languages

MDPI

Article

Yeyi: A Phylogenetic Loner in Eastern Bantu

Hilde Gunnink [1,*], Natalia Chousou-Polydouri [2] and Koen Bostoen [1,*]

[1] UGent Centre for Bantu Studies (BantUGent), Ghent University, 9000 Gent, Belgium
[2] Institute for Mediterranean Studies, Foundation for Research and Technology—Hellas, GR-741 32 Rethymno, Greece; nchousoupolydouri@gmail.com
* Correspondence: hilde.gunnink@ugent.be (H.G.); koen.bostoen@ugent.be (K.B.)

Abstract: While major advances in the subclassification of Bantu languages have been made thanks to comprehensive, lexicon-based classifications, there are still several important uncertainties obscuring not only the diachronic linguistic processes that gave rise to Bantu diversification, but also the population dynamics of ancestral Bantu speakers underlying them. In this paper, we address one of these persisting mysteries of Bantu genealogy, i.e., the unclassified Yeyi (R41) language of southern Africa. While the Bantu origin of Yeyi is straightforward and undisputed, its closest relatives are unknown, as is the major Bantu branch to which it belongs. We use a lexicon-based, Bayesian phylogenetic approach, comparing Yeyi to languages of the wider geographic region, including even more far-flung languages that have previously been hypothesized to bear a close relationship to Yeyi. The resultant linguistic phylogeny shows that Yeyi is part of the Wider Eastern Bantu branch as its own clade with Narrow Eastern Bantu languages as its closest relatives and none of its nearest neighbors. We argue that this relatively isolated position of Yeyi within Eastern Bantu suggests an early migration into southern Africa from the putative Wider Eastern Bantu homeland, which was followed by the loss of Yeyi's putative earlier sister languages, presumably through a shift to Bantu languages spoken by more recent migrants.

Keywords: Bantu; Bayesian phylogenetics; language classification

check for updates

Academic Editor: John Hajek

Received: 23 October 2024
Revised: 23 February 2025
Accepted: 8 March 2025
Published: 21 March 2025

Citation: Gunnink, H., Chousou-Polydouri, N., & Bostoen, K. (2025). Yeyi: A Phylogenetic Loner in Eastern Bantu. *Languages*, *10*(4), 55. https://doi.org/10.3390/languages10040055

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Scholars have recognized the unity of the Bantu language family since the nineteenth century CE (Bostoen & Van de Velde, 2019, p. 6). In recent decades, our understanding of Bantu subclassification has been refined thanks to Bantu-wide lexicon-based quantitative approaches, first lexicostatistics (Bastin et al., 1999) and later Bayesian phylogenetics (Grollemund et al., 2015; Koile et al., 2022). These internal classifications relying on basic vocabulary have solidified the consensus that the highest diversity within Bantu lies in the northwest. The languages in the large eastern and southern parts of the Bantu-speaking area are numerous and spoken by the vast majority of Bantu speakers, but they form a relatively late and homogenous branch (Grollemund et al., 2015; Koile et al., 2022), reminiscent of what is called 'Savanna Bantu' in earlier literature (Ehret, 1999, 2001). Although Grollemund et al. (2015) further subdivide this clade into Southwestern and Eastern Bantu, this subdivision does not match with discrete monophyletic groups in the 'ladderized' or onion-like topology of their tree, i.e., clades successively splitting off from the central backbone (Pacchiarotti & Bostoen, 2020, p. 163). In Koile et al. (2022), who use the same lexical dataset and cognacy judgements as Grollemund et al. (2015) but different parameters in their phylogenetic model, Southwestern and Eastern Bantu do emerge as

two well-supported discrete monophyletic groups, which form sister clades more closely related to each other than to the remainder of Bantu.

Despite recent advances, some areas of Bantu subclassification remain hazy, and some individual Bantu languages continue to defy classification. In this study, we focus on such a classificatory rebel, namely Yeyi (R41), a Bantu language of Botswana and Namibia. On the one hand, the Bantu affiliation of Yeyi is undeniable. It has a large number of lexemes of clear Bantu origin with cognates in Bantu languages nearby and further away (Gowlett, 1992, 1997). Beyond lexicon, it also has, for instance, a full system of noun classes marked through prefixes on nouns and agreement prefixes on dependents, which is unmistakably inherited from Proto-Bantu (Seidel, 2008, pp. 101–129). The same holds for several other domains of Yeyi grammar. On the other hand, Yeyi's genealogy within Bantu is still unclear. None of its Bantu neighbors are similar enough to be considered a close relative. It is even not clear to which major Bantu subclade Yeyi belongs, as it features in none of the lexicon-based Bantu-wide genealogical classifications, except for the lexicostatistical study of Heine et al. (1977).

In this paper, we resolve some of the uncertainties surrounding Yeyi's classification and history through a new linguistic phylogeny explicitly focused on elucidating Yeyi's genealogy. In Section 2, we introduce the Yeyi language, its modern and historical sociolinguistic setting, especially with regard to contact with other Bantu languages and non-Bantu languages, and earlier attempts at or suggestions about Yeyi classification. In Section 3, we introduce our dataset and methodology. In Section 4, we present the resultant phylogenetic classification, while we interpret these results in terms of the population history of Yeyi speakers in Section 5. In Section 6, we summarize our conclusions.

2. Previous Approaches to Yeyi Classification

Yeyi is spoken in northwestern Botswana and northeastern Namibia. In Namibia, the Yeyi speech community is concentrated in the southern part of the Zambezi region (formerly known as the East Caprivi). In Botswana, Yeyi speakers are mostly found in the Northwest or Ngamiland district, living in and around the Okavango delta. The Yeyi people consider their Namibian habitat to be their original homeland, with the expansion into Botswana dated to around 1650 (Westphal, 1964; Tlou, 1985; Larson, 1989).

In both Botswana and Namibia, Yeyi is spoken in a highly multilingual setting. In Namibia, Yeyi is in contact with Fwe (K402), Totela (K41), and Subiya (K42), all members of the western branch of the closely related and relatively well-studied Bantu Botatwe (Bostoen, 2009; de Luna, 2010), which forms a discrete subgroup within Eastern Bantu in both Grollemund et al. (2015) and Koile et al. (2022), though with only three eastern members represented, i.e., Lenje (M61), Soli (M62), and Tonga (M64). Lozi (K21) is the regional lingua franca, a Bantu language that developed when Southern Sotho speakers from South Africa fled to Western Zambia in the mid-nineteenth century CE, where they came into contact with speakers of Luyi (K31). The resultant Lozi language combines a distinctly Luyi-influenced phonology with a basic lexicon and grammatical system that betrays its Sotho origin (Gowlett, 1989). Mbukushu (K333), which is in contact with Yeyi in both Botswana and Namibia, is spoken across southern Angola, northern Namibia, western Zambia, and northwestern Botswana, and is part of the Luyana cluster of languages spoken across large parts of Western Zambia (Lisimba, 1982). Another Bantu language that borders on Yeyi in Botswana is Tswana (S31), the national language of Botswana, especially its Tawana variety. Tswana's dominance in Botswana is triggering language shift (Sommer, 1995), and, much more than Namibian Yeyi, Botswana Yeyi counts as endangered, in spite of recent and ongoing grassroots revitalization attempts (Nyati-Ramahobo, 2002; Nyati-Saleshando, 2019).

