola]. Ridgewood, NJ: Gregg.

Clarence-Smith, W. (1990). The hidden costs of labour on the cocoa plantations of São Tomé and Príncipe, 1875-1914. *Portuguese Studies*, *6*, 152-172.

Da Silva, L. (2012). A variação na concordância de gênero no sintagma nominal nos Tongas (em São Tomé, na África): Um contraste com o português brasileiro [Gender agreement variation in the noun phrase of Tonga (São Tomé, Africa): A comparison with Brazilian Portuguese]. In E. Costa, N. Lopes, & Y. Castro (Eds.), *II Seminario Internacional Acolhendo as Línguas Africanas* (pp. 1-10). Salvador, Brazil: Editora da Universidade do Estado da Bahia.

Da Silva, L., & Baxter, A. (2011). Os tongas: Contribução para o estudo do português brasileiro [The Tongas: A contribution for the study of Brazilian Portuguese]. In *Simpósio Nacional de Letras e Linguística e III Simpósio Internacional de Letras e Linguística, 2011, Uberlândia. Anais do SILEL*. Uberlândia, Brazil: Editora da Universidade Federal de Uberlândia. Retrieved from http://www.ileel.ufu.br/anaisdosilel/wp-content/uploads/2014/04/silel2011_351.pdf

de Campos, V. (2010). As marcas de crioulos de São Tomé na narrativa de Fernando Reis [The Creole features of São Tomé in the narrative of Fernando Reis]. *Lumen et Virtus*, *1*, 85-105.

Dorian, N. (1973). Grammatical change in a dying dialect. *Language*, *49*, 413-438.

Dorian, N. (Ed.). (1992). *Investigating obsolescence: Studies in language contraction and death*. Cambridge, UK: Cambridge University Press.

Espírito Santo, C. (1985). Situação actual da Língua Portuguesa nas ilhas de S. Tomé e Príncipe [Present situation of the Portuguese language in the islands of São Tomé and Príncipe]. In *Congresso sobre a Situação Actual da Língua Portuguesa no mundo* (vol. 1, pp. 253-260). Lisbon: Instituto de Cultura e Língua Portuguesa.

Eyzaguirre, P. (1986). *Small farms and estates in São Tomé, West Africa* (Doctoral dissertation). Yale University, New Haven, CT (ProQuest dissertation & Theses. Order No. 303507718).

Ferraz, I. (1979). *The Creole of São Tomé*. Johannesburg, South Africa: WitwatersrandUniversity Press.

Gonçalves, P. (2012, February 7-10). *Lusofonia em Moçambique: com ou sem glotofagia?* [Lusophony in Mozambique: With or without glottophagy?] II Congresso Internacional de Linguística Histórica. Homenagem a Ataliba Teixeira de Castilho, São Paulo, Brazil.

Günther, W. (1973). *Das portugiesische Kreolisch der Ilha do Príncipe* [The Portuguese Creole of Principe Island]. Marburg, Germany: Marburger Studien zur Afrika und Asienkunde.

Hagemeijer, T. (2007). *Clause structure in Santome* (Doctoral dissertation). Universidade de Lisboa, Lisbon, Portugal.

Hagemeijer, T. (2009). Initial vowel agglutination in the Gulf of Guinea Creoles. In E. O. Aboh & N. Smith (Eds.), *Complex processes in new languages* (pp. 29-50). Amsterdam, The Netherlands: Benjamins.

Hagemeijer, T. (2013). The Gulf of Guinea Creoles: Genetic and typological relations. In P. Bhatt & T. Veenstra (Eds.), *Creole languages and linguistic typology* (pp. 163-205). Amsterdam, The Netherlands: Benjamins.

Hamp, E. (1992). On signs of health and death. In N. Dorian (Ed.), *Investigating obsolescence: Studies in language contraction and death* (pp. 197-210). Cambridge, UK: Cambridge University Press.

Hodges, T., & Newitt, M. (1988). *São Tomé and Príncipe: From Plantation colony to microstate*. Boulder, CO: Westview Press.

Huth, K. (1991). Portugiesische Lehnwörter im grammatischen system des Kimbundu (Angola) [Portuguese loan words in the grammatical system of Kumbundu (Angola)]. In M. Perl & A. Schönberger (Eds.), *Portugiesischen in Afrika und Asien* (pp. 21-37). Berlin, Germany: TFM Publishing House.

Instituto Nacional de Estatística. (2013). *Distribução da População Segundo a Língua Materna* [Population distribution according to mother tongue]. Available from http://www.ine.gov.ao

Köpke, B., & Schmid, M. (2007). Language attrition: The next phase. In B. Köpke, M. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language attrition: Theoretical perspectives* (pp. 1-46). Amsterdam, The Netherlands: Benjamins.

Köpke, B., Schmid, M., Keijzer, M., & Dostert, S. (2007). *Language attrition: Theoretical perspectives*. Amsterdam, The Netherlands: Benjamins.

Lambert, R., & Freed, B. (Eds.). (1982). *The loss of language skills*. Rowley, MA: Newbury House.

Lewis, M., Simons, G., & Fennig, C. (Eds.). (2014). *Ethnologue: Languages of the world* (18th ed., online version). Dallas, TX: SIL International. Retrieved from http://www.ethnologue.com/language/umb

Lopes, A. (2004). The language situation in Mozambique. In R. Baldauf & R. Kaplan (Eds.), *Language policy and planning, Africa, vol. 1: Botswana, Malawi, Mozambique and South Africa* (pp. 155-196). New York, NY: Multilingual Matters.

Lopes, N. (2009). Um estudo do gênero nos tongas e em Helvêcia: uma comparação [A study of gender in the Tongas and Helvêcia: A comparison]. *Papia*, *19*, 141-151.

