Erratic velars in West-Coastal Bantu

Explaining irregular sound change in Central Africa

Sara Pacchiarotti and Koen Bostoen
Ghent University

In this article, we present the first quantitative study of what we call multiple unconditioned reflexes (MUR) in Bantu, more specifically of Proto-Bantu velar stops *k and *g in the West-Coastal Bantu (WCB) branch of the Bantu language family. MUR, also known as "doubles reflexes" in Bantu studies, represent a situation where one and the same proto-sound has two or more reflexes in a given language which cannot be accounted for by phonological conditioning and/or lexical borrowing. This diachronic irregularity has been explained in Bantu historical linguistics, and Niger-Congo studies more broadly, by reconstructing either an additional series of consonants (phonemic merger) or a latent conditioning that went lost (phonemic split). We show that MUR should not be explained, but rather taken as an indicator of the same pervasive irregularity of sound change reported in other parts of the world that are highly multilingual and lack a neat overlap between distinct languages and communities. Along with widespread multilingualism, we assess lexical diffusion, substrate influence, and spread-over-spread events in Bantu language history as complementary explanations for the rise of MUR in WCB.

Keywords: multiple unconditioned reflexes, irregular sound change, West-Coastal Bantu, substrate interference, small-scale multilingualism, spread-over-spread events, lexical diffusion

1. Introduction

Since its emergence in Indo-European studies in the second half of the 19th century, the Comparative Method "has been the key tool for investigating linguistic prehistory" (Weiss 2015: 127), and this in spite of many challenges. One main threat to the Comparative Method has been the calling into question of its underlying assumption, i.e., the so-called "Neogrammarian hypothesis" of exceptionless

https://doi.org/10.1075/jhl.20054.bos
Journal of Historical Linguistics ISSN 2210-2116 | E-ISSN 2210-2124

regularity in phonological change, upon the recurrent observation of irregular sound shifts across the world's languages (cf. Wang 1969; Labov 1981; Durie & Ross 1996 and chapters therein; Harrison 2003). Another serious challenge, closely linked to the previous, is the disputation of the Comparative Method's universality. It has been questioned whether its applicability is not too strongly dependent on the specific historical context of Indo-European, not only in terms of language evolution (i.e., dispersal and divergence), but also in terms of documentation of both language and history (i.e., presence of old written records). Doubts have been raised on the successfulness of the Comparative Method in parts of the world where the current-day linguistic landscape does not primarily result from the differentiation of a common proto-language and where no ancient texts are available to reconstruct language evolution and the more general historical backdrop against which this happened (cf. Durie & Ross 1996 and chapters therein; Weiss 2015: 136–139). In this respect, Sub-Saharan Africa certainly offers fertile ground for testing the Comparative Method's suitability as a tool to examine (linguistic) prehistory.

One of the Comparative Method's success stories in Africa is no doubt Bantu, especially for the reconstruction of Proto-Bantu (PB) (cf. Meinhof 1899, 1906; Greenberg 1948; Meeussen 1967, 1969; Schadeberg 2002, 2003; Bostoen & Bastin 2016; Bostoen 2019). Bantu is a relatively young "spread zone" (Güldemann 2011), namely an estimated 4000–5000 years, which is the outcome of migration from a homeland on the border between Nigeria and Cameroon towards Eastern and Southern Africa (Bostoen 2018). Its rapid spread over a huge area no doubt favored the successful reconstruction of ancestral stages thanks to close geneaological relatedness, high identifiability of cognates, and well-documented internal variation (cf. Bostoen 2019: 308–309). Then again, when it comes to internal Bantu classification, the Comparative Method has proven to be less effective (cf. Nurse & Philippson 2003; Schadeberg 2003; Philippson & Grollemund 2019). Shared innovations turn out to be rarely diagnostic for genealogical subgrouping, because isoglosses have the tendency to crosscut rather than to overlap. This situation, likely the most common worldwide, has been attributed to longstanding and intensive Bantu-internal contact and intense multilingualism (cf. Schadeberg 2003: 158–159). One specific indicator of prehistoric contact between Bantu languages and/or shifts from one Bantu language to another would be what Möhlig (1977, 1981a) calls "hybrid sound shifts" as opposed to "linear sound shifts". Irregularity in sound correspondences would emerge through the hybridization of regular inherited sound systems, for instance in the case of language shift (*Sprachübernahme*), or the adoption of pronunciation habits (*Aussprachegewohnheiten*) of other speech communities (Möhlig 1981b: 88). According to this stratification model (Möhlig 1977, 1979, 1981b), the difficulty in establishing genealogical

Bantu subgroups through shared phonological innovations is simply due to the fact that the sound systems of present-day Bantu languages did not regularly evolve from a single ancestral language, but are actually intricate composites of superimposed historical layers (see also Nurse & Masele 2003). Briefly put, as in other parts of the world, such as New Caledonia (Grace 1996) and Papua New Guinea (Ross 1996), the role of intensive language contact in the rise of irregular phonological change as a structural and inherent feature of Bantu sound systems has been recognized since at least the late 1970s.

Alongside this recognition of irregularity in sound change, from the early 1980s there has been much debate in Bantu/Niger-Congo scholarship about so-called "double reflexes", a controversy which essentially fails to admit the possibility of irregular sound change. In several languages, especially in the northwestern part of the Bantu domain (though also in South-Western and Eastern Bantu, cf. Kanyamibwa 1982 and Botne 1992 respectively), certain PB consonants have two or more recurrent correspondences that are divergent but without a clear phonological conditioning. Broadly speaking, this synchronic situation has been accounted for by two mutually exclusive diachronic explanations: phonemic merger vs. phonemic split (cf. Janssens 1993: 1–18 for a detailed overview of previous scholarship on Bantu double reflexes). In the merger scenario, the sound system of Proto-Bantu would have had more contrasts than those reconstructed. The "double reflexes" would then be archaisms reflecting those phonemic oppositions. In languages not having "double reflexes", the ancient contrasts would have merged and got lost. John Stewart, the most ardent advocate of the merger scenario, argued that the contrast between so-called fortis and lenis consonants, which he had reconstructed for Proto-Volta-Congo (Stewart 1973), was not only retained in Kwa languages but also in several northwestern Bantu languages and by extension also in Proto-Bantu (Van Leynseele & Stewart 1980; Stewart 1989, 1993; see also Gerhardt 1986; Hedinger 1987; Bancel 1988 for related claims). In the split scenario, scholars assume that a reconstructed Proto-Bantu consonant having more than one phonologically unconditioned reflex in a given language underwent a divergent evolution through a conditioning that subsequently disappeared. Different scholars have proposed different conditioning factors, such as the following vowel (short vs. long) (Guthrie 1967: 58, but refuted, amongst others, by Nsuka-Nkutsi 1980; Van Leynseele & Stewart 1980; Janssens 1986) or the preceding prefix (e.g., Bachmann 1989; Miehe 1989; Blanchon 1991; Janssens 1991, 1993). In the end, Stewart would have discarded the merger scenario in favor of the split hypothesis, at least as far as Bantu is concerned (Philippson 2018). Oddly enough, both sides of the "double reflexes" debate stick hard to the Neogrammarian hypothesis and never consider the possibility of irregular sound change.

In the end, the jury is out on how to account historically for irregular sound correspondences in Bantu. Whether they are seen as an indicator of intensive language contact or as "double reflexes" whose regularity is salvaged by either ancient contrasts or latent conditioning, the different accounts all suffer from both a lack of sufficient language data and a clear genealogical framework against which the comparative evidence can be interpreted diachronically. Most studies mentioned in the preceding paragraphs are based on relatively little data, lack a systematic quantitative assessment of assumedly unconditioned reflexes of a proto-sound, and are carried out on an arbitrarily selected set of languages, whether or not spoken in a well delimited part of the Bantu domain, but certainly never within a given genealogical subgroup of the Bantu family.

This is why in this article, we present the first quantitative study of what we call "multiple unconditioned reflexes" (henceforth MUR) in the West-Coastal Bantu branch of the Bantu language family. The term MUR refers to a situation where one and the same proto-sound appears to have two or more synchronic reflexes in a given language without any phonological conditioning environment to tease them apart. We focus here on the MUR of the PB velar stops *k and *g in WCB, which we consider to be illustrative of the broader picture in the northwestern part of the Bantu domain. Our study focuses on one specific branch within the Bantu family whose genealogical unity and internal structure are well established, mainly on the basis of lexicon-based quantative approaches (Vansina 1995; Bastin, Coupez & Mann 1999; de Schryver et al. 2015; Grollemund et al. 2015; Pacchiarotti et al. 2019), but also corroborated by a uniquely shared phonological innovation (Pacchiarotti & Bostoen 2020). What is more, we have a decent knowledge of the diachronic sound changes different WCB languages underwent (Daeleman 1977; Rottland 1977; Hombert & Mouélé 1988; Idiata-Mayombo 1993; Mouélé 1997; Nguimbi-Mabiala 1999; Koni Muluwa 2010: 117–161; Bostoen & Koni Muluwa 2011; Crane, Hyman & Tukumu 2011: 255–270; Koni Muluwa & Bostoen 2011, 2012; Bostoen & Goes 2019; Goes & Bostoen 2019; Pacchiarotti & Bostoen 2020, 2021). Finally, thanks to dedicated documentation work as part of the successive KongoKing[1] and BantuFirst[2] projects, we have an unequalled archive of digitial language sources for this specific branch of the Bantu family and thus access to plenty of comparative language data.

---

[1] http://kongoking.net/

[2] https://www.bantufirst.ugent.be/

Our study serves four purposes. The first is to provide strong comparative evidence, both qualitative and quantitative, to prove that MUR are really the outcome of irregular sound change and need not be explained by positing either additional consonantal series in PB or latent conditioning factors that went lost. The second is to show that MUR in WCB are not so different from cases of pervasive irregularity in sound change reported in other areas of the world (Blust 1996; Grace 1996; Ross 1996) and that the common thread underlying lack of Neogrammarian regularity in geographically distant and genetically unrelated language families seems to be found in the sociocultural situation in which speech communities develop. The third is to demonstrate that intensive language contact indeed needs to be factored in as the main source for irregular sound change in Bantu, but not in such a way that it would undermine the basic Neogrammarian principle of regularity underlying the Comparative Method (cf. Campbell 1996). At least the "regularity assumption" remains intact (cf. Harrison 2003: 219–220). The fourth and last is to show that prehistoric language contact within the Bantu domain is not necessarily to be understood as the interaction – whether mutual or not – between neatly delimited speech communities speaking distinct languages, but could also be conceived as a multilingual context without "a close association between linguistically defined units and communities" (Grace 1996: 175). Language contact scenarios to which irregular sound change is commonly attributed, such as lexical diffusion and substrate influence, tend to start out from the (implicit) assumption that a given language is spoken by a given speech community or ethnic group living in a well circumscribed territory. If one departs from this traditional view of what a language is socially, conceptual borders between lexical diffusion and substrate influence become fuzzy. While lexical diffusion is typically seen as process happening within a speech community, substrate influence tends to be seen as an external influence, i.e., the impact of a recently lost source language on a recently acquired target language through language shift. Besides lexical diffusion and substrate influence, we entertain the possibility that the sound change irregularity observed in WCB might also be due to the stratified nature of Bantu language history.

In line with these objectives, this article is organized as follows. In Section 2, we give some background on WCB and present the data and methodology. In Section 3, we show the results of the first quantitative study of MUR for Proto-Bantu velar stops *k and *g in 40 WCB languages (Sections 3.1–3.2) and argue that they cannot be accounted for by either phonological conditioning factors (Section 3.3) or borrowing (Section 3.4). In Section 4, we consider different possible historical explanations to account for the synchronic situation emerging from the case study in Section 3: lexical diffusion (Section 4.1), substrate influence

(Section 4.2), widespread multilingualism (Section 4.3), and spread-over-spread events in Bantu language history (Section 4.4). Conclusions are in Section 5.