In addition to Bantu speakers, Yeyi speakers are also in contact with speakers of various Khoisan languages, particularly Khwe, a language of the Khoe–Kwadi family, and Ju, a language of the Kx'a family. This has led to heavy Khoisan influence on Yeyi, which is seen in its lexicon (Sommer & Voßen, 1992; Gunnink et al., 2015), verbal morphology (Gunnink, 2022), but most clearly in its extensive use of clicks. Clicks are cross-linguistically very rare phonemes that are ubiquitous in Khoisan languages, but are also found in some languages spoken relatively close to Khoisan, such as Bantu languages of southern Africa. This makes the use of clicks in Bantu languages a clearly recognizable sign of Khoisan interference (Vossen, 1997; Sands & Güldemann, 2009; Pakendorf et al., 2017). Click inventories in Yeyi differ between Botswana Yeyi and Namibian Yeyi, and especially in Botswana Yeyi, they seem to be subject to a great degree of regional and idiolectal variation (Sommer, 2017), possibly linked to the ongoing process of language loss (Sommer & Vossen, 1995). Botswana Yeyi distinguishes up to 22 different clicks, including four click types, dental, alveolar, lateral, and palatal (Sommer & Voßen, 1992; Fulop et al., 2003), making Botswana Yeyi the only living Bantu language to have phonemic palatal clicks (Pakendorf et al., 2017). The Namibian Yeyi click inventory is smaller, does not make use of palatal clicks, and uses lateral clicks only marginally (Donnelly, 1991, pp. 13–15; Seidel, 2008, pp. 40–43).

The extensive Khoisan influence on Yeyi may be one of the reasons why Yeyi is hard to classify within Bantu, especially if this contact-induced influence has replaced inherited Bantu material that would have been able to shed light on Yeyi's closest Bantu relatives. The unclear genealogical affinities of Yeyi have led to wildly different proposals. Andersson, a Swedish traveler who visited the Okavango delta in the mid-nineteenth century CE and collected some of the earliest written Yeyi data, considered Yeyi to share similarities with Herero (R31), a Bantu language spoken in Namibia (Andersson, 1855, pp. 19–20). However, this suggestion was probably prompted by Andersson's longer experience with Herero compared to other Bantu languages. A similar western orientation for Yeyi was implicitly suggested by Guthrie (1948), who classified Yeyi into zone R together with Umbundu (R10), Wambo (R20), and Herero (R30), all spoken in Angola and Namibia, but as a separate subgroup (R40), therefore emphasizing its relatively isolated status. Johnston (1919), on the other hand, links Yeyi to the Luyana (K30) cluster spoken in western Zambia. These languages all form part of Southwestern Bantu (Grollemund et al., 2015; Koile et al., 2022).

Other attempts at Yeyi classification have rather suggested its inclusion in Eastern Bantu. Heine et al. (1977), for example, include Yeyi in their "Osthochland-Gruppe", which roughly corresponds to Eastern Bantu in recent phylogenies (Grollemund et al., 2015; Koile et al., 2022). It has also been suggested that Yeyi may be affiliated with the Bantu Botatwe subgroup, although these suggestions do not seem to be based on much more than geographical contiguity (Tlou, 1985; Gowlett, 1997). Seidel (2005) compares Yeyi to other languages spoken in the (then) Caprivi strip of Namibia. Although this limited geographic focus prevents the identification of Yeyi relatives spoken further away, it is one of the few attempts at Yeyi classification that employs a clear methodology, comparing Yeyi to surrounding languages in terms of phonology and basic lexicon. Layered language genesis, consisting of different strata of influence from various Khoisan as well as Bantu languages, would then account for the unique profile of Yeyi (Seidel, 2009).

In sum, previous proposals of Yeyi classification have pointed in very different directions, but mostly miss a consistent methodology, or use only a relatively small sample of languages to compare Yeyi to. Bantu-wide lexicon-based classifications since 1999 have unfortunately not included Yeyi, which means that the possibility of Yeyi sharing a closer genealogical relationship with languages now spoken further away has not yet been investigated.

3. Data and Methodology

In order to clarify the genealogical position of Yeyi with respect to other Bantu languages, we provide a new, lexicon-based, Bayesian phylogenetic classification. We collected linguistic data for nine Yeyi doculects, i.e., varieties documented by different authors in different times and places. These consist of five sources on Botswana Yeyi (Livingstone, 1851; Andersson, 1855; Schapera & van der Merwe, 1942; Sommer, 1995; Lukusa, 2009) and four on Namibian Yeyi (Donnelly, 1991; Gowlett, 1992; Seidel, 2005, 2008), spanning a period of some 150 years.

Some of these sources represent Yeyi varieties that are similar enough to each other to be considered the same language. We therefore created two datasets, one where all Yeyi doculects were kept separately and one where Yeyi doculects collected in similar areas during a similar period were merged. In this latter dataset, we merged the Namibian data published in Donnelly (1991) with those published in Seidel (2005, 2008), and we merged the Botswana Yeyi data published in Schapera and van der Merwe (1942) with those published in Sommer (1995). Although Andersson (1855) and Livingstone (1851) were working in the area of modern-day Botswana, we kept their nineteenth-century CE data separate given their age and the lack of details on the exact location of data collection. As Gowlett (1992, 1997) occasionally combines Yeyi data from Namibia and Botswana, we removed this source from the second dataset. Lukusa (2009) is not entirely clear on where the data were collected, although the use of palatal and lateral click phonemes in this source suggests that it represents the Botswana variety. These types of clicks occur in Yeyi of Botswana, but not that of Namibia. Given the homogenous nature of the data, we included this source in the second dataset, albeit without a definitive classification as either Botswana or Namibia.

Our non-Yeyi dataset was created to test the different affiliations for Yeyi that have been proposed in the literature, combined with languages spoken in the wider geographic area where Yeyi is spoken today. Given earlier suggestions that Yeyi might be of Eastern Bantu origin, we also included 38 Bantu languages of wider Eastern Bantu affiliation, despite being spoken far away from Yeyi. Finally, we added languages from the known subgroups of Bantu according to Grollemund et al. (2015), e.g., North-Western, Central-Western, West-Western (also called West-Coastal in Pacchiarotti et al., 2019, among others), and we rooted the tree with Duala (A24), a North-Western Bantu language. The complete list of languages used in our classification can be found in the Supplementary Materials.

For each of the languages in our dataset, we collected words for a list of 100 basic concepts. We focused on basic lexicon because it is more resistant to contact-induced change, and, therefore less likely to include borrowings which may obscure Yeyi's genealogical origins. This list is loosely based on the 92-word list used by Bastin et al. (1999) and the 100-word list used by Grollemund et al. (2015), but has been adapted in several ways.

Firstly, words that, at least in our dataset, appear to be prone to recurrent sound-symbolic associations were removed. For instance, the concept 'round' is frequently expressed by words containing rounded vowels and/or bilabial consonants. The occurrence of similar phonetic shapes for words with this meaning, therefore, could be either a sign of shared inheritance or of similar form-meaning associations which arose independently in different languages. Words which, during our initial data collection phases, appeared to be particularly prone to these kinds of sound-symbolic associations were therefore excluded from our word list.

Secondly, a number of concepts appeared, at least in our dataset, to be more prone to borrowing than expected. For instance, 'to swim', while found to be relatively borrowing-resistant cross-linguistically (Haspelmath & Tadmor, 2009), frequently occurs in the Bantu

languages in our dataset with phonemes which are described as only occurring in loanwords in the languages in question.

These problematic concepts were therefore replaced by concepts which, at least in the area under study, tend to be basic, such as 'elephant', 'snake', and 'house'.

The forms for each concept were sorted into cognate sets, using as much as possible available analyses on the sound changes that the languages in question have undergone, but using more intuitive judgments where necessary. Subsequently, each cognate set was recoded as a binary character tracking the presence or absence of each particular root-concept association.