Lopes, N., & Baxter, A. (2011). A concordância verbal variável no português dos Tongas [Variable verbal agreement in Tonga Portuguese]. *Papia*, *21*, 39-50.

Lorenzino, G. (1998). *Angolar Creole Portuguese: Its grammar and sociolinguistic history* (Doctoral dissertation). City University of New York–Graduate Center, New York.

Lorenzino, G. (2006). Linguistic, historical and ethnographic evidence on the formation of the Angolares: A maroon descendant community in São Tomé (West Africa). *Portuguese Studies Review*, *14*, 1-34.

Lorenzino, G. (2007). Angolar Creole Portuguese or lingua Ngolá [Angolar Creole Portuguese or the Angolar language]. In J. Holm & P. Patrick (Eds.), *Comparative Creole syntax* (pp. 1-24). London, England: Battlebridge.

Maher, J. (1991). A crosslinguistic study of language contact and language attrition. In H. Seliger & R. Vago (Eds.), *First language attrition* (pp. 67-84). Cambridge, UK: CambridgeUniversity Press.

Mata, I. (1998). *Diálogo com as Ilhas (sobre cultura e literatura de São Tomé e Príncipe)* [Dialogue with the islands (on culture and literature of São Tomé and Príncipe)]. Lisbon, Portugal: Colibri.

Mata, I. (2001). A prosa de ficção são-tomense: a presença obsidiante do colonial [The prose of Sãotomense fiction and the besieging colonial presence]. *Revista de Filología Románica*, *11*, 207-244.

Maurer, P. (1995). *L'angolar: Un créole afro-portugais parlé à São Tomé* [Angolar: An Afro-Portuguese Creole spoken in São Tomé]. Hamburg, Germany: Helmut Buske Verlag.

Maurer, P. (2009). *Principense (Lung'ie): Grammar, texts, and vocabulary of the Afro-Portuguese Creole of the island of Príncipe, Gulf of Guinea*. London, England: Battlebridge.

Morais-Barbosa, J. (Ed.). (1967). *Estudos linguísticos crioulos* [Creole linguistic studies]. Lisbon: Academia Internacional de Cultura Portuguesa.

Morgado, N. (1957). Contribução para o estudo do problema demográfico de S. Tomé e Príncipe [Contribution to the study of the demographic problem of São Tomé and Príncipe]. *Garcia de Orta*, *5*, 633-658.

Mufwene, S. (2004). Language birth and death. *Annual Review of Anthropology*, *33*, 201-222.

Myers-Scotton, C. (1992). Codeswitching as a mechanism of deep borrowing, language shift, and language death. In M. Brenzinger (Ed.), *Language death: Factual and theoretical explorations with special reference to East Africa* (pp. 31-58). Berlin, Germany: Mouton de Gruyter.

Nazare Ceita, M. (1991). *Ensaio para uma reconstituição histórico-antropológica dos angolares de S. Tomé* [Essay for the historical and anthropological reconstruction of the Angolares of São Tomé]. Lisbon, Portugal: Centro de Estudos Africanos.

Rougé, J.-L. (1992). Les langues des Tongas [The languages of the Tongas]. In E. D'Andrade & A. Kihm (Eds.), *Actas do Colóquio sobre 'Crioulos de Base Lexical Portuguesa* (pp. 171-175). Lisbon, Portugal: Colibri.

Rougé, J.-L. (2004). *Dictionnaire étymologique des créoles portugais d'Afrique* [Etymological dictionary of the Portuguese Creoles of Africa]. Paris, France: Karthala.

Sasse, H.-J. (1992). Language decay and contact-induced change: Similarities and differences. In M. Brenzinger (Ed.), *Language death: Factual and theoretical explorations with special reference to East Africa* (pp. 31-58). Berlin, Germany: Mouton de Gruyter.

Schadeberg, T. (1986). Tone cases in Umbundu. *Africana Linguistica*, *10*, 423-447.

Schadeberg, T. (1990). *A sketch of Umbundu*. Köln, Germany: Köpper.

Schmid, M. (2002). *First language attrition, use and maintenance: The case of German Jews in anglophone countries*. Amsterdam, The Netherlands: Benjamins.

Seliger, H., & Vago, R. (Eds.). (1991). *First language attrition*. Cambridge, UK: CambridgeUniversity Press.

Silva, H. (1959). Nomes vulgares de algumas plantas de São Tomé e Príncipe [Common names of some plants of São Tomé and Príncipe]. *Garcia de Orta*, *7*, 293-323.

Sommer, G. (1992). A survey on language death in Africa. In M. Brenzinger (Ed.), *Language death: Factual and theoretical explorations with special reference to East Africa* (pp. 301-417). Berlin, Germany: Mouton de Gruyter.

Sommer, G. (2003). Western Savanna. In D. Nurse & G. Philipson x(Eds.), *The Bantu Languages* (pp. 566-580). New York, NY: Routledge.

Tenreiro, F. (1961). *A ilha de São Tomé* [The island of São Tomé]. Lisbon, Portugal: Junta de Investigações do Ultramar.

Thomason, S., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Valente, J. (1964). *Gramática Umbundu* [Umbundu grammar]. Lisbon, Portugal: Junta de Investigações do Ultramar.

Wurm, S. (1991). Language death and disappearance: Causes and circumstances. In H. Robins & E. Uhlenbeck (Eds.), *Endangered languages* (pp. 1-18). Oxford, UK: Berg.

Author Biography

Gerardo Augusto Lorenzino is associate professor of Spanish and Portuguese at TempleUniversity, Philadelphia, USA. His previous publications include *The Angolar Creole Portuguese of São Tomé* and *The Morphosyntax of Spanish-Lexified Creoles*, as well as many articles on language contact in Africa and Latin America.