2. Data and methodology

Ever since Vansina (1995), lexicostatistical and lexicon-based phylogenetic approaches to Bantu internal classification have supported the existence of a major branch within the Bantu family called West-Coastal Bantu (henceforth WCB) (Bastin, Coupez & Mann 1999; Bostoen et al. 2015; de Schryver et al. 2015; Bostoen & de Schryver 2018a, 2018b), also known as West-Western (Grollemund et al. 2015). Geographically, this branch spans across Gabon, the Republic of the Congo, the Democratic Republic of the Congo (henceforth DRC), and northern Angola. In terms of Guthrie's referential (i.e., non-genealogical) classification (cf. Guthrie 1971; Maho 2009), this branch includes the B40–80, H10, and H30 language groups as well as Hungan H42 and Samba L12a. Figure 1 presents major subgroups (represented by different colors) according to the most recent and comprehensive lexicon-based phylogenetic classification of WCB in general (Pacchiarotti, Chousou-Polydouri & Bostoen 2019) and the Kikongo Language Cluster (henceforth KLC) in particular (de Schryver et al. 2015; Bostoen & de Schryver 2018a, 2018b).

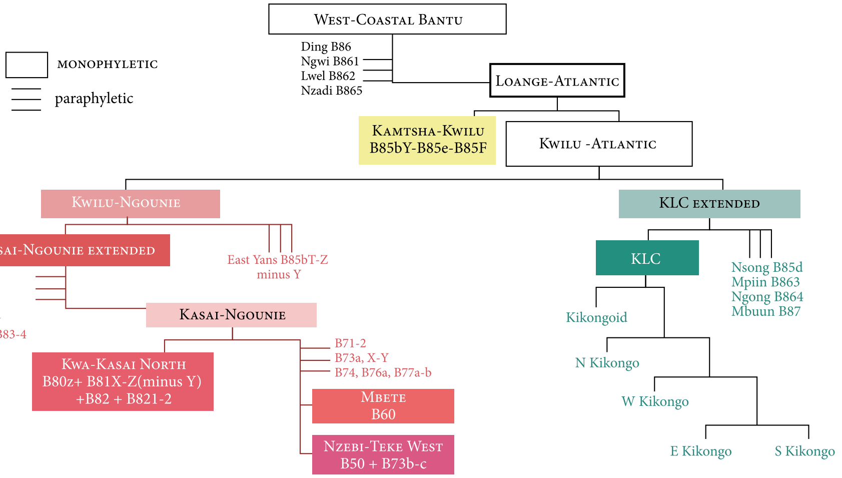


Figure 1. Internal lexicon-based phylogenetic classification of WCB (de Schryver et al. 2015; Bostoen & de Schryver 2018a, 2018b; Pacchiarotti, Chousou-Polydouri and Bostoen 2019)

We choose the WCB branch for our quantitative case study because, unlike others, this branch has been the target of in-depth historical phonological work (Daeleman 1977; Rottland 1977; Hombert & Mouélé 1988; Idiata-Mayombo 1993; Mouélé 1997; Nguimbi-Mabiala 1999; Koni Muluwa 2010: 117–161; Bostoen & Koni Muluwa 2011; Crane, Hyman & Tukumu 2011: 255–270; Koni Muluwa & Bostoen 2011, 2012; Bostoen & Goes 2019; Goes & Bostoen 2019; Pacchiarotti & Bostoen 2020, 2021). Particularly relevant to our purposes is a recent study based on 41 WCB language varieties which shows that when not preceded by a nasal, PB *k and *g merged into *k minimally at Proto-WCB (PWCB) level (Pacchiarotti & Bostoen 2020). While PWCB *k is stable and retained as such root-initially (i.e., C_1 in a $C_1V_1C_2V_2$ structure) virtually in every WCB language, there is much more variation and thus innovation in the developments of PWCB *k in root-medial position (i.e., C_2 in a $C_1V_1C_2V_2$ structure). Besides retentions of the PWCB velar stop *k, two major innovations in C_2 position are (i) lenitions to different kinds of fricatives in the back of the oral cavity, and (ii) loss (i.e., zero). However, very few WCB varieties are consistently conservative or innovative. Pacchiarotti and Bostoen (2020) note that PWCB *k in C_2 position has multiple unconditioned reflexes (MUR) in many present-day WCB languages. In this article, we focus on one specific branch of the Bantu family and on the MUR of PB velar stops in C_2 position. Nonetheless, we consider our case study as representative of the broader phenomenon of MUR affecting the reconstructed series of PB voiceless and voiced stops (*p, *t, *k and *b, *d, *g) in C_1 and/or C_2 position, which has been examined since the 1980s especially in the northwestern part of the Bantu domain, i.e., Cameroon and Gabon, in relation to languages found in Guthrie's referential zones A and B (see references in Section 1). An example of MUR (both in C_1 and C_2 positions) reconstructed by Bancel (1988: 10) for the putative most recent common ancestor of the so-called Ewondo-Fang languages of Guthrie's A70 group is given in (1). The fact that MUR can be reconstructed to an ancestral stage indicates that present-day Ewondo-Fang languages manifest the same variation and that this diachronic phonological irregularity must have a certain time depth. In (1), noun classes (CL) are given in parentheses. BLR in (1) stands for Bantu Lexical Reconstructions 2/3 (Coupez, Bastin & Mumba 1998; Bastin et al. 2002); see discussion below. A blank cell means 'does not apply'.

(1)

Proto-Bantu		Proto-A70	MUR in C_1	MUR in C_2
BLR 1179	*dʊ́k ‘vomit’	*yô	*d > *y	*k > *Ø
BLR 1109	*dók ‘rain’	*e-lók ‘rain’ (CL 7/8)	*d > *l	*k > *k
BLR 2476	*pèèpè ‘wind’	*m-fə̂p (CL 3/4)	*p > *f	*p > *p
BLR 2430	*pèdim ‘flash’	*və́s	*p > *v	
BLR 2684	*pùkò ‘bag’	*m-fə̂k (CL 3/4)	*p > *f	*k > *k
BLR 2682	*púk ‘be mature’	*vú	*p > *v	*k > *Ø
BLR 2827	*tèk ‘be soft’	*tə̀k	*t > *t	*k > *k
BLR 3105	*túkʊ̀ ‘night’	*a-lû (CL 5/6)	*t > *l	*k > *Ø
BLR 2811	*tátʊ̀ ‘three’	*lâl	*t > *l	*t > *l
BLR 2808	*tààté ‘father’	*tàtə́ (CL 1a/2)	*t > *t	*t > *t
BLR 2816	*tédam ‘stand’	*tə́lə́	*t > *t	*d > *l
BLR 2727	*tádè ‘iron ore’	*e-láɛ̀ (CL 7/8)	*t > *l	*d > *Ø
BLR 3101	*túd ‘forge’	*lúì	*t > *l	*d > *Ø

To develop a systematic examination of such MUR in a well-defined though large enough branch of the Bantu family, we took as a point of departure for the present case study the 66 cognate sets in Pacchiarotti and Bostoen (2020) that have reflexes of PB *k and *g (= PWCB *k) in C_2 position in 41 different WCB varieties. The cognate sets in Pacchiarotti and Bostoen (2020: 175–189) are chosen based on the most widespread protoforms across WCB. To these, we added for the same convenience sample (minus Teke Tyee B73d see Appendix B) language-specific reflexes of roots containing PB *k or *g in C_2 position that are less widely distributed (see Appendix A). To find these additional reflexes in each of the language varieties for these two PB velar stops, we used a top-down approach. This means that instead of positing protoforms based on form (and meaning) of reflexes found in a given cognate set, we relied on the already existing protoforms in the *Bantu Lexical Reconstructions* 2/3 database (Coupez, Bastin & Mumba 1998; Bastin et al. 2002) and posited synchronic forms in individual languages as likely reflexes of those protoforms. We chose to develop a MUR case study of velar protosegments for reasons of convenience, as it allows us to build on published cognate sets. Nevertheless, the phenomenon of MUR illustrated here is certainly not epiphenomenal for reflexes of velar consonants in WCB. It is also found in the reflexes of PB labial stops in C_1 position (especially reflexes of PB *b and *p) and alveolar stops in C_2 position (especially reflexes of PB *d).

In Appendix B, we list all varieties included in this study with their corresponding alphanumeric code, their phylogenetic subgroup according to Pacchiarotti, Chousou-Polydouri and Bostoen (2019) and their data sources.[3] In Appendix C, we show the MUR of PWCB *k (the merged reflex of PB *k and *g) in C_2 for each variety in our 40-language convenience sample. The three types of MUR we have identified are (i) voiceless velar stop /k/, (ii) dorsal fricatives, and (iii) zero. All MUR listed in Appendix C are to be understood as reflexes in the absence of a specific conditioning environment, such as high vowels, which commonly trigger distinctive sound shifts in Bantu (Schadeberg 1995; Labroussi 1999; Bostoen 2008). We illustrate this in (2) with data from Nzebi B52, which has /k/ and /x/ as reflexes of both PB *k and *g in C_2.[4] In (2) and throughout the paper, meanings of reflexes are indicated only when they differ from the meaning assigned to the protoform. Most reconstructions are taken from the BLR database, see (1). Those which say “Proto-Duma” are taken from Mouélé’s (1997: 349 and ff.) reconstruction of lexicon in Proto-B50, the most recent common ancestor of Wanzi B501, Duma B51, Nzebi B52, and Tsaangi B53. As can be seen by comparing for example *Ø-kókò* ‘chicken’ < BLR 1904 *kókó ‘chicken’ with kòxó ‘log’ < BLR 1427 *gògó ‘trunk, bridge’, it is hard to find a phonological conditioning environment to tease the /k/ and /x/ reflexes apart. We discuss this in detail in Section 3.3.

(2) Nzebi B52 (Marchal-Nasse 1989; Mouélé 1997)

PB *k in C_2	BLR 66	*bák ‘build’	> *ì-báxà* ‘wall’
	BLR 67	*bák ‘get, catch, rob’	> *ù-báxà* ‘get, receive, earn’
	BLR 820	*dáká ‘tongue, language’	> *n-dáxà* ‘language’
	BLR 9590	*dákò ‘house (for men)’	> *mù-lákà* ‘camp’
	BLR 9642	*káká ‘foot’	> *lè-kákà* ‘hand’

---

[3] To carry out this quantitative study, we collapsed data from more than one dialectal variety of a single language (distinguished by means of capital X, Y, and Z after the alphanumeric code in Appendix B). For example, data on Nzebi (Lébamba) B52Z from Marchal-Nasse (1989) and on Nzebi (Mbigou) B52Y from Mouélé are collapsed under Nzebi B52 in Appendix C. This paper also provides original fieldwork data on three poorly documented Bantu languages spoken in the DRC, namely Mpe B821, Nunu B822, and Ngwi B861.

[4] Different authors report different fricatives as the reflex of PB *g and *k in C_2 in Nzebi B52. Marchal-Nasse (1989) reports [x], while Mouélé (1997) reports [ɣ]. These could be dialectal differences. In (2) we adapt the transcription of the data from Mouélé (1997) to [x] for convenience.