Our lexical data, along with cognate judgements, are available in the Supplementary Materials. We constructed two lexical datasets, one keeping all Yeyi doculects separate (YeyiFull) and one with some Yeyi doculects merged, as explained above (YeyiMerged). We used lexedata v. 1.0.8 (Kaiping et al., 2022) to facilitate editing and annotation of the lexical datasets, as well as for the conversion of cognate sets into coded matrices. We analyzed the resulting binary matrices with MrBayes 3.2.7a (Ronquist et al., 2012). We used a simple binary model (also known as restriction site model), and default priors. All analyses were run for 20 million generations, and convergence was assessed with Tracer v.1.7.1 (Rambaut et al., 2018) and the standard deviation of split frequencies, which fell below 0.01 in all cases. We tested for the addition of gamma-distributed rate variation across characters using Bayes Factors. We estimated the marginal probabilities with and without rate variation using the steppingstone method, as implemented in MrBayes 3.2.7a. All analyses were performed using the resources of CIPRES Science Gateway (Miller et al., 2010).

4. Results

Our preliminary analyses with the full, unmerged dataset (YeyiFull) showed that all doculects of Yeyi form a highly supported clade and are very closely related (see file YeyiFull.pdf in online Supplementary Material). The results of the merged dataset (YeyiMerged) were essentially identical, and we proceeded to use this dataset for the rest of the analyses. The comparison with Bayes Factors indicated decisive support for the inclusion of gamma rate heterogeneity (the estimated natural log of the marginal likelihood with gamma rate heterogeneity was −17,939.35, while the corresponding value without gamma rate heterogeneity was −18,315.92, giving a lnBF of 376.57) (Kass & Raftery, 1995).

Figure 1 represents the linguistic phylogeny and the position of Yeyi.

Our phylogeny is rooted with Duala (A24), a North-Western Bantu language that is part of the primary clade in earlier Bantu phylogenies (Grollemund et al., 2015; Koile et al., 2022). Sister to Duala are one branch comprising the seven languages in our sample representative of two sub-branches in earlier phylogenies, i.e., West–Coastal/West–Western and Central–Western (Grollemund et al., 2015; Pacchiarotti et al., 2019; Koile et al., 2022), and another branch including the remainder of our sample. As it contains languages spoken outside of the Congo Forest, we call this third primary clade in our phylogeny "Savanna Bantu", a label used in earlier historical Bantu studies (cf. Section 1), though not necessarily covering exactly the same languages. Our Savanna Bantu corresponds to the superclade including Southwestern and Eastern Bantu languages in previous Bantu phylogenies (Grollemund et al., 2015; Koile et al., 2022), which do not use this label. However, its internal topology in our results is different from theirs, especially with regard to the position of zone D languages of the eastern DRC and Luba (L31a) of the south–central DRC, which are the first two branches to split off in our classification, but have a more embedded position in previous Bantu phylogenies (Grollemund et al., 2015; Koile et al., 2022). This is probably due to the very incomplete representation of their closest relatives in our sample.

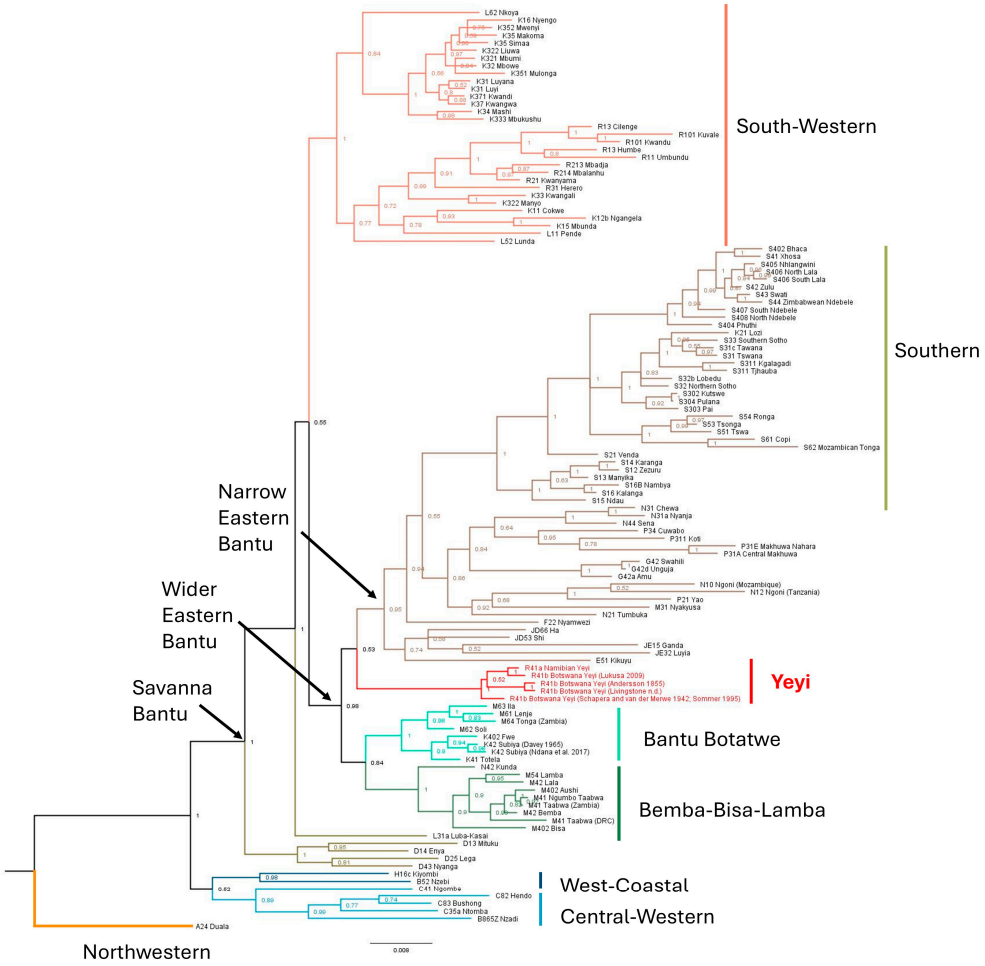


Figure 1. Bantu phylogeny highlighting the classification of Yeyi.

After the split off of zone D languages and Luba (L31a), the remainder of Savanna Bantu is a large clade splitting in two major subclades, which we label "Southwestern Bantu" and "Wider Eastern Bantu". The languages included in our Southwestern Bantu correspond largely to Southwestern Bantu in previous phylogenies (Grollemund et al., 2015; Koile et al., 2022), except for Luba (L31a), as discussed above. "Wider Eastern Bantu" comprises a first clade uniting two subclades, i.e., "Bantu Botatwe" and "Bemba–Bisa–Lamba". Bantu Botatwe in southwestern Zambia has already been recognized as a distinct subgroup on the basis of both lexicon and diachronic phonology (Bostoen, 2009; de Luna, 2010). The Bemba–Bisa–Lamba subclade unites languages from northeastern Zambia belonging to Guthrie's referential M40 (Bemba) and M50 (Bisa-Lamba) groups. The second major clade within Wider Eastern Bantu comprises a subclade uniting all Yeyi doculects in our sample and a sister one which we call "Narrow Eastern Bantu" and clusters the remainder of Bantu languages spoken in eastern and southern Africa in our sample. All of these have also previously been classified as part of Eastern Bantu (Grollemund et al., 2015; Koile et al., 2022). Our Narrow Eastern Bantu also includes the Southern Bantu languages of southeastern Africa, which, as in previous phylogenies, are a low-level subgroup of Eastern Bantu (Gunnink et al., 2023).

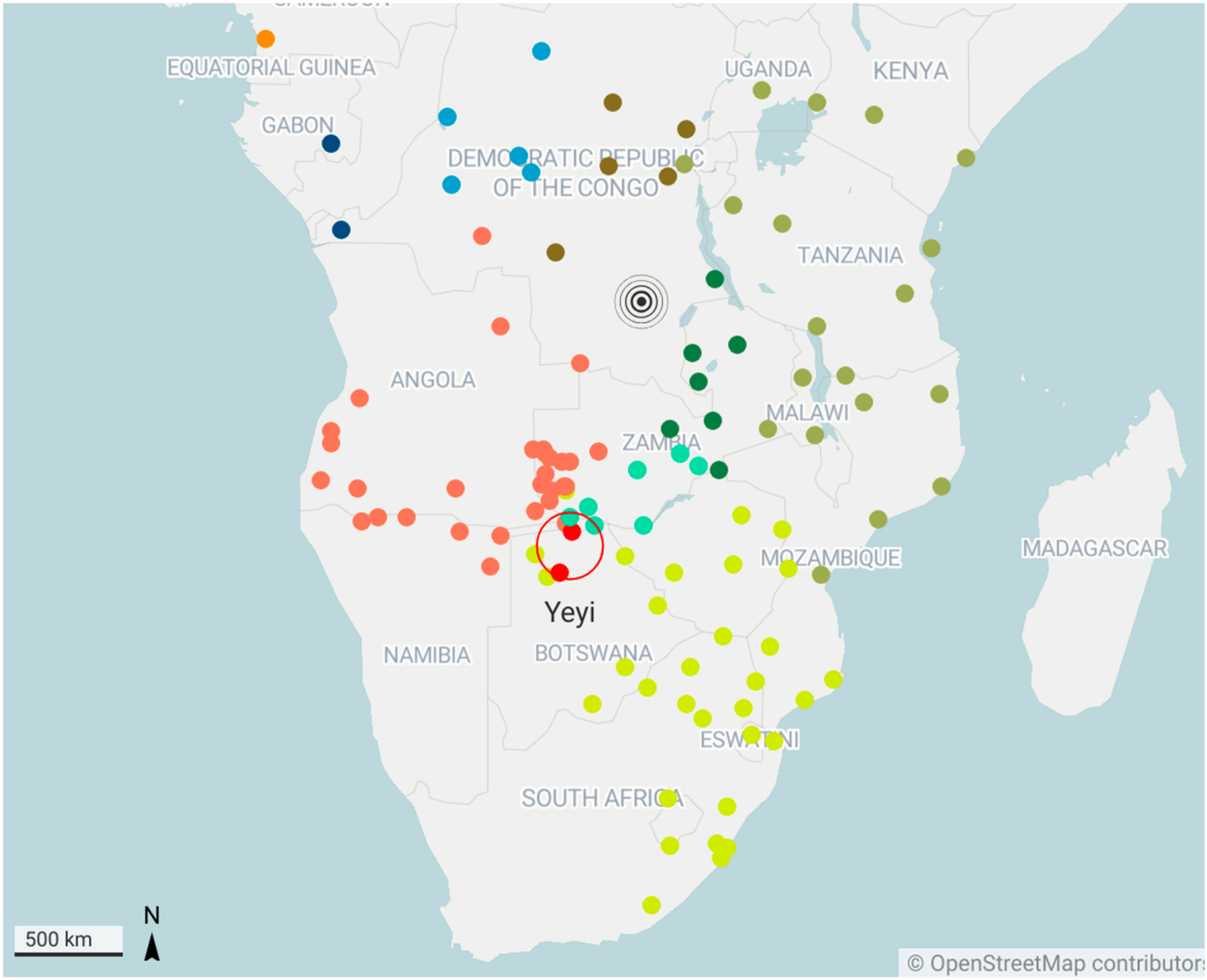
All Yeyi doculects in our study cluster closely together in a well-supported clade. As for its internal classification, this Yeyi clade does not show a clear split between Yeyi doculects of Botswana and Namibia. Namibian Yeyi appears most closely related to the variety of Yeyi documented by Lukusa (2009), which is not explicit about the origin of the data. The differentiation between Yeyi as spoken in Botswana and Namibia may, therefore,

be seen in other domains of the language than the lexicon. As for its external classification, our phylogeny offers clear support for the classification of Yeyi as part of Wider Eastern Bantu. Lexical innovations that define Wider Eastern Bantu, and are also shared with Yeyi, include the root *dòpà 'blood' (BLR 1144) (Bastin et al., 2002) and the *ini* root for 'liver' that has reflexes in many Wider Eastern Bantu languages, e.g., *li-ini* in Manda (N11), *li-ini* in Lenje (M61), *ɪk-ɪnye* in Nyakyusa (M31), *ini* in Swahili (G42), and *ini* in Kikuyu (E51) (cf. Nurse & Hinnebusch, 1993; Samson & Schadeberg, 1994). A cognate form *l-ine* or *l-inye* is also attested in Yeyi.

Yeyi's closest relatives within Wider Eastern Bantu are the entirety of the languages which cluster as Narrow Eastern Bantu, and not the languages spoken in its immediate vicinity. The clade we label Southwestern Bantu, with languages spoken to the west and north of Yeyi, is well-supported in our phylogeny and excludes Yeyi. For example, lexical innovations defining Southwestern Bantu include *cʊdɪ 'liver' (BLR 4655), which has reflexes throughout Southwestern Bantu but not in Yeyi, or the loss of the reflex of *jínò 'tooth' (BLR 3472), which is maintained in Yeyi but lost throughout Southwestern Bantu.

Evidence for the inclusion of Yeyi in Bantu Botatwe is also not found. Bantu Botatwe languages cluster together as a well-supported clade, which belongs to a distinct Wider Eastern Bantu branch, and thus excludes Yeyi. One of the characteristic lexical innovations of Bantu Botatwe is the replacement of the inherited reflex of the root *tɪ́ 'tree' (BLR 2881) with a new root exclusively found in Bantu Botatwe, tentatively reconstructable as *camu on the basis of its reflexes such as *ci-ʃamu* in Fwe and *ci-samu* in Lenje (see also de Luna, 2008, 2010). Yeyi, however, does not share in this innovation, but has retained *mu-ti* 'tree' as a reflex of *tɪ́. Furthermore, there are sound changes that set Yeyi apart from Bantu Botatwe, such as the loss of nasals before voiceless consonants, which has affected Yeyi (Gowlett, 1997) but not Bantu Botatwe (Bostoen, 2009).

In other words, within Wider Eastern Bantu, Yeyi appears to be somewhat of a loner, having diverged from other Eastern Bantu languages relatively early. This is supported by the many lexical innovations in its basic vocabulary that Yeyi does not share with any other Bantu language, such as *in-goro* 'arm', *li-dzundzo* 'cloud', *hweta* 'speak', *shi-poro* 'mouth', *mashira* 'be sick', *mu-nyana* 'man', and *ma-shuta/ma-shita* 'milk'. These words are also not recognizable as Khoisan loanwords (although some Khoisan loans did enter Yeyi's basic vocabulary, see Table 1 below). The loner status of Yeyi is also reflected in its geographical isolation from its closest Narrow Eastern Bantu relatives, except for Tswana and Lozi (see Figure 2). However, Southern Bantu, to which Tswana and Lozi belong, probably emerged in southern Africa not earlier than about a millennium ago (Gunnink et al., 2023). The presence of Tswana in the Okavango delta, more specifically its Tawana variety, is even much more recent, i.e., the early nineteenth century CE (Tlou, 1985). Similarly, the Lozi language only emerged after the migration of Sotho-speaking groups to western Zambia in the nineteenth century CE (Gowlett, 1989) and is therefore also a relative newcomer. Furthermore, Yeyi is not part of the Southern Bantu subgroup, and its position within southern Africa can therefore not be taken to be part of the same northward expansion that brought Tswana and Lozi into the Botswana/Namibia/Zambia border region.