	BLR 1685	*kààká ‘grandparent’	> *Ø-kààxá*[5]
	BLR 1906	*kòòkó ‘grandparent’	> *kòòkó*
	BLR 7983	*kéèkéè ‘little, small’	> *mù-ɣééɣè* ‘youngest sibling’
	BLR 2286	*nók ‘rain (v.)’	> *ù-nóxò* ‘rain (v.)’
	BLR 647	*còká ‘axe’	> *Ø-tsòkó* ‘traditional axe’
	BLR 2967	*tòk ‘boil’	> *ù-tòxò*
	BLR 1904	*kókó ‘chicken’	> *Ø-kókò*
	Proto-Duma	*kòkò ‘domestic cat’	> *mu-kòxò*
	BLR 1355	*gègò ‘molar’	> *kèkə̀*
	BLR 145	*béek ‘put (away)’	> *ù-bééka*
	BLR 2828	*ték ‘draw (water)’	> *ù-téxè*
PB *g in C_2	BLR 1041	*dígì ‘string’	> *nzíxì*
	BLR 1828	*kígè ‘eyebrow’	> *Ø-kíkì*
	BLR 60	*bàgú ‘stumbling block’	> *bàká*
	BLR 55	*bààg ‘tear’	> *ù-bàːkà*
	BLR 1427	*gògó ‘trunk, bridge’	> *kòxó* ‘log’
	BLR 808	*dàg ‘show (v.)’	> *ù-làxà*
	BLR 812	*dàgá ‘promise’	> *làxà*
	BLR 316	*bógà ‘village, path’	> *mbóxà* ‘at someone’s place, village’
	BLR 315	*bóg ‘plaster, dig foundations’	> *ù-bókà* ‘dig’
	BLR 900	*dègè ‘weaver bird’	> *ndèkè*
	BLR 2180	*mìg ‘try (v.)’	> *ù-mèxà* ‘taste (v.)’
	BLR 3338	*jíg ‘learn, imitate’	> *ù-yíxà* ‘learn’

In Appendix C, MUR for each variety are ordered according to frequency, from the most frequent (see column “Reflex 1”) to the least frequent (see column “Reflex 3”). Each “Reflex” column has four subcolumns. The first subcolumn is the actual phonetic form of the reflex of PWCB *k in a given variety, e.g., /ɣ/. The second subcolumn shows the ratio of BLR roots reconstructed with a PB *k or

---

[5] Pacchiarotti and Bostoen (2020: 185) give *ngáxà* ‘family, relative, friend’ as the Nzebi reflex of BLR 1685 *kààká ‘grandparent’. This is a mistake. Mouélé (1997: 387, 390) reconstructs Proto-Duma *kààxá ‘grandparent’ as a reflex of BLR 1685 *kààká, and Proto-Duma *ŋgàxá ‘parent’ as not linkable to any existing reconstruction in the BLR database.

*g in C_2 displaying that reflex in a given variety over the total amount of reflexes identified for that variety. For instance the "Reflex 1" column for Punu B43 in Appendix C reads that /ɣ/ is Reflex 1 (the most frequent reflex) of PWCB *k, and that we found 47 reflexes in Punu out of 58 total reflexes of BLR reconstructions containing a PB *k in C_2 having /ɣ/ as a reflex (i.e., *k: 47/58), as well as 23 reflexes out of 34 total reflexes of BLR reconstructions containing a PB *g in C_2 having /ɣ/ as a reflex (i.e., *g: 23/34). The third subcolumn converts each fraction (for each PB velar stop) into a percentage for the purpose of comparing percentages across varieties (i.e., in Punu B43 47/58 equals 81% of total identified reflexes of PB *k in C_2; 23/34 equals 68% of total identified reflexes of PB *g in C_2). The fourth subcolumn gives the combined percentage of roots with PB *k and *g in C_2 having a given reflex. For example, in Punu, we identified 92 reflexes of PB reconstructions containing either PB *k (58) or *g (34) in C_2. Of these, a total of 70 have /ɣ/ as a reflex (47 for *k and 23 for *g): 70/92 equals 76% of all reflexes.

3. Results

In Map 1, we plot the MUR of PWCB *k in C_2 for each variety in our convenience sample. Specifically, each variety is represented by a color-coded pie chart displaying the total percentage of each reflex (see subcolumn 4 in each "Reflex" column in Appendix C).[6] By looking at the percentages in Appendix C and the pie charts in Map 1, a first observation is that some varieties have only one reflex for PWCB *k in C_2, which is always Ø (=zero) (Ngungwel B72a, Eboo-Nzikou B74, and Mfinu B83). Other varieties have predominantly one reflex and only a smaller percentage of a second reflex, but ratios differ across languages:

- 9:1 ratio: 90% or more lexical items with /k/ as a shared retention and only 10% or less with a zero reflex in the most conservative varieties (Boma Yumu B80z, Tiene B81, East Yans B85b, Ngong B864, and all KLC varieties of zone H, i.e., Hangala H111, Sikongo H16a, Manyanga H16b, Yombe H16c, Ntandu H16g, and Yaka H31); over 90% fricative reflexes and 10% Ø in some innovative varieties (e.g., Mpe B821, Nunu B822, North Boma B82, and Ngwi B861) and 87% Ø and 13% /k/ in Yaa B73c;
- 8:2 ratio: nearly 80% Ø and 20% /k/ in Ding B86 and Mbuun B87 or nearly 80% fricatives and 20% /k/ in Tsaangi B53;

---

[6] Map 1 was created by using the Free and Open Source Sotfware QGIS, a professional Geographic Information System (GIS) application downloadable at https://download.qgis.org.

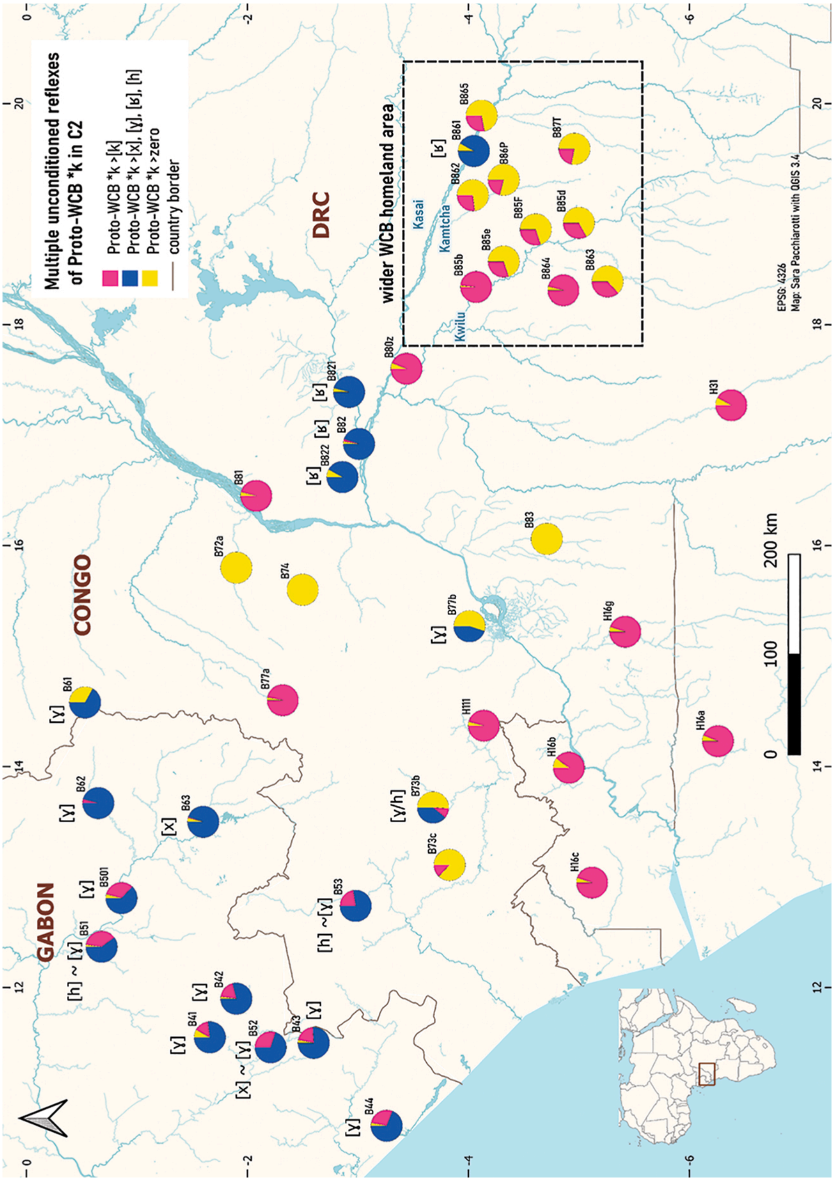
- 7:3 ratio: nearly 70% Ø and 30% /k/ in Lwel B862, Nzadi B865, and Nsong B85d and 70% of fricatives and 30% /k/ in Nzebi B52;
- 6: 4 ratio: 63% Ø and 37% /k/ in Mpiin B863 or 60% fricatives, 38% /k/ and 2% Ø in Duma B51.

In only a couple of varieties, the ratio is almost 5:5 (Fumu B77b) or 5:4:1 (Laali B73b). Excluding Laali, in varieties with up to three MUR, one of these three is usually lower than 10% (e.g., Nduumo B63 with 1% /k/ or Punu B43 and Lumbu B44 with 3% Ø). Thus, a relevant question that has never been considered so far in the Bantu literature on so-called "double reflexes" is what ratio should be taken as threshold to claim the existence of MUR in a given Bantu language. Whatever the answer to this question might be, some of these ratios seem to be areally concentrated in Map 1.

A second, visually prominent aspect emerging from Map 1 has to do with cycles of innovations. Pacchiarotti and Bostoen (2020) posit the lenition chain in (3) to account for the innovation of fricatives and Ø after the fragmentation of PWCB.

(3) PB *k and *g > PWCB *k > Post-PWCB x/ɣ/ʁ/h > Post-PWCB Ø

However, Map 1 provides evidence for two distinct and possibly separate paths of innovation of PWCB /k/. The first one is PWCB *k > dorsal fricative > Ø, as proposed by Pacchiarotti and Bostoen (2020). The second one is PWCB *k > Ø. Although the abrupt loss of simple stops in intervocalic position might be rare cross-linguistically, we do not have any synchronic evidence to posit an intermediate lenition stage. Each of these chains operates within the same language variety and creates MUR for one and the same protosound. The lenition chain in (3) is supported by the B40 languages, Wanzi B501, Duma B51, Nduumo B63, and Laali B73b, where we find /k/, dorsal fricatives and zero as reflexes of PWCB *k (see Map 1). At the same time, there is evidence in favor of a direct shift from *k > Ø without going through an intermediate fricative stage. The twelve varieties with the highest number of Ø as reflexes of PWCB *k in C_2 ranging between 100% and 63% do not have any fricative reflexes. None of these languages has a dorsal fricative in its phonological inventory. These are: the first paraphyletic offshoots Ding B86, Lwel B862, and Nzadi B865; the KLC extended varieties Nsong B85d, Mbuun B87, Mpiin B863; the Kamtsha-Kwilu varieties Nsambaan B85F, Mpur B85e; and the Kasai-Ngounie extended paraphyletic varieties Ngungwel B72a, Yaa



Map 1. Multiple unconditioned reflexes of PWCB *k in C_2

B73c,[7] Eboo-Nzikou B74, Mfinu B83. If they have a second reflex, this is a retention of PWCB *k (ranging from 13% to 37%). These languages occur in all major subbranches of WCB. In our view, this is a strong indication for a direct PWCB *k > Ø sound shift. Similarly, the eleven most conservative varieties having between 98.5–90% retentions of PWCB *k only have zero as an alternative reflex but never a fricative (i.e., within Kwilu-Ngounie: Kukwa B77a, Boma Yumu B80z, Tiene B81, East Yans B85b; within KLC extended: Ngong B864, Hangala H111, Sikongo H16a, Manyanga H16b, Yombe H16c, Ntandu H16g, Yaka H31). We take this as one more indication that the *k > Ø innovation can start out by simply losing /k/ without going through an intermediate fricative stage.[8]

A third consideration emerging from Map 1 related to the two cycles of innovation described above is that different languages are synchronically at different stages in the chain PWCB *k > fricative (> Ø). In some languages, PWCB *k in C_2 only underwent lenition to a dorsal fricative. Other languages went a step further and are losing the fricative in C_2 position. Among those which only innovated a fricative, this change appears to have applied to more than 90% of all retreivable lexical items in Nduumo B63, Mbaama B62, and North Boma B82. Languages which innovated a fricative but preserve a more significant percentage of roots which still have /k/ are several Nzebi-Teke West varieties, such as Wanzi B501, Duma B51, Nzebi B52, and Tsaangi B53 as well as the West-Kongo B40 languages of the KLC. In Shira B41, Punu B42, and Sangu B43, 10–30% of lexical items have /k/ instead of a fricative. To a neglectible extent, zero is also attested alongside /k/

---

[7] Yaa B73c stands out for being the only variety within Nzebi-Teke West to have almost completed the innovation to zero without showing any trace of a fricative as the reflex of PWCB *k in C_2.