North-Western Bantu
Central-Western Bantu
Bemba-Bisa-Lamba
South-Western Bantu
Yeyi
Southern Bantu
West-Coastal Bantu
D10-20-40-L31
Bantu Botatwe
Narrow Eastern Bantu
Wider Eastern Bantu homeland

Figure 2. Geographic location and classification of languages included in this phylogeny.

Table 1. Basic lexemes in Yeyi of Khoe origin.

Yeyi	Khoe Etymology	Yeyi	Bantu Etymology
ldi!'owo 'belly' (Lukusa, 2009, p. 95)	Naro*!'oo* 'inside, stomach' (Visser, 2001, p. 74)	*uwura* 'belly' (Seidel, 2008, p. 383; Donnelly, 1991, p. 41; Gowlett, 1992, p. 162) *ora* 'belly' (Andersson, 1855, p. 96)	*dà 'abdomen, intestines' (BLR 773) (Bastin et al., 2002)
g!uli 'big' (Sommer, 1995, p. 406; Seidel, 2008, p. 142), *g!udi* 'big' (Schapera & van der Merwe, 1942, p. 5)	Khwe *gùrí* 'be big' (Kilian-Hatz, 2003, p. 49)	*kuru* 'big' (Sommer, 1995, p. 406)	*kód 'grow up' (BLR, 1997) (Bastin et al., 2002)
mu-tj'awu 'sand' (Seidel, 2008, p. 113), *ʃi-c'au* 'sand' (Sommer, 1995, p. 410)	Ts'ixa *ts'áó* 'sand' (Fehn, 2019)	*movo* 'sand' (Andersson, 1855, p. 99), *muvu* 'sand' (Sommer, 1995, p. 410)	*bú 'soil' (BLR 355) (Bastin et al., 2002)
mugǁekwa 'woman' (Chebanne et al., 2007, p. 58; Sommer, 1995, p. 405), *mogǁekwa* 'woman' (Schapera & van der Merwe, 1942, p. 5)	Khwe *ǁgɛ̀ɛ̀-khòè* 'woman' (Kilian-Hatz, 2003, p. 211)	*mukazi* 'woman' (Sommer, 1995, p. 405; Donnelly, 1991, p. 49; Seidel, 2008, p. 110), *mokaz* 'woman' (Andersson, 1855, p. 100)	*kádí 'woman' (BLR, 1674) (Bastin et al., 2002)
hwe 'white' (Lukusa, 2009, p. 34), *wù-hwé* 'white' (Sommer, 1995, p. 413), *hwe* 'white' (Seidel, 2008, p. 132), *chwe* 'white' (Schapera & van der Merwe, 1942, p. 78)	Khwe *xùɛ́ci* 'be bright white, be very white' (Kilian-Hatz, 2003, p. 147)	*tɯa* 'white' (Schapera & van der Merwe, 1942, p. 78)	*tʊb 'be white' (BLR, 4527) (Bastin et al., 2002)
lene 'small' (Chebanne et al., 2007, p. 54), *nleni* 'small' (Schapera & van der Merwe, 1942, p. 5), *nl'ene* 'small' (Sommer, 1995, p. 407; Seidel, 2008, p. 142)	Khwe *líní* 'to be thin' (Kilian-Hatz, 2003, p. 155)		
munlu 'leg' (Chebanne et al., 2007, p. 27), *monlu* 'leg' (Schapera & van der Merwe, 1942, p. 5), *mon'o* 'leg' (Andersson, 1855, p. 98), *munglu* 'lower leg' (Donnelly, 1991, p. 22)	Khwe *nlgúu* 'lower leg' (Kilian-Hatz, 2003, p. 168)		
leki 'one' (Schapera & van der Merwe, 1942, p. 5), *lhiki* 'one' (Chebanne et al., 2007, p. 28; Seidel, 2008, p. 143), *mo'keke* 'one' (Andersson, 1855, p. 100)	Khwe *lúí*, Ts'ixa *lúí*, Naro *lúí* 'one'		

However, there is a theoretical possibility that ongoing contact with both Tswana and Lozi may have influenced the apparent Narrow Eastern Bantu affiliation of Yeyi. If Yeyi had borrowed extensively from Lozi and Tswana, and these similarities would be coded as cognates in our dataset, this could have created the false impression that Yeyi has inherited substantial Narrow Eastern Bantu material. However, while the Tswana and Lozi influence on Yeyi is undeniable (Seidel, 2009), it is not identifiable in Yeyi's basic vocabulary. Yeyi shares 48 cognates with either Lozi or Tswana, but almost all of these have a corresponding Proto-Bantu reconstruction and are more widely distributed than Narrow Eastern Bantu. Furthermore, there are marked phonological differences between these cognates in Yeyi and in Tswana and Lozi, which result from the very different historical phonological development of these languages. For instance, Tswana, Lozi and the other languages of

the Sotho-Tswana cluster (S30) have not undergone Bantu Spirantization (Creissels, 1999, 2007), the change of stops to fricatives under the influence of a reconstructed high vowel (Schadeberg, 1994/1995; Janson, 2007; Bostoen, 2008). This widespread sound change did affect Yeyi (Gowlett, 1992). The effects of Bantu Spirantization clearly show that, even when Yeyi forms are historically cognate with forms attested in Lozi and Tswana, Yeyi did not borrow these forms from Tswana or Lozi, in which case no spirantization would be attested. Rather, Yeyi forms consistently show spirantization, which indicates that they are inherited forms rather than borrowings, e.g., *jédì 'moonlight' (BLR 3283) (Bastin et al., 2002) > Yeyi *ukw-ezi* 'moon' vs. Lozi *kweli* 'moon', Tswana *kgwedi* 'moon'.

Another remarkable sound change is the palatalization of bilabial consonants when followed by a rounded vowel, which has affected the Sotho languages Lozi and Tswana (Kotzé & Zerbian, 2008), but not Yeyi, e.g., *bóà 'dog' (BLR 282) (Bastin et al., 2002) > Yeyi *om-bwa* 'dog' vs. Tswana *n-tʃa* 'dog', Lozi *n-ja* 'dog; *bòè 'stone' (BLR 285) (Bastin et al., 2002) > Yeyi *li-we* 'stone' vs. Lozi *li-cwe* 'stone'.

The phonological differences between Yeyi forms and their cognates in Lozi and Tswana therefore clearly show that Lozi or Tswana influence has not impacted Yeyi's basic vocabulary, and therefore the analysis of the Yeyi cluster as a sister branch to Narrow Easter Bantu is unlikely to be an artefact of extensive contact-induced change.

Another possibility is that extensive contact-induced changes from Khoisan languages would have contributed to the relatively isolated status of Yeyi within Eastern Bantu. Yeyi has seen extensive influence from surrounding Khoisan languages in its phonology, morphology, and lexicon (Sommer & Voßen, 1992; Gunnink et al., 2015; Gunnink, 2022). If inherited words in the basic lexicon were replaced by loans, this would have obscured the precise Bantu origins of Yeyi. However, horizontal Khoisan influence was not strong enough to obscure Yeyi's Bantu origins. In the 100-word list used for our phylogeny, eight Yeyi words have an identifiable Khoisan origin. These words all come from languages of the Khoe branch of the Khoe–Kwadi family and are listed in Table 1. Five of these concepts are also represented by a synonym in Yeyi with a clear Bantu origin, as shown by its corresponding Bantu lexical reconstruction, either in the same doculect as the Khoisan loan or in a different doculect. So, even though Khoisan influence on Yeyi was strong enough to interfere with its basic vocabulary, most of its basic vocabulary of Bantu origin is also still maintained. Furthermore, as mentioned above, Yeyi has acquired many lexical innovations in its basic vocabulary which do not have an identifiable Khoisan origin.