[8] Two anonymous reviewers asked whether we knew of languages outside of Bantu where the loss of a velar happens without going through an intermediate lenition stage such as a fricative. We are not aware of any diachronic cases in intervocalic position, only in consonant clusters. Cser (2015: 199) states that "the gradual weakening [...] of sounds that may lead to loss in the long run [...] is not always distinguished terminologically from the more abrupt and categorical kinds of loss often encountered in clusters." Cser (2015: 199–200) offers multiple examples of deletion (without an intermediate "lenited" stage) of consonants in consonant clusters, such as Latin *rupta* > French *route* 'road, way', *costa* > *côte* 'coast, rib'. However, in some present-day Flemish dialects, /k/ is realized as a glottal stop in intervocalic position, e.g., *bakken* 'to bake' > ['baʔᵊn] (cf. De Wulf, Goossens & Taeldeman 2005: 13). This is the closest instance we found of a stop in intervocalic position being realized almost as zero. Additionally, an anonymous reviewer points out that *g is known to disappear more readily than *k. While it is entirely possible that PWCB *k underwent voicing to /g/ before the lenition chain started, we do not have any synchronic evidence for this intermediate stage. While several WCB varieties retained PWCB *k in C2, none of them provides evidence for /g/ (see Pacchiarotti & Bostoen 2020: 175–189).

and a fricative in the B40 group, Wanzi B501, and Duma B51. Perhaps the variety that best illustrates the PWCB *k > fricative > Ø chain is Laali B73b, where 10% of roots still show a conservative /k/, 39% have a fricative while in the remaining 51% the fricative already became zero. Mpe B821, Nunu B822, Ngwi B861, Mbete B61, and Fumu B77b are the only varieties that went a step further in the PWCB *k > fricative > Ø chain. They no longer have roots with /k/ in C_2, but only fricatives and zero as a further development out of the fricative innovation.

After these general observations, we now assess whether the two main innovations of PWCB C_2 *k, i.e., zero (Section 3.1) and dorsal fricatives (Section 3.2), can be reconstructed as shared innovations going back to certain ancestral nodes within WCB. This assessment obviously presupposes that we assign some validity to the phylogenetic WCB subgroupings emerging in Pacchiarotti, Chousou-Polydouri and Bostoen (2019). In Section 3.3 and Section 3.4, we assess whether we can exclude the possibility of explaining MUR discussed in Section 3.1 and Section 3.2 via phonological conditioning (Section 3.3) and/or borrowing (Section 3.4), two factors often held responsible for the apparent irregularity of sound change. Because we deal with 40 varieties in this study, it is not feasible to discuss these possibilities for each and every language, also because our linguistic knowledge is not equal across varieties. Thus, we conveniently choose our own fieldwork data on Ngwi B861 spoken in the homeland region for the argumentation in Section 3.3 and Section 3.4. Most of what we argue for Ngwi is no doubt applicable to other WCB varieties labelled with Guthrie's alphanumeric codes B85 and higher spoken in the wider homeland region (see data in Pacchiarotti & Bostoen 2020: 175–189 and Appendix A).

3.1 Zero reflexes

As observed above, languages having zero reflexes ranging between 100% and 63% are found in all major WCB branches (see Figure 1), namely the first paraphyletic offshoots, KLC extended, Kamtsha-Kwilu, and Kwilu-Ngounie. Because this innovation is at different stages within different branches and given that there are varieties within these branches which are highly conservative in that they mostly have shared retentions of PWCB *k, it is impossible to reconstruct the zero innovation to PWCB level. Rather, the scattered distribution of this innovation across the different branches suggests that it should be considered as a series of independent parallel innovations across branches. This hypothesis is supported by the fact that distantly related WCB varieties such as Mbete B61 and Mbaama B62 on one hand, and Yans B85 and Ngong B864 on the other, synchronically lose velar consonants in C_2 depending on the dialect, as shown in (4)–(6). Sources for these data are specified for each variety in Appendix B. The lack of a BLR index in (6) means that the form is a new tentative reconstruction going back to PWCB.

(4) BLR 70 *bàkàdà 'man, male' > B85aY *bakæl*
B85bV *bɛ̯à:l`*
B864W *bàgál*
B864X *ábá:l*

(5) BLR 3604 *jʊ́g 'hear' > B61Y *ɣo-yuɣa*
B61Z *ŋò-yúwá*

(6) *còkó 'liver' > B62Y *le-sɔɔ*
B62Z *lè-sɔ́ɣɔ́*

This synchronic pattern strongly suggests that loss as a further development of PWCB *k in C_2 in different groups should be explained as distinct, parallel innovations. Considering that other consonants besides velars, e.g., some nasal clusters and some coronals, are also lost in many Teke varieties as well as in several languages closer to the WCB homeland (e.g., Ngwi B861), the recurrent loss of PWCB *k in C_2 was likely favored by a weak phonotactic C_2 position (cf. Hyman 2004, 2010; Teil-Dautrey 2008; Bostoen & Mundeke 2011).

Although PWCB *k > Ø in C_2 should be posited as a parallel independent innovation in several WCB branches, this does not necessarily imply that no cases of PWCB *k loss happened at a deeper ancestral stage. In the case of conservative KLC varieties of Guthrie's zone H, for instance, a legitimate question is whether words which have a zero instead of /k/ as a reflex of PWCB *k unexpectedly lost this velar stop already at a much earlier ancestral stage and inherited a root with Ø in C_2 from their most recent common ancestor. As shown in (7), a root like BLR 2911 *tígad 'remain' lost PB *g (=PWCB *k) everywhere in C_2. This could indicate that the sound was lost at the ancestral node uniting the KLC and the paraphyletic KLC Extended languages (cf. Figure 1).

(7) BLR 2911 *tígad 'remain' > KLC B43 *syaal-a*
B44 *sial*
H111 *sààl-á*
H16a *ku-ssaàl-a*
H16b *sal-a*
H16c *syáál-a*
H16g *sáál*
H31 *sáál-á*
KLC Extended B864X *sal*

The root BLR 3423 *jígu 'hear' in (8) also lost PB *g everywhere in C_2 except in Ngong B864 which maintained PWCB *k. Hence, one could argue that the loss of PWCB *k in C_2 had certainly happened already in Proto-Kikongo, the most recent common ancestor of the KLC, and possibly even slightly earlier.

(8)	BLR 3423 *jígu 'hear' > KLC		H111	*yw-á*
			H16a	*kú-ú-a*
			H16b	*w-a*
			H16c	*w-á*
			H16g	*w-a*
			H31	*w-a*
		KLC Extended	B85dZ	*wʊ:*
			B863Y	*wʊ:*
			B864X	*wuk*
			B87W	*wʊ:*

As for the reflexes of BLR 3527 *jògà 'mushroom' in (9), PWCB *k is retained in two paraphyletic varieties belonging to the KLC Extended branch, namely Mpiin B863 and Ngong B864. Within the KLC proper, PWCB *k has voiced velar fricatives as reflexes in Punu B43 and Lumbu B44. Based on these attestations, it is impossible to reconstruct loss of PB *g (=PWCB *k) in BLR 3527 *jògà 'mushroom' at a deeper ancestral stage uniting at least some of the languages which have a reflex of this protoform in (9). Rather, the Ø reflex of PWCB *k in this lexical item must have emerged independently within the KLC (see e.g., B42, H16b, H16g, H31) on one hand and in the KLC Extended paraphyletic languages such as Nsong B85d and Mbuun B87 on the other.

(9)	BLR 3527 *jògà 'mushroom' > KLC		B42	*by-oowu*
			B43	*bo-oɣu*
			B44	*bu-oɣu*
			H16b	*bù-wa*
			H16g	*bu-wá*
			H31	*bu-wá*
		KLC Extended	B85dZ	*b-ɔ́:*
			B863Y	*bò-óku*
			B864X	*b-ɔ̌k*
			B87W	*b-óó*

3.2 Dorsal fricative reflexes

Dorsal fricative reflexes in C_2 are found close to the wider WCB homeland between the Kwilu and Kasai Rivers (cf. Pacchiarotti, Chousou-Polydouri & Bostoen 2019) and mostly to the northwest of the homeland area (see Map 1).[9] Following the phylogeny in Figure 1, it is untenable to posit the development of fricatives in C_2 as a single shared innovation inherited from a most recent common ancestor. Languages having fricative reflexes of PWCB *k are too scattered across the family tree. This innovation is found in Ngwi B861, one of the languages to split off first within WCB, as well as in languages belonging to much lower nodes, such as the KLC within the KLC extended branch, and Mbete, Nzebi-Teke West, and Kwa-Kasai North within the Kasai-Ngounie branch. Several of these subgroups also include languages which did not innovate PWCB *k in C_2. For example, the Kwa-Kasai North subgroup includes the conservative Boma Yumu B80z and Tiene B81 alongside the innovative Mpe B821, Nunu B822, and North Boma B82 (see Map 1). In the KLC, only the B40 group shows fricative reflexes while all other zone H languages are conservative. Velar fricativization therefore occurred in all likelihood independently and recurrently in WCB as a parallel innovation. This does not mean, however, that all present-day languages having such fricatives developed them individually. Within the Kasai-Ngounie (Extended) subclade (see Figure 1), for instance, several languages did not develop fricative reflexes of PWCB *k as an innovation. Hence, while the innovation cannot be reconstructed back to Proto-Kasai-Ngounie (Extended), it is nevertheless widespread in its three main subgroups. Within Kwa-Kasaï North, North Boma B82, Mpe B821, and Nunu B822 share voiced uvular fricatives. This might indicate that they are more closely related amongst each other than to Boma Yumu B80z and Tiene B81, which also belong to that subclade but by and large retained PWCB *k in C_2. In the two other lower Kasai-Ngounie subgroups, i.e., the Mbete (B60) and Nzebi-Teke West (B50+B73b-c), fricative reflexes of PWCB *k in C_2 are omnipresent. These languages most likely inherited the innovation from their respective most recent common ancestor, or possibly from a slightly older ancestor shared by the two subgroups.

Moreover, certain dorsal fricatives may also have emerged as the result of contact-induced change. Dorsal fricative reflexes of PWCB *k are most pervasive in the northwestern part of the WCB domain. They are not only attested in the

---

[9] We call it wider because the putative WCB homeland area is supposed to be located between the Kamtcha and Kasai rivers (Pacchiarotti, Chousou-Polydouri & Bostoen 2019: 193).