Having shown that language contact did not affect Yeyi's basic vocabulary strongly enough to account for its position as an evolutionary "loner" within Wider Eastern Bantu, we therefore consider these results as robust and now turn to the interpretation of this linguistic classification in terms of historical scenarios.

5. Discussion

Firstly, the analysis of Yeyi as a member of the Wider Eastern Bantu branch allows us to identify the homeland of the Wider Eastern Bantu clade as the most likely area from where ancestral Yeyi speakers would have migrated into their current habitat. This Wider Eastern Bantu homeland would be situated somewhere in the Katanga region of what is now the southern DRC (see map in Figure 2), which is where the center of highest diversity within the Wider Eastern Bantu sub-clade including Yeyi is situated. What we call Wider Eastern Bantu corresponds largely to branch 11 in Grollemund et al.'s (2015, 2023) phylogeny, whose homeland they situate to the east of Lake Tanganyika, from where speakers of "Sabi-Botatwe" languages, similar to our Bemba–Bisa–Lamba and Bantu Botatwe branches, expanded in a southwestern direction. This also corresponds to the "Osthochland-Gruppe" in the lexicostatistical classification of Heine et al. (1977), though

not exactly, as the latter also includes Luba of Kasai (DRC) and the zone D languages of eastern DRC. Heine et al. (1977, p. 65) situate the homeland of this subgroup, from where most of eastern and southern Africa was colonized, between the Kasai and Lualaba Rivers, i.e., an area roughly corresponding to the territories of the modern Kasai, Kasai-Central, Kasai-Oriental, Sankuru, and Lomami provinces of the DRC (formerly Kasai-Occidental and Kasai-Oriental). This is slightly to the north of Katanga, where we situate the Wider Eastern Bantu homeland, and immediately to the east of where Pacchiarotti et al. (2019) situate the West–Coastal Bantu homeland. Southwestern Bantu's center of origin would, then, have been somewhere south of the West–Coastal Bantu homeland and west of our Wider Eastern Bantu homeland.

According to Heine et al. (1977, p. 65), the fragmentation of the "*Osthochland-Nukleus*" some two millennia ago must have been explosive and led in a very short time to the "flooding" ("*überfluten*") of the vast area between the Laikipia Plateau in Kenya and the Karoo in South Africa by the different subgroups emerging from that "punctuational burst" (cf. Atkinson et al., 2008). One of these early offshoots may have been the most recent common ancestor of Yeyi and its extinct closest relatives.

However, the exact timing of the Yeyi migration into southern Africa is difficult to establish. On the one hand, the lack of genealogical affiliation between Yeyi and its geographic neighbors of the Bantu Botatwe, Bemba–Bisa–Lamba, or Southwestern Bantu subgroups clearly supports a scenario of a separate migration. If this migration happened relatively recently, we would expect Yeyi to have clearly identifiable sister languages within Wider Eastern Bantu as a whole. Other cases of such recent long-distance migrations of speech communities originated in nineteenth-century CE South Africa, from where Sotho-speaking groups ancestral to Lozi speakers migrated to Western Zambia, and Nguni-speaking groups ancestral to Ndebele migrated to western Zimbabwe. The modern Lozi language, however, still shows a very clear genealogical affiliation to the Sotho language cluster of Southern Africa, and Zimbabwean Ndebele is closely embedded in the Nguni cluster (see Figure 1). If the Yeyi language would also result from such a relatively recent, relatively long-distance migration, it would be expected to still show an identifiable affiliation to specific Eastern Bantu languages, which is not the case.

On the other hand, if the ancestors of modern-day Yeyi speakers entered southern Africa through a relatively early migration, this would not only explain its lack of close affiliation to its neighbors, but would also account for the accumulation of linguistic differences between Yeyi and its Narrow Eastern Bantu relatives.

The lack of close genealogical relatives to Yeyi, either within the general geographic vicinity or within Wider Eastern Bantu, also raises the question of how the Yeyi cluster has apparently resisted diversification. One possibility is that the Yeyi speech communities were surrounded by other ethnolinguistic groups, which prevented them from migrating and loosening their social and linguistic bonds. Another possibility is that the Yeyi cluster did once belong to a larger and more diverse group of languages, but that its putative earlier sister languages disappeared through language shift. This would be especially likely if the Yeyi resulted from a relatively early migration, where surrounding Bantu languages arrived as part of a separate and more recent expansion. Among the possible targets of shift are Bantu Botatwe languages, which border on Yeyi in the northeast, and languages of the Luyana cluster (K30), spoken to the north of Yeyi. Both groups are also among the languages that have previously been proposed as Yeyi's possible relatives (see Section 2). While a close genealogical relationship between Yeyi and either Luyana or Bantu Botatwe is not supported in our phylogeny, it is possible that the linguistic similarities that exist are rather the result of a substrate of a now extinct Bantu language closely related to Yeyi.

Yeyi as the remnant of an early migration phase, partially replaced by later migrants speaking different Bantu languages, is similar to other scenarios of spread-over-spread proposed in different parts of the Bantu-speaking area. In the Congo rainforest, two separate population expansions were separated by a population collapse in the first millennium CE, suggesting that many of the modern-day Bantu languages spoken in the area derive from this second population expansion, with many of the languages spoken by the first Bantu migrants having gone extinct (Seidensticker et al., 2021; Bostoen et al., in press). In southern Africa, based on mismatches between archaeology and linguistic phylogeny, Gunnink et al. (2023) propose a spread-over-spread scenario for the Shona language community. Judging from their ceramic traditions (Huffman, 2007), the ancestors of modern-day Shona speakers would have arrived in southern Africa in the first millennium CE as part of the earliest wave of Bantu Expansion. However, judging from the phylogeny of Southern Bantu, of which Shona is integral part (see also Figure 1), and the fact that the ceramic traditions of other Southern Bantu speech communities do not extend back beyond 1000 CE (Schoeman, 2013; Mitchell, 2024), Shona ancestors likely shifted to ancestral Southern Bantu languages when these were introduced as part of a later wave of Bantu Expansion in the second millennium CE. In the same vein, comparing archaeological and linguistic data, Phillipson (1977) has already suggested that modern-day Eastern Bantu speakers more generally trace their origin to the Late Iron Age rather than the Early Iron Age.

The case of Yeyi could represent another instance of an early Bantu migration, possibly followed by more recent migrations of communities speaking different Bantu languages. However, unlike the ancestors of modern-day Shona speakers, the ancestors of modern-day Yeyi speakers did not shift to more recently introduced Bantu languages of newcomers such as Southern Bantu or to expanding languages of Bantu-speaking groups settled earlier on, such as Bantu Botatwe or Luyana. Yeyi forebears stuck to their own ancestral language, while neighboring communities speaking closely related languages or possibly even part of the ancestral Yeyi speech community itself did shift and were gradually assimilated into those language communities. As a result, Yeyi ended up as a clade of its own within the Wider Eastern Bantu sub-branch of the Bantu family tree's major Savanna Bantu branch.

This hypothesis of Yeyi as the sole survivor of an early phase of Bantu Expansion in the wider vicinity of the Okavango Delta needs further testing, on the basis of both other types of historical linguistic evidence and bodies of evidence from other disciplines, such as archaeology and population genetics. For instance, future research into the material culture of modern or historically attested Yeyi communities, such as ceramic traditions, could shed further light on their historical development, continuity and migration. Furthermore, in order to test the hypothesis that Yeyi's putative earlier sister languages disappeared as the result of language shift, a potential Yeyi-like substrate could be investigated in Bantu languages spoken to the northeast of modern-day Yeyi communities.