Mbete (B60) and Nzebi-Teke West (B50+B73b-c)[10] subgroups of Kasai-Ngounie, but also in the neighboring West-Kikongo B40 languages, which belong to an entirely distinct WCB branch, i.e., the KLC (Extended). In the remainder of the KLC (Extended), fricative reflexes of PWCB *k are entirely absent. The fact that only the most extreme northwestern KLC varieties which are in direct contact with Nzebi-Teke West (B50+B73b-c) languages acquired velar fricatives could point towards an areal feature. What is most striking is that the B40 languages which acquired velar fricatives as a result of contact are the most innovative WCB varieties: they are the only ones to have innovated velar fricatives also in C_1 position (Pacchiarotti & Bostoen 2020: 150). Contact with neighboring Laali B73b could also account for the fact that Fumu B77b is the only the paraphyletic Kasai-Ngounie variety to have voiced velar fricatives (see Map 1).

As we previously observed for zero reflexes, the fact that fricative reflexes of PWCB *k are best analyzed as parallel independent innovations (except in the B40-B50-B60 convergence area) does not exclude that for some lexemes the fricative innovation might be reconstructed at some common ancestral node. We investigate this possibility in Appendix D, where we present cognate series for two lower-node Kasai-Ngounie subgroups, i.e., Mbete and Nzebi-Teke West. We also include data from Fumu B77b, whose most recent common ancestor with the other languages in Appendix D would be Proto-Kasai-Ngounie itself according to the latest lexicon-based phylogeny (Pacchiarotti, Chousou-Polydouri & Bostoen 2019).

As far as data are available, roots 1–7 in Appendix D seem to have innovated a fricative out of PWCB *k across all languages. Hence, the most parsimonious explanation would be that the shift to a fricative started in the most recent common ancestor uniting Fumu with the Mbete and Nzebi-Teke West. This cannot be Proto-Kasai-Ngounie itself as several languages belonging to Kwa-Kasai North, the other monophyletic Kasai-Ngounie subgroup (cf. Figure 1), have retained PWCB *k in C_2 (cf. Boma Yumu B80z and Tiene B81 in Map 1). In other words, the fricative shift could be a shared innovation indicating that the Mbete and Nzebi-Teke West subgroups and the Kasai-Ngounie paraphyletic languages like Fumu B77b (cf. Figure 1) are actually more closely related to each other than to Kwa-Kasai North and thus resolve the paraphyly that emerged relying on basic vocabulary. More solid comparative data are needed to firmly prove this point.

---

[10] Yaa B73c, which is part of Nzebi-Teke West according to lexicon-based phylogeny in Figure 1, stands out amongst other languages within this group in that it does not have fricative reflexes of PWCB *k at all. It rather underwent loss of PWCB *k in C2. This could be interpreted as a challenge to the phylogeny, but this specific innovation could also be considered as a further development in a chain such as PWCB *k>fricative>Ø.

On the other hand, as far as roots 8–19 in Appendix D are concerned, none of them seems to have consistently undergone the PWCB *k > fricative (> Ø) innovation across all languages. Hence, for all of these roots, it is necessary to posit PWCB *k as a retention at some higher ancestral node within the Kasai-Ngounie branch. This suggests that while this chain of innovation had started in the most recent common ancestor of Fumu, Nzebi-Teke West, and Mbete (cf. roots 1–7 in Appendix D), it started out gradually, and affected only some targetable lexical items at a time (cf. roots 8–19 in Appendix D). While the roots that had not been affected earlier on did undergo the shift quite consistently in Fumu B77b and Mbete B60, this does not seem to have been the case in Nzebi-Teke West, where PWCB *k was retained across all languages in several other roots, see the reflexes of *bʊ̀gí 'squirrel', *gègò '(molar) tooth', *kígì 'eyebrow', *káká 'foot', and *páágʊ̀ 'tree fork'.

3.3 Excluding phonological conditioning as an explanation for MUR

In this section we assess whether we can exclude the possibility of explaining MUR discussed in Section 3.1 and Section 3.2 via phonological conditioning using Ngwi B861 as a test case. Because of its relevance for the following discussion, readers should be aware that like other WCB varieties spoken in the homeland area, Ngwi underwent systematic final vowel loss (Pacchiarotti & Bostoen 2021). By this we mean the historical loss of the word-final vowel segment in the reflexes of nouns and verbs reconstructed as *CVCV or *CVNCV in BLR3. Historical comparative research on final vowel loss indicates that this diachronic change is relatively recent and happened only *after* the loss of intervocalic consonants in several languages of the homeland area. This means that final vowel loss did not apply to CVV shapes created from the loss of a C_2 consonant in a $*C_1V_1C_2V_2$ template, see e.g., *Ø-mfúù* 'mouse' in (10). As can be seen in Map 1, PWCB *k (the merged reflex of PB *k and *g) in C_2 has /ʁ/ as the most common reflex in Ngwi and zero as a second, less common reflex. We illustrate these two reflexes in (10). The fact that there are more zero reflexes for reconstructions featuring PB *k in C_2 is an artifact of the BLR3 database which contains many more reconstructions with *k in C2 than *g in C2 (see Pacchiarotti & Bostoen 2020: 162). In noun forms, we separate the noun class prefix from the simple noun stem. Infinitive verb forms in Ngwi consist of the root only.

(10) Ngwi B861 (own fieldwork)

PB *k in C_2	BLR 1179	*dʊ́k 'vomit'	> *lúà*	/Ø/
	BLR 5333	*pʊkʊ 'burrow'	> *ì-pûʁ* 'garbage hole'	/ʁ/
	BLR 2642	*pʊ́kʊ̀ 'mouse'	> *Ø-mfúù*	/Ø/
	BLR 5464	*cúkʊ̀ 'sauce'	> *ò-súʁ*	/ʁ/

	BLR 5339	*tʊ́kì 'insult'	> *ì-tî̂*	/Ø/
	BLR 3536	*jókà 'snake'	> *Ø-ndʒúà*	/Ø/
	BLR 7413	*cókì 'saliva'	> *à-súì*	/Ø/
	BLR 9461	*cákú 'safou'[11]	> *è-súú* 'safou (fruit)'	/Ø/
	BLR 9605	*pákù 'honey'	> *Ø-mpâʁ*	/ʁ/
	BLR 1904	*kókó 'chicken'	> *Ø-ŋkɔ́ʁ*	/ʁ/
	BLR 1685	*kààká 'grandparent'	> *Ø-ŋkíáŋkíá* 'paternal aunt'	/Ø/
PB *g in C_2	BLR 2634	*pʊ̀gí 'pot'	> *Ø-mpǔʁ*	/ʁ/
	BLR 761	*cúgù 'day'	> *è-ʃúù*	/Ø/
	BLR 4992	*tʊ́gʊ́ 'ashes'	> *ò-túʁ*	/ʁ/
	BLR 316	*bʊ́gà 'path'	> *Ø-mbûʁ*	/ʁ/
	BLR 814	*dàgò 'promise'	> *ì-lâʁ*	/ʁ/
	BLR 900	*dègè 'weaver bird'	> *è-lɛ̂ʁ*	/ʁ/
	BLR 2433	*pègà 'shoulder'	> *ì-pɛ́àʁ*	/ʁ/
	BLR 1248	*dʊ́g 'paddle'	> *lûʁ*	/ʁ/
	BLR 1607	*jògù 'elephant'	> *Ø-ndzɔ́ɔ̀*	/Ø/

As can be seen in (10), the secondary zero reflex of PWCB *k in Ngwi is found in those lexical items which were not possible targets for final vowel loss because they irregularly lost PWCB *k in C2, while in most other lexemes this protosound evolved to /ʁ/. Perhaps one of the most striking examples showing the impossibility of finding a conditioning environment to tease apart the /ʁ/ and zero reflexes of Ngwi are *lúà* 'vomit' < BLR 1179 *dʊ́k 'vomit' and *lûʁ* 'paddle' < BLR 1248 *dʊ́g. Due to the merger of PB *k and *g, at PWCB stage these two protoforms were (supra)segmentally identical, namely *dʊ́k. However, in *lúà* 'vomit' PWCB *k was lost while in *lûʁ* 'paddle' PWCB *k evolved into a voiced velar fricative.

Let us now examine the noun forms in (10) in search of a possible phonological conditioning. The word for 'mouse' *Ø-mfúù* < BLR 2642 *pʊ́kʊ̀ 'mouse' is among those which display the zero reflex. In this word, the historical noun class prefix of class 9 *N-, a homorganic nasal, was reinterpreted as part of the root. Synchronically, the word belongs to class 7Ø/8Ø. One might think that this morphological reanalysis played a role in the irregular change PWCB *k >Ø in C_2 in Ngwi. In fact, two lexical items displaying the zero reflex in (10) have a fossilized nasal as part of their simple stem: *Ø-ndʒúà* 'snake' < BLR 3536 *jókà 'snake', *Ø-ndzɔ́ɔ̀* < BLR 1607 *jògù 'elephant'. However, there are also several items with a fos-

---

[11] The scientific name of this fruit tree is *Dacryodes edulis*. In West and West-Central African countries, it is also known as atanga, ube, African pear, bush pear, African plum, nsafu, bush butter tree, and butterfruit (cf. Bostoen 2014: 134–135).

silized nasal as part of their simple stem which display the /ʁ/ reflex, e.g., *Ø-mbûʁ* 'path' < BLR 316 *bʊ́gà, *Ø-ŋkɔ́ʁ* 'chicken' < BLR 1904 *kókó. One might then posit that it was the particular quality of historical V_1 or V_2 (or a combination thereof) which had an effect on the reflex of PWCB *k in C_2 in Ngwi. This hypothesis is untenable as shown by pairs such as *è-súú* 'safou (fruit)' < BLR 9461 *cákú 'safou' and *Ø-mpâʁ* 'honey' < BLR 9605 *pákù 'honey' where we have the same historical vowels in V_1 and V_2 (*a and *u respectively) and two different reflexes for PWCB *k in C_2 (see also *ì-pûʁ* 'garbage hole' < BLR 5333 *pʊkʊ 'borrow' vs. *Ø-mfúù* < BLR 2642 *pʊ́kʊ̀ 'mouse').[12] The initial consonant of the $C_1V_1C_2V_2$ template does not make a good candidate as a phonological conditioner either, cf. *è-ʃúù* 'day' < BLR 761 *cúgù vs. *ò-súʁ* 'sauce' < BLR 5464 cúkʊ̀ or *ò-túʁ* 'ashes' < BLR 4992 *tʊ́gʊ́ vs. *ì-tîì* insult < BLR 5339 *tʊ́kì. Yet another possibility is that the different reflexes have to do with internal reduplication, but the /ʁ/ reflex of BLR 1904 *kókó 'chicken' and the zero reflex of BLR 1685 *kààká 'grandparent' speak against it.[13] A tonal conditioning factor distinguishing the two reflexes must also be discarded based on the data in (10). Zero reflexes of PWCB *k in C_2 in Ngwi are found in reflexes of reconstructions with the following tone patterns: *HL as in BLR 761 *cúgù, *LL as in BLR 1607 *jògù, *HH as in BLR 9461 *cákú, and *LH as in BLR 1685 kààká. Nevertheless, reflexes of reconstructions with these tone patterns can also display the /ʁ/ reflex of PWCB *k in C_2, see e.g., BLR 9605 *pákù 'honey', BLR 900 *dègè 'weaver bird', BLR 1904 *kókó 'chicken', and BLR 2634 *pʊ̀gí 'pot'.

3.4 Excluding lexical borrowing as an explanation for MUR

Having excluded all possible segmental and suprasegmental conditioning environments, we now turn to lexical borrowing to salvage the regularity hypothesis and account for the apparent MUR in WCB varieties. The idea is that in languages like Ngwi, the more widespread /ʁ/ reflex would be the regular reflex of PWCB *k in C_2, while zero would have been introduced through borrowed words from neighboring varieties where zero is the most common reflex of PWCB *k in C_2 such as Ding B86, Lwel B862, or Nzadi B865 (see Map 1). According to this hypothesis, sound change is regular, despite the irregularity of the reflexes in one

---

[12] The historical *p in BLR 2642 *pʊ́kʊ̀ 'mouse' became [f] in the Ngwi reflex *mfúù*, but not in the Ngwi reflex *pûʁ* < BLR 5333 pʊkʊ. This fricativization triggered by a following PB *i, *ɪ, *u, and *ʊ is known to happen sporadically and irregularly in several WCB varieties spoken in the homeland area (Pacchiarotti & Bostoen 2020: 153).

[13] See Section 3.4 for a discussion of the formal (ir)regularities of *Ø-ŋkíáŋkíá* 'paternal aunt' as a reflex of BLR 1685 *kààká 'grandparent'.

and the same language: regular sound changes in different languages are blurred by later borrowings from related languages where the regular sound change is not at the same stage (cf. dorsal fricative and zero reflexes in Section 3.1 and Section 3.2). In (11), we present some easily detectable borrowings in Ngwi which have not been included in the present case study, i.e. nouns and verbs with a non-derivational final vowel.[14] Often times speakers themselves are aware that certain lexical items are borrowings and indicate the donor language. We add this information whenever available in parentheses next to each entry. Kongo Ya Leta is a vehicular variety of Kongo used as a *lingua franca* in the region where Ngwi is spoken (cf. Swartenbroeckx 1973).[15]

(11) Ngwi B861 (own fieldwork)

BLR 4998	*kʊ̀gʊ́	‘sugar cane’	> *ò-kùʁú* (cf. Kongo *mùkùkú*)[16]
BLR 7402	*tòkó	‘mat’	> *ì-tòʁó* (cf. Kongo *kítókò*)
BLR 1905	*kòòkò	‘sheep’	> *ò-kòkò* (cf. Kongo *kókò* ‘ram’)
BLR 6213	*jíkʊ́	‘porcupine’	> *Ø-ʒíʁú*
BLR 4574	*bʊ́dʊ́gʊ́	‘dwarf antelope’	> *Ø-mblúʁú* ‘antelope sp.’ (cf. Kongo *mbùlúkù* ‘dwarf antelope’)
BLR 2368	*pàkàcà	‘buffalo’	> *Ø-mpàkàsà* (cf. Kongo *mpàkásà*)
BLR 2967	*tòk	‘boil up’	> *tòkìsà* (cf. Kongo *tòkísà*)
			> *àvòká* (French *avocat*, probably via Kongo *àvòká*)

As can be seen in (11), all borrowings in Ngwi preserve the final vowel. But there are also other features which make them easily detectable. First, non-nativized borrowed items have /k/ as a reflex of PWCB *k in C2, e.g., *Ø-mpàkàsà* ‘buffalo’. In partially nativized borrowings the original /k/ shifts to /ʁ/, the most common

---

[14] Thanks to a dedicated study (Pacchiarotti & Bostoen 2021), we were aware of similar easily identifiable borrowings in all homeland varieties which underwent systematic final vowel loss, namely Yans B85, Mpur B85e, Nsong B85d, Nsambaan B85F, Ding B86, Lwel B862, Mpiin B863, Ngong B864, and Nzadi B865.

[15] Although we report the tonal notations in Swartenbroeckx (1973), these should not be taken at face value. The author claims to note High tone as [á] and low tone as [à]. However, throughout the dictionary one also finds [â] as well as vowels without any tonal notation. Judging from the discussion in Swartenbroeckx (1973: viii–x), there might be some conflation of tone and tonic accent in his suprasegmental transcriptions. He himself admits that these notations should be improved by tone specialists in the languages of the area such as his fellow-Jesuit Jan Daeleman.

[16] The speaker is aware of a native word, i.e., *ò-fûŋ*, which is a reflex of BLR 5111 *cʊ̀ngʊ̀ ‘Graminaceous spp.’

reflex of PWCB *k in C_2, e.g., *ì-tòʁó* 'mat'. None of the borrowings in (11) displays the zero reflex of PWCB *k found in (10). Second, borrowed words usually do not display the tonal innovations typical of native words. For example, PB *LL became mostly HL in Ngwi, see e.g., BLR *dègè 'weaver bird'> B861 *è-lɛ̂ʁ*; BLR 1607 *jògù 'elephant' > B861 *Ø-ndzɔ́ɔ̀* in (10). In borrowed words, the inherited PB *LL(L) pattern was maintained as such, see e.g., *ò-kɔ̀kɔ̀* 'sheep' < BLR 1905 *kòòkò.

Now that we have established that Ngwi words having zero as the reflex of PWCB *k are definitely not recent borrowings from the region's vehicular language Kongo ya Leta, we should also exclude the possibility that they might be loanwords from neighboring WCB languages. As Map 1 clearly shows, Ngwi is the only language in the wider homeland region to have a dorsal fricative as the predominant reflex of PWCB *k. In all others, the zero reflex prevails. Hence, Ngwi words manifesting the zero reflex may well be borrowings from one of these languages. We address this possibility based on the comparative data in Table 1, where we show the Lwel B862, Nzadi B865, and Ding B86 translation equivalents of the Ngwi words in (10) with a zero reflex of PWCB *k. Those three languages are spoken in the same territories as Ngwi (see Map 1) and all have zero as the most common reflex of PWCB *k in C_2 position. Note that even though we give in the first column of Table 1 the protoform from which a given Ngwi reflex originated, translation equivalents in the other languages are not necessarily cognate. For instance, while Nzadi *ò-dzwó* 'snake (sp.)' is in all likelihood also a reflex of BLR 3536 *jókà 'snake', Lwel *n-tààl* and Ding *n-tɛ̀ɛ̀l* are reflexes of BLR 2733 tààdí 'snake'. A question mark in Table 1 means lack of data.

Table 1. Equivalents of selected Ngwi lexemes with a zero reflex in neighboring languages

	Ngwi B861	**Lwel B862**	**Nzadi B865**	**Ding B86**
BLR 7413 *cókì 'saliva'	*à-súì*	*mə̀-tyɛ́*	*à-tɛ́*	*mà-tɛ́*
BLR 9461 *cákú 'safou'	*è-súú 'safou (fruit)'*	*?*	*ò-té ó mfùŋ*	*lu-say*
BLR 3536 *jókà 'snake'	*Ø-ndʒúà*	*n-tààl*	*ò-dzwó*	*n-tɛ̀ɛ̀l*
BLR 5339 *tʊ́kì 'insult'	*ì-tíì*	?	*è-pwɔ̌n*	*mu-tsœœ*
BLR 3050 *tʊ́k 'to insult'	*túyè*	*tʃwɛ́*	*ò-twâ*	*ò-sààr*
BLR 761 *cúgù 'day'	*è-ʃúù*	*lə̀-ʃú*	*è-súù*	*è-tý́*
BLR 2642 *pʊ́kʊ̀ 'mouse'	*Ø-mfúù*	*m-pú*	*m-púù*	*m-pú*
BLR 1607 *jògù 'elephant'	*Ø-ndzɔ́ɔ̀*	*n-dzòò*	*n-dzɔ̀ɔ̀*	*n-dzòò*
BLR 1179 *dʊ́k 'vomit'	*lúà*	*líír*	*ò-lwâ*	*ò-lwá*
BLR 1685 *kààká 'grandparent'	*Ø-ŋkíáŋkíá* 'paternal aunt'	*ŋ̀-kàá*	*ŋ-kàá*	*ŋ-kǎ* 'aunt'

We now discuss the likelihood of borrowing for each of the Ngwi words in Table 1. The first four cases stand out as immediately discardable: *à-súì* 'saliva', *è-súú* 'safou (fruit)', *Ø-ndʒúà* 'snake', and *ì-tîì* 'insult'. In the case of *à-súì* 'saliva', all three neighboring languages have a reflex of a different protoform for the same concept, namely BLR 6231 *tâì 'saliva, spittle'. As for *è-súú* 'safou (fruit)', only Ding *lu-say* comes from the same protoform, but the two cognates do not look alike. Similarly, Nzadi *ò-dzwó* is in all likelihood cognate with Ngwi *Ø-ndʒúà*, but the two forms have a different initial consonant (/dz/ vs. /dʒ/, the latter preceded by an erstwhile fossilized nasal noun class prefix in Ngwi), a different final vowel (/a/ vs. /o/), and a different root tone pattern (HL vs. H). As for *ì-tîì* 'insult (n.)', Lwel and Nzadi have a reflex of BLR 3050 *tʊ́k 'to insult', *tʃwɛ́* and *òtwâ* respectively, but for the verb form. Nzadi uses the reflex of a different protoform for the noun 'insult'. We unfortunately lack this information for Lwel. In Ding, the noun *mu-tsœœ* is also derived, just like in Ngwi, from BLR 5339 *tʊ́kì 'insult': the long /œ/ vowel is the result of umlaut effects triggered by V2 *ì on V1 *ʊ́ in *tʊ́kì (Bostoen & Koni Muluwa 2014). On the other hand, the verb to 'insult' in Ding is a reflex of a different protoform. Thus, the only real cognate of Ngwi *ì-tîì* is Ding *mu-tsœœ*, but these two forms hardly bear any formal similarity. Lwel *tʃwɛ́* and Nzadi *òtwâ* are also cognate with Ngwi *túyè* 'to insult'.[17] In this case too, the degree of formal similarity among these three cognate forms does not support a borrowing scenario.

The remaning forms in Table 1 require a more in depth discussion. The Ngwi form *è-ʃúù* is not a borrowing from Ding *è-tý* (likely a reflex of BLR 3156 *túkù 'day of 24 hours') as suggested by the lack of formal similarity between the two. However, the words for 'day' in Nzadi, Lwel, and Ding are clearly cognate with Ngwi *è-ʃúù*. It is unlikely that *è-ʃúù* is a Nzadi borrowing. While in Nzadi PB *c in C1 position has /s/ as a reflex (Crane, Hyman & Tukumu 2011: 256–257), in Ngwi the same protosound has two phonologically conditioned reflexes, /s/ when PB *c was followed by a non-high vowel (namely *a, *o, *e) and /ʃ/ elsewhere, e.g., BLR 394 *càbʊk 'cross river' > *sǎβ*, BLR 631 *còbó 'intestines' > *ò-sɔ̌β*, BLR 508 *cèd 'be slippery' > *sɛ̂r* vs. BLR 5110 *cʊ́ngʊ́ 'tree, bark' > *ò-ʃúŋ* 'tree', BLR 604 *cíd 'be finished' > *ʃîr* 'to finish'.[18] Thus, if Ngwi had borrowed the word for 'day' from Nzadi we would expect it to have a root-initial /s/ and not /ʃ/. While it is possible that /ʃ/ is

---

[17] The <y> (IPA [j]) in *túyè* might be epenthetic, but its conditioning is still unclear. Besides *túyè* 'insult', we only find it in the following lexical items: *búyè* 'break' (< BLR 372 *búg 'break, snap'), fúyè 'grow up' (< BLR 1997 *kʊ́d 'grow up'), *Ø-mvúyè* 'rain' (< BLR 368 *búdà), *víyè* 'invite, call' (< BLR 177 bíd 'call, announce'), as well as *kùyé* 'besmear, polish', *ò-ntúyè* 'mushroom (sp.)', and *ì-pfúyé* 'flea' for which no protoform is available.