6. Conclusions

In this paper, we have investigated the genealogical relationships of Yeyi, a Bantu language of previously unknown subclassification. Using a lexicon-based Bayesian phylogenetic classification, we have provided clear evidence that Yeyi is part of the Wider Eastern Bantu branch within Savanna Bantu but forms a single clade that is sister to the entirety of Narrow Eastern Bantu. Yeyi does not share any closer genealogical relationship to the Wider Eastern Bantu and Southwestern Bantu languages in its close vicinity. The status of Yeyi as an evolutionary loner within Eastern Bantu suggests a long period of separate development, possibly as the result of a relatively early migration of ancestral Yeyi speakers. The Yeyi language as it survives today would be the last remaining member of a presumably larger clade of languages, whose members disappeared, possibly as the

result of language shift to Bantu languages spoken by more recent migrants or Bantu languages of older expanding groups. More research is needed to further test this hypothesis, by investigating a possible Yeyi-like substrate in surrounding Bantu languages and by interdisciplinary triangulation.

Supplementary Materials: Supplementary materials for this paper, including all lexical data, cognate judgments, input files for phylogenetic analyses and all results files, are available at https://osf.io/tksrh/?view_only=e775d8fdb0d640589bcbd71e16c2e7e8.

Author Contributions: Conceptualization, methodology, writing, H.G, N.C.-P. and K.B.; investigation, H.G. and K.B.; data curation H.G. and N.C.-P.; formal analysis, N.C.-P.; visualization, H.G. and N.C.-P. All authors have read and agreed to the published version of the manuscript.

Funding: Hilde Gunnink's research was funded by a grant by the Research Foundation Flanders (FWO 12P8423N).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available in the Supplementary Materials (https://osf.io/tksrh/?view_only=e775d8fdb0d640589bcbd71e16c2e7e8).

Conflicts of Interest: The authors declare no conflict of interest.

References

Andersson, C. J. (1855). Explorations in South Africa, with route from Walfish Bay to Lake Ngami, and ascent of the Tiogé river. *Journal of the Royal Geographical Society of London*, *25*, 79–107. [CrossRef]

Atkinson, Q. D., Meade, A., Venditti, C., & Greenhill, S. J. (2008). Languages evolve in punctuational bursts. *Science*, *319*(5863), 588. [CrossRef] [PubMed]

Bastin, Y., Coupez, A., & Mann, M. (1999). *Continuity and divergence in the Bantu languages: Perspectives from a lexicostatistic study*. Royal Museum for Central Africa.

Bastin, Y., Coupez, A., Mumba, E., & Schadeberg, T. C. (2002). *Bantu lexical reconstructions 3*. Royal Museum for Central Africa. Available online: https://www.africamuseum.be/en/research/discover/human_sciences/culture_society/blr (accessed on 21 June 2021).

Bostoen, K. (2008). Bantu spirantization: Morphologization, lexicalization and historical classification. *Diachronica*, *25*, 299–356. [CrossRef]

Bostoen, K. (2009). Shanjo and Fwe as part of Bantu Botatwe: A diachronic phonological approach. In O. Akinloye, & L. Moshi (Eds.), *Selected proceedings of the 39th annual conference of African linguistics* (pp. 110–130). Cascadilla Proceedings Project.

Bostoen, K., Pacchiarotti, S., & Hubau, W. (in press). À la recherche des sentiers perdus du passé en forêt. Réflexions sur l'effondrement de la population en Afrique centrale il y a 1500 ans. *Proceedings of the Royal Academy for Overseas Sciences*.

Bostoen, K., & Van de Velde, M. (2019). Introduction. In M. Van de Velde, K. Bostoen, D. Nurse, & G. Philippson (Eds.), *The Bantu languages* (2nd ed., pp. 1–13). Routledge.

Chebanne, A. M., Rodewald, M., & Nyati-Ramahobo, L. (2007). *Mutjango wu Shiyeyi (The Shiyeyi writing system)*. The Botswana Society.

Creissels, D. (1999). Remarks on the sound correspondences between Proto-Bantu and Tswana (S.31), with particular attention to problems involving *j (or *y), *I and sequences *NC. In J.-M. Hombert, & L. M. Hyman (Eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (pp. 297–334). CSLI Publications.

Creissels, D. (2007). L'influence des voyelles sur les évolutions des consonnes en tswana (S31). *Africana Linguistica*, *13*, 33–52. [CrossRef]

de Luna, K. M. (2008). *Collecting food, cultivating persons: Wild resource use in central African political culture, c. 1000 B.C.E. to c. 1900 C.E* [Ph.D. dissertation, Northwestern University].

de Luna, K. M. (2010). Classifying Botatwe: M60 languages and the settlement chronology of south central Africa. *Africana Linguistica*, *16*, 65–96. [CrossRef]

Donnelly, S. S. (1991). *Phonology and morphology of the noun in Yeeyi* [Bachelor's thesis, University of Cape Town].

Ehret, C. (1999). Subclassifying Bantu: The evidence of stem morpheme innovations. In J.-M. Hombert, & L. M. Hyman (Eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (pp. 43–147). CSLI Publications.

Ehret, C. (2001). Bantu expansions: Re-envisioning a central problem of early African history. *The International Journal of African Historical Studies*, *34*(1), 5–41. [CrossRef]