[18] Many "exceptions" exist to this conditioning environment, that is, PB *c in C_1 also shows MUR in Ngwi.

a nativization strategy, there are several Ngwi seemingly native words with /s/ followed by /u/ (e.g., *ò-sûɲ* 'meat' < BLR 3778 *cùnì 'meat, flesh', *sú* < BLR 9826 *ì-cú-è '1PL') which speak against this hypothesis. It is much harder (if not impossible) to exclude with certainty the possibility that *è-ʃúù* is not a borrowing from Lwel *lə-ʃú*. If Ngwi speakers had borrowed this word from Lwel speakers, they must have done so at a time when Lwel had not shortened CVV stems historically derived from *CVCV stems to CV, i.e., *cúgù > *ʃúù* > *ʃú*. Otherwise, it would be impossible to account for the HL tone pattern on *è-ʃúù* considering that the present-day Lwel form has a H tone only, at least based on the only available phonological account of Khang Levy (1979). The same reasoning holds for *Ø-mfúù* 'mouse': if this word had been borrowed from Lwel or Ding, the loan event must be old, due to the fact that apparently both languages shortened an erstwhile form formally equivalent to Nzadi *mpúù* to *mpú*. That *Ø-mfúù* is a borrowing from Nzadi *m-púù* is unlikely due to the presence of a prenasalized fricative /mf/ in Ngwi instead of the expected Nzadi /mp/. In Ngwi, noun roots and stems with an initial /mp/, /mf/, and even /mpf/ followed by /u/ are all attested in native words, e.g., *m-pûy* 'skins' (< BLR *pʊ̀cʊ̀), *Ø-mpúŋ* 'eagle' (< BLR 2657 *pʊ́ngʊ́), *Ø-mfûr* 'land turtle' (< BLR 2108 *kúdʊ̀), *Ø-mfŭr* 'bird' (< BLR 3962 *pùdú), *m-pfŭɲ* 'stomachs' (< BLR 1545 *kùndú), *m-pfûʁ* 'doors' (< BLR 5467 *kúk 'cover'). This makes it unlikely that the word for 'mouse' got borrowed from Nzadi as *m-púù* and was then nativized to *Ø-mfúù*, first because native Ngwi words attest both /mp/ and /mf/ followed by /u/, and second because of Bantu Spirantization (Schadeberg 1995; Bostoen 2008). This sound change, which transforms stops into fricatives or affricates when they are followed by reflexes of PB *i and *u, happened only very occasionally in the history of Ngwi and yielded /pf/, not /f/, e.g., *ò-pfŭɲ* 'burial' (< BLR 2125 *kùnd 'to bury'), *è-pfûy* 'calabash' (< BLR 5395 *gútù), *ò-pfûy* 'fart' (< BLR 3959 *pùdì). Moving on to *Ø-ndzɔ́ɔ̀*, the only feature which seems to exclude borrowing is the HL tone pattern on the Ngwi reflex of BLR 1607 *jògù 'elephant'. As discussed earlier on in this section and in Section 3.3, PB LL noun stems evolved into HL in Ngwi, see e.g., *ì-lâʁ* 'promise' < BLR 814 *dàgò, *è-lɛ̂ʁ* 'weaver bird'< BLR 900 *dègè, *ì-pɛ́àʁ* 'shoulder' < BLR 2433 *pègà in (10), but also *è kúù* 'leg' < BLR 1490 *gʊ̀dʊ̀, *ò-lúù* 'bitterness' < BLR 1168 *dʊ̀dʊ̀. On the other hand, borrowed words with a LL tone pattern are preserved as such, see examples in (11). While tonal nativization cannot be excluded with certainty, we do not have any examples of this alleged phenomenon in our database.

When it comes to *lúà* 'vomit', the formal similarity of the Ngwi reflex with its Nzadi and Ding cognates *ò-lwâ* and *ò-lwá* respectively is striking. We have evidence that elsewhere in Ngwi, sequences of two vowel nuclei such as /u/ and /a/ as in *lúà* 'vomit' arose, among others, from the reanalysis of diphthongs into sequences of two full vowels (see Pacchiarotti, Maselli & Bostoen 2021 for synchronic phonological evidence in favor of this analysis). This process affected words which lost a consonant in C_2 position. In these contexts, whenever a long vowel was created, it then broke into a diphthong which in turn crystallized into a sequence of two vowel nuclei, e.g., BLR 6882 *jòbó > *dzòó* > *dzwǒ* > *Ø-dzùó*, BLR 893 *ndédé 'white man' > *ndɛ́ɛ́* > *ndyɛ́* > *ò-ndíé*. Two vowel nuclei were also created off of historical *CVV, e.g., *ò-kúá* 'salt' < BLR 1521 *gúá, *Ø-mvúá* 'dog' < BLR 282 *bʊ́à. While these data do not exclude the possibility that Ngwi *lúà* 'vomit' might be a borrowing from either Nzadi *ò-lwâ* and Ding *ò-lwá* where the diphthong was reanalyzed as a sequence of two vowels, they do not offer evidence in support of the borrowing hypothesis either. This is because the process described above happened consistently in the history of Ngwi to words which underwent (irregular) loss of a PB C_2, namely *b, *d, *c, and *j.

Lastly, *Ø-ŋkíáŋkíá* 'paternal aunt' is also a difficult case. At first sight, it seems unlikely that the Ngwi form might have been borrowed from one of its three cognate forms in Lwel, Nzadi, and Ding. In all three languages, reflexes of BLR 1685 *kààká 'grandparent' are regular in that: (i) they have zero as a reflex of PWCB *k in C_2; and (ii) they preserve the LH tone pattern of the reconstruction. The Ngwi reflex *ŋkíáŋkíá* is a full reduplication of BLR 1685 *kààká with a fossilized historical class 9 homorganic nasal noun class prefix N- reanalyzed as part of the simple noun stem. We have evidence that whenever a long vowel [aː] was created in earlier stages of Ngwi due to the loss of a C_2 consonant in the language, it underwent palatal on-glide diphthongization and became /ya/. This diphthong in turn was reanalyzed as a sequence of two vowel nuclei as discussed in the preceding paragraph, e.g., *kààká > *kààá* >*kàá*> *kyǎ* >*kìá*. Similar examples are: BLR 1662 *kádà 'charcoal' > *ì-kíà*, BLR 1294 *gádà 'fingernail' > *è-kíà*, BLR 406 cádá 'feather' > *è-síà*, BLR 1557 *jàdà 'rubbish heap' > *Ø-dʒìà*, BLR 1555 *jàdà 'hunger' > *Ø-ndzìà*. However, the HH tone pattern of *Ø-ŋkíáŋkíá* is irregular and could indeed point towards borrowing. In Ngwi, *LH as in BLR 1685 *kààká is usually preserved, cf. BLR 2634 *pʊ̀gí 'pot' > *Ø-mpǔʁ* in (10), or BLR 4570 *bʊ̀dʊ́ 'swart, pimple' > *è-bùú*. Nevertheless, suprasegmental MUR are also common in Ngwi and not only for reflexes of reconstructed *LH nouns (see Philippson 1999 for a discussion of what he calls tonal double reflexes in Bantu languages). If we posit that *Ø-ŋkíáŋkíá* was borrowed from Nzadi, Lwel, or Ding, it must have been borrowed with the LH tone pattern and then have shifted to HH. The fact that the noun stem was fully reduplicated might have played a role in the development of an irregular tone pattern. The only other fully reduplicated noun stem in our database is *ì-ŋkìàŋkìà* 'crow' which has an unusual LLLL tone pattern.[19] What-

---

[19] Partial stem reduplication is much more common in Ngwi. In this type of reduplication, the first syllable of the simple noun stem is reduplicated and preserves its tone.

ever the case might be, while the possibility of borrowing cannot be excluded for *Ø-ŋkíáŋkíá*, it is also possible that this word might have been created language internally, by analogy with noun stems which lost a consonant in word medial position and were reconstructed with *a in both V_1 and V_2 positions such as BLR 1662 *kádà 'charcoal' > *ì kíà*.

As we have shown in this section, most zero reflexes of PWCB *k in Ngwi cannot be explained by resorting to borrowing from neighboring languages. This of course does *not* mean that no lexical item displaying MUR in our dataset can be explained via borrowing. But at least in Ngwi, borrowing does not serve as an explanation for several of the words with a zero reflex of PWCB *k in C_2.

4. Discussion

By looking at the evolutions of PWCB *k in C_2 discussed in Sections 3.1–3.2 and the argumentation in Sections 3.3–3.4, it is hard to maintain the Neogrammarian view of pervasive regularity in sound change. The MUR in Map 1, sometimes up to three for one and the same proto-sound, do not have different phonological or grammatical conditionings. In other words, it is impossible to provide a Neogrammarian explanation for their synchronic co-occurrence as the outcome of a single PWCB phoneme. In earlier Bantu scholarship, this kind of double (or more) consonant reflexes were accounted for either by positing the existence of two series of PB consonants (cf. Van Leynseele & Stewart 1980; Gerhardt 1986; Hedinger 1987; Bancel 1988; Stewart 1989, 1993) or by supposing that the original environment conditioning the split of a single reconstructed consonant became synchronically opaque (e.g., Guthrie 1967: 58; Bachmann 1989; Miehe 1989; Blanchon 1991; Janssens 1991, 1993). Although they advocated for two mutually exclusive diachronic accounts to explain divergent synchronic reflexes, i.e., phonemic merger vs. phonemic split, scholars of both convictions ardently stick to the Neogrammarian hypothesis. They failed to admit that irregularity in sound change can occur under specific sociocultural circumstances without necessarily jeopardizing the use of the Comparative Method as such for reconstructing linguistic prehistory. Our study of MUR of PWCB *k in C_2 shows that irregular sound change is the rule, not the exception. As Map 1 shows, WCB languages having a unicolored pie chart, i.e., only one unconditioned reflex of PWCB *k in C_2, are rare. This pervasiveness of MUR urges us to accept irregular sound change as a historical-linguistic reality and to search for underlying historical sociocultural scenarios which may account for this synchronic outcome. In this section, we consider two traditional explanations that fall within the broad paradigm of language contact, but in structurally different ways. In Section 4.1, we consider lexical diffusion, which is traditionally seen as contact-induced innovation within a speech community. In Section 4.2, we assess substrate influence, which is traditionally seen as an external contact-induced influence on a community's language through the inclusion of foreign speakers shifting to that language. In Section 4.3, we argue that these two accounts are not necessarily mutually exclusive and that the prehistoric linguistic landscape of Bantu-speaking Africa was possibly so multilingual that distinguishing contact between speakers from one and the same community from contact between speakers of different speech communities presumes a too neat one-to-one correspondence between linguistically defined units (i.e., languages) and socially defined units (i.e., peoples). In Section 4.4, we discuss interdisciplinary evidence suggesting that the history of Bantu languages is not tree-like but rather stratified, consisting of multiple layers of migratory events over time. We suggest that MUR could also be a result of this historical composition.

4.1 Lexical diffusion

Based on empirical evidence from several sound changes in Philadelphian English, Labov (1981) argues that two types of sound change can coexist at different levels within a variety. Low-level output changes such as raising, lowering, fronting and backing tend to show Neogrammarian regularity, that is, sound change is phonetically gradual in that it procedes in imperceptible increments and lexically abrupt in that it affects all lexical items at once. On the other hand, more abstract phonological changes such as vowel split show ongoing lexical diffusion type of change, that is, sound change is phonetically abrupt but lexically gradual. Labov (1981) claims that sound changes such as the split of a long vowel into a short and a long raised one are abstract in the sense that they often involve a whole set of phonetic features for their realization (such as length, height, fronting, directions and contours of glides, etc.). In the case of consonantal changes, lexical diffusion appears to be common in changes of place of articulation, while changes in manner of articulation would usually display Neogrammarian regularity (Labov 1981: 302–303). However, it seems to us that changes in manner of articulation also involve numerous subchanges, for example a different position of the velum, a different constriction of the oral cavity to produce a different degree of aperture, movements of the glottis, etc. Although certain studies did indeed show that shifts in place of articulation tend to spread gradually through the lexicon (cf. Kinkade 1973; Li 1982), several others provided evidence for the sensitivity of lenitions, which involve a shift in mode of articulation, to lexical diffusion conditioned by frequency (cf. Phillips 1984; Bybee 2001, 2002; Phillips 2006; cited in Garrett 2015).

In the case of the evolutions of PWCB *k in C_2, a change such as *k > ʁ, implies a set of subchanges in manner of articulation, place of articulation, voicing and duration. Since many subsystems are affected by the innovation, one would expect to find lexical diffusion type of change, where not all /k/ in a given language have become /ʁ/ yet, a snapshot of sound change caught in the process of happening. In languages where all /k/ became some sort of fricative and then some fricatives became zero, we also capture an ongoing second cycle of change. For Labov (1999: 542), the deletion of obstruents would be a typical instance of a lexical diffusion type of sound change, along with the shortening and lengthening of segments, diphthongization of mid and low vowels, the metathesis of liquids and stops, and consonant changes in place of articulation. Such sound changes would result from "the abrupt substitution of one phoneme for another in words that contain that phoneme. The older and newer forms of the word will usually differ by several phonetic features. This process is most characteristic of the late stages of an internal change that has been differentiated by lexical and grammatical conditioning, has developed a high degree of social awareness, or shows borrowings from other systems ("change from above")" (Labov 1999: 542). The lexical diffusion type of change contrasts with regular sound change "resulting from a gradual transformation of a single phonetic feature of a phoneme in a continuous phonetic space" (Labov 1999: 542). This type of "change from below" usually happens in the early stages of a change within the system, without any kind of conditioning or social awareness. Vowel shifts in place of articulation, diphthongization of high vowels, consonant changes in manner of articulation, vocalization of liquids and deletion of glides and shwa would typically fit into this category.

Unfortunately, our study cannot take into account any sociolinguistic variables, such as age, sex, social status, etc., which might have influenced the existence of MUR of PWCB *k in different present-day WCB languages, because we are dealing with sound changes for which no empirical historical evidence is available. However, it is striking that the initial sound shift which we take as a point of departure, i.e., the merger of PB *g and *k, was fully regular and affected the entire lexicon of PWCB to the extent that it can be considered diagnostic for the genealogical unity of this Bantu branch (Pacchiarotti & Bostoen 2020). Intervocalically, WCB languages manifest no trace whatsoever of PB *g as distinct from PB *k.[20] The shift from PB *g to PWBC *k required the change of a single pho-

---

[20] An anonymous reviewer points out that in the Nzadi B865 grammar by Crane, Hyman and Tumuku (2011), one reads "as with the other stops, *k remained *k* in C_2 position, while *g alternatively devoiced (very rarely) or disappeared (much more common)" (Crane et al. 2011: 261).

netic feature, i.e., voicing. However, all later transformations of PWCB *k, i.e., fricativization and deletion, were phonetically much more complex in that they required the mutation of several phonetic features at once. These further sound changes might have been more prone to lexical irregularity in the sense that word frequency could have played a role in whether a lexical item was affected by the innovation or not.

The fact that certain MUR patterns, such as k/Ø and k/dorsal fricative(/Ø), are geographically clustered and neighboring languages manifest variable ratios in the lexical distribution of certain reflexes suggests that the diffusion of a given phonological innovation across the lexicon did not only happen within a given speech community. Conceivably, it also spilled over to neighboring communities speaking closely related varieties. In this process, sometimes called dialect borrowing, it is difficult to say whether a phonological innovation expands because words having undergone it are copied from a more innovative variety into a more conservative one (Thomason 2007: 43), or rather due to the spread of a (more prestigious) pronunciation norm (Campbell 1998: 191). In any event, the pervasivennes of MUR across WCB and the geographical clustering of certain specific MUR patterns seem to indicate that after the initial expansion of WCB, interaction between closely related varieties gradually intensified and stimulated sociolinguistic processes such as dialect borrowing leading to the uneven contact-induced spread of further evolutions of PWCB *k. This phenomenon affected only a part of the speech community and only a portion of the lexicon. The innovation of PWCB *k therefore appears to have proceeded "in a quasi-wavelike fashion along the social and geographic dimensions of the speech community, and through the linguistic system itself" (Harrison 2003: 220).

4.2 Substrate influence

While the geographic distribution of specific MUR patterns points to language contact processes such as lexical diffusion and dialect borrowing, the geographic

---

Then, the following examples are given: BLR 2433 *pègà 'shoulder'> *i-pek*, BLR 1248 *dúg 'paddle' > *o-dwâ*, BLR 1621 *jʊ̀gʊ́ 'groundnut' > *e-dzuu*, BLR 3423 *jígu 'hear' > *o-zwâ*. However, the claim that PB *k in C_2 position remained /k/ is not supported by any data in Crane et al. (2011). Rather, all reflexes of PB *k in C_2 found in Crane, Hyman and Tukumu (2011: 258–260) show that PB *k in C_2 became zero: BLR 3536 *jókà > *o-dzwó* 'snake', BLR 1044 *dìɪk > *o dzya*, BLR 5304 *bʊ́kʊ̀ > *e-bvûu* 'fish (sp.)', BLR 67 *bák > *o-bâa* 'get'. As can be seen from the 66 cognate sets in Pacchiarotti and Bostoen (2020: 175–189), while most PB *k (and *g) in C_2 became zero in Nzadi, some were retained as /k/, e.g. BLR 429 *càkʊd 'to weed' > *o-sakul* and BLR 2433 *pègà 'shoulder' > *i-pek*. Thus, the Nzadi data offers evidence in support of the merger of PB *k and *g to /k/ in both C_1 and C_2 at PWCB level.

clustering of certain specific reflexes of PWCB *k, whether part of a series of MUR or not, could also be the outcome of another type of contact-induced change. As we argued in Section 3.2, the current-day geographic distribution of fricative reflexes of PWCB *k, especially in the northwestern part of the WCB domain, where B40-B70 belonging to different generalogical subgroups are in contact, can best be accounted for as an areal feature. Perhaps the scattered presence of fricative reflexes of PWCB *k throughout WCB, including the homeland area (i.e., in Ngwi B861), could be considered as parallel innovations with an areal bias. Outside of the northwestern corner of the Bantu-speaking area (i.e., Cameroon, Gabon, the Republic of the Congo, and the DRC), fricative reflexes of PB *k and *g – whether velar [x]/[ɣ], uvular [ʁ] or glottal [h] – are rare, as can be seen from the maps in Guthrie (1967: 74–75).

In the northwestern part of the Bantu domain, apart from WCB, they are attested in several of Guthrie's B10–30 languages from Gabon, which belong to the North-Western Bantu (NWB) branches (cf. Grollemund et al. 2015) and are spoken in close proximity of Guthrie's B40–60 WCB languages. In B10 and B30, [ɣ] occurs as an intervocalic reflex of PB *k and *g (cf. van der Veen 1991, 2003; Mouguiama-Daouda & van der Veen 2005: 110). In B20, [ɣ] is the only velar fricative sporadically attested (in less than ten words in each of the following varieties) as the phonetic realization of /g/ in Ntumbəde B22d, Wumbvu B24, Siwu B202, Saké B251, and Ndasa B201 (see Mokrani 2016: 189, 263, 320, 381, 389). Velar and glottal fricatives ([ɣ] and [h] respectively) as reflexes of PB *k and *g are also attested in North-Western Bantu (NWB) languages spoken in Cameroon, e.g., Bubi A31 and Bafia A53 (Janssens 1993: 25 and ff., 144 and ff.). Within the Central-Western Bantu (CWB) branch (cf. Grollemund et al. 2015), the voiced velar fricative is attested in some varieties of Mboshi C25 spoken in the Republic of the Congo (i.e., Akwa, Koyo, Mboko, and Ngare), where [ɣ] is the phonetic realization of /g/ intervocalically, while in other varieties [ɣ] is realized as zero (Ndongo Ibara 2000: 40), and the Sakata cluster C34 in the DRC (cf. De Witte 1955: 81–141). These two CWB languages border with WCB varieties. Otherwise, fricative reflexes of PB *k and *g are absent from CWB languages (cf. Grégoire 2003), except maybe for Guthrie's Zone C languages spoken by hunter-gatherer communities. According to Motingea Mangulu (2021), *k >x > h > ʔ >Ø would be a feature shared by the so-called "Pygmy" languages from the Central Congo Basin. This claim needs further documentation and research.

Considering that: (i) NWB, CWB, and WCB are different branches of the Bantu language family tree; and (ii) dorsal fricatives are attested in a scattered fashion in these branches, it is impossible to posit their development as a single innovation which happened at an ancestral node uniting these three western branches before they split off. Nevertheless, given the overall rarity of fricatives as reflexes of PB *k and *g outside of the northwestern Bantu-speaking area and their reported occurrence in several hunter-gatherer Bantu languages, their relatively high frequency within WCB (especially in the northwestern part), as well as NWB and CWB immediately bordering them, could be considered an areal feature originating in substrate influence from no longer spoken non-Bantu languages of extinct autochthonous hunter-gatherer groups. This hypothesis cannot be substantiated from a linguistic point of view, as no descendants of the languages spoken by Central African hunter-gatherers before they shifted to Bantu languages have survived the ravages of time. However, several genetic studies have shown that ancestral Bantu speakers did admix with indigenous groups and that this exogamy was biased towards the integration of hunter-gatherer women into the communities of Bantu-speaking settlers (Destro-Bisol et al. 2004; Wood et al. 2005; Quintana-Murci et al. 2008; Batini et al. 2010; Verdu et al. 2013; Patin et al. 2014, 2017). Although ethnographic data on such mixed marriages are largely missing – let alone historical data (cf. Bahuchet 2012: 36), children of double descent probably grew in the Bantu-speaking village of their father, while also acquiring the non-Bantu language of their mother. Such intermarriages not only led to the cohabitation of men and women speaking different and unrelated languages, but also generated bilingual children. Depending on the number of speakers and speech communities involved and the social hierarchies ruling between them, this bi/multilingualism may have continued in the society during several generations, possibly with complementary functions for the different languages. When this stable bi/multilingualism eventually collapsed, it resulted in complete language shift towards the dominant Bantu language and in the death of the original non-Bantu language(s) of the mothers, at least within that specific speech community (cf. Winford 2003: 27). Contact-induced phonological change happens more readily in the case of language shift under the imposition of the sound system of the original language (Van Coetsem 1988; Winford 2003: 377) than in situations of superficial contact between communities without widespread bilingualism (Winford 2003: 55). In this regard, it is not unlikely that the geographical clustering of dorsal fricatives (whether or not as reflexes of PB *k and/or *g) in the northwestern Bantu languages of the Central African rainforest is indeed an areal feature that rose as a parallel innovation through substrate influence. This hypothesis finds partial support in a recent genetic study targeting DRC individuals speaking WCB varieties addressed in this article. As discussed in Section 3.2 and shown in Map 1, dorsal fricatives are pervasive in the northwestern part of the WCB domain, where they are found in the Nzebi-Teke West (B50+B73b-c), Mbete (B60), and the KLC B40 languages. Fortes-Lima et al. (2021) show that amongst the sampled populations, individuals speaking Nzebi B52, Shira B41, and

Punu B43 show high degrees of genetic admixture with western rainforest hunter-gatherer groups.[21]

4.3 Intrinsic multilingualism as the norm

It is clear from Sections 4.1 and 4.2 that the pervasiveness of irregular sound change in Bantu-speaking Central Africa is to a large extent determined by the region's specific "ecology of language evolution" (Mufwene 2001), namely the impact of social factors in particular and extra-linguistic factors more generally. It has long been acknowledged that linguistic change is inevitably speaker-based and that communal languages are only abstract extrapolations from idiolects (cf. Milroy 1992). However, in Central Africa, even what are commonly considered to be distinct (Bantu) languages are mostly abstract – not to say artificial – extrapolations or consolidations of very time-specific and often outsider intepretations of complex linguistic realities. Inventories of African languages and