Fehn, A.-M. (2019). *A dictionary of Ts'ixa: A Khoe language of northern Botswana* [Unpublished manuscript].
Fulop, S. A., Ladefoged, P., Liu, F., & Vossen, R. (2003). Yeyi clicks: Acoustic description and analysis. *Phonetica*, *60*, 231–260. [CrossRef]
Gowlett, D. F. (1989). The parentage and development of Lozi. *Journal of African Languages and Linguistics*, *11*, 127–149. [CrossRef]
Gowlett, D. F. (1992). Yeyi reflexes of proto-Bantu. In D. F. Gowlett (Ed.), *African linguistic contributions: Presented in honour of Ernst Westphal* (pp. 122–188). Via Afrika.
Gowlett, D. F. (1997). Aspects of Yeyi diachronic phonology. In W. H. G. Haacke, & E. E. Elderkin (Eds.), *Namibian languages: Reports and papers* (pp. 235–263). Rüdiger Köppe.
Grollemund, R., Branford, S., Bostoen, K., Meade, A., Venditti, C., & Pagel, M. (2015). Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences*, *112*(43), 13296–13301. [CrossRef]
Grollemund, R., Schoenbrun, D., & Vansina, J. (2023). Moving histories: Bantu language expansions, eclectic economies, and mobilities. *The Journal of African History*, *64*(1), 13–37. [CrossRef]
Gunnink, H. (2022). Morphological Khoisan influence in the Southern African Bantu language Yeyi. *Journal of African Languages and Linguistics*, *43*(1), 3–35. [CrossRef]
Gunnink, H., Chousou-Polydouri, N., & Bostoen, K. (2023). Divergence and contact in Southern Bantu language and population history: A new phylogeny in cross-disciplinary perspective. *Language Dynamics and Change*, *13*(1), 74–131. [CrossRef]
Gunnink, H., Sands, B., Pakendorf, B., & Bostoen, K. (2015). Prehistoric language contact in the Kavango-Zambezi transfrontier area: Khoisan influence on southwestern Bantu languages. *Journal of African Languages and Linguistics*, *36*(2), 193–232. [CrossRef]
Guthrie, M. (1948). *The classification of the Bantu languages*. Published for the International African Institute by the Oxford University Press.
Haspelmath, M., & Tadmor, U. (2009). *World loanword database.* Max Planck Institute for Evolutionary Anthropology. Available online: http://wold.clld.org (accessed on 24 September 2020).
Heine, B., Hoff, H., & Vossen, R. (1977). Neuere Ergebnisse zur Territorialgeschichte der Bantu. In *Zur Sprachgeschichte und Ethnohistorie in Afrika* (pp. 57–72). Dietrich Reimer.
Huffman, T. N. (2007). *Handbook to the Iron Age. The archaeology of pre-colonial farming societies in Southern Africa*. University of KwaZulu-Natal Press.
Janson, T. (2007). Bantu spirantisation as an areal change. *Africana Linguistica*, *13*, 79–116. [CrossRef]
Johnston, H. H. (1919). *A comparative study of the Bantu and semi-Bantu languages*. Oxford University Press.
Kaiping, G. A., Steiger, M. S., & Chousou-Polydouri, N. (2022). Lexedata: A toolbox to edit CLDF lexicaldatasets. *Journal of Open Source Software*, *7*(72), 4140. [CrossRef]
Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795. [CrossRef]
Kilian-Hatz, C. (2003). *Khwe dictionary*. Rüdiger Köppe.
Koile, E., Greenhill, S. J., Blasi, D. E., Bouckaert, R., & Gray, R. D. (2022). Phylogeographic analysis of the Bantu language expansion supports a rainforest route. *Proceedings of the National Academy of Sciences*, *119*(32), e2112853119. [CrossRef]
Kotzé, A. E., & Zerbian, S. (2008). On the trigger of palatalization in the Sotho languages. *Journal of African Languages and Linguistics*, *29*, 1–28. [CrossRef]
Larson, T. J. (1989). The Bayeyi of Ngamiland. *Botswana Notes and Records*, *21*, 23–42.
Lisimba, M. (1982). *A Luyana dialectology* [Ph.D. thesis, University of Wisconsin].
Livingstone, D. (1851). *Comparative vocabulary of eight Bantu languages: Bakhoba (Bayeyi), Bashubea, Balojazi, Ba(ma)ponda, Barotse, Batoka, Banyenko, Betshwana*. Grey Collection, South African Public Library.
Lukusa, S. T. M. (2009). *Shiyeyi—English dictionary*. Lincom.
Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). *Creating the CIPRES science gateway for inference of large phylogenetic trees*. 2010 Gateway Computing Environments Workshop (GCE) (pp. 1–8), New Orleans, LA, USA. [CrossRef]
Mitchell, P. (2024). *The archaeology of Southern Africa* (2nd ed.). Cambridge University Press.
Nurse, D., & Hinnebusch, T. J. (1993). *Swahili and Sabaki: A linguistic history*. University of California Press.
Nyati-Ramahobo, L. (2002). From a phone call to the high court: Wayeyi visibility and the Kamanakao association's campaign for linguistic and cultural rights in Botswana. *Journal of Southern African Studies*, *28*(4), 685–709. [CrossRef]
Nyati-Saleshando, L. (2019). Shiyeyi language planning in Botswana: International connections and local imperatives. *Journal of Language, Identity & Education*, *18*(3), 176–189.
Pacchiarotti, S., & Bostoen, K. (2020). The proto-west-coastal Bantu velar merger. *Africana Linguistica*, *26*, 139–195.
Pacchiarotti, S., Chousou-Polydouri, N., & Bostoen, K. (2019). Untangling the West-Coastal Bantu mess: Identification, geography and phylogeny of the Bantu B50-80 languages. *Africana Linguistica*, *25*, 155–229.
Pakendorf, B., Gunnink, H., Sands, B., & Bostoen, K. (2017). Prehistoric Bantu-Khoisan language contact: A cross-disciplinary approach. *Language Dynamics and Change*, *7*(1), 1–46. [CrossRef]
Phillipson, D. W. (1977). The spread of the Bantu language. *Scientific American*, *236*(4), 106–115. [CrossRef]

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Systematic Biology*, *67*(5), 901–904. [CrossRef]

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology*, *61*(3), 539–542. [CrossRef]

Samson, R., & Schadeberg, T. C. (1994). Kiinimacho cha mahali: Kiambishi tamati cha mahalini. *Afrikanistische Arbeitspapiere*, *37*, 127–138.

Sands, B., & Güldemann, T. (2009). What click languages can and can't tell us about language origins. In R. P. Botha, & C. Knight (Eds.), *The cradle of language* (pp. 204–218). Oxford University Press.

Schadeberg, T. C. (1994/1995). Spirantization and the 7-to-5 vowel merger in Bantu. *Belgian Journal of Linguistics*, *9*, 71–84. [CrossRef]

Schapera, I., & van der Merwe, D. F. (1942). *Notes on the noun-classes of some Bantu languages of Ngamiland (Yeei, Subia, Gova, and Gcereku)*. School of African Studies, University of Cape Town.

Schoeman, A. (2013). Southern African Late Farming Communities. In P. Mitchell, & P. Lane (Eds.), *Oxford handbook of African archaeology* (pp. 928–941). Oxford University Press.

Seidel, F. (2005). The Bantu languages of the Eastern Caprivi: A dialectometrical analysis and its historical and sociolinguistic implications. *South African Journal of African Languages*, *4*, 207–242. [CrossRef]

Seidel, F. (2008). *A grammar of Yeyi*. Rüdiger Köppe.

Seidel, F. (2009). Layered language genesis in the "catch basin" of the Linyanti and Okavango swamps: The case of Yeyi. *Sprache und Geschichte in Afrika*, *20*, 231–264.

Seidensticker, D., Hubau, W., Verschuren, D., Fortes-Lima, C., de Maret, P., Schlebusch, C. M., & Bostoen, K. (2021). Population collapse in Congo rainforest from 400 CE urges reassessment of the Bantu expansion. *Science Advances*, *7*(7), eabd8352. [CrossRef] [PubMed]

Sommer, G. (1995). *Ethnographie des sprachwechsels: Sozialer wandel und sprachverhalten bei den Yeyi (Botswana)*. Rüdiger Köppe.

Sommer, G. (2017). Variation in Shiyeyi (Botswana/Namibia). In R. Kramer, & R. Kießling (Eds.), *Mechthildian approaches to Afrikanistik: Advances in language based research on Africa. Festschrift für Mechthild Reh* (pp. 321–337). Rüdiger Köppe.

Sommer, G., & Voßen, R. (1992). Schnalzwörter im Yei (R.41). *Afrika und Übersee*, *75*, 1–42.

Sommer, G., & Vossen, R. (1995). Linguistic variation in Siyeyi. In A. Traill, R. Vossen, & M. Biesele (Eds.), *The complete linguist. Papers in memory of Patrick J. Dickens* (pp. 407–479). Rüdiger Köppe.

Tlou, T. (1985). *A history of Ngamiland. 1750 to 1906. The formation of an African state*. Macmillan.

Visser, H. (2001). *Naro dictionary: Naro-English, English-Naro*. Naro Language Project.

Vossen, R. (1997). What clicks sounds got to do in Bantu. In B. Smieja, & M. Tasch (Eds.), *Human contact through language and linguistics* (pp. 353–366). Peter Lang.

Westphal, E. O. J. (1964). *An example of complex language contacts in Ngamiland, BP*. Colloque sur le Multilingualisme/Symposium on Multilingualism: The Second Meeting of the Inter-African Committee on Linguistics, Brazzaville, 16–21 August 1962 (pp. 205–210), Commission pour Coopération Technique en Afrique (CCTA); Conseil Scientifique pour l'Afrique (CSA).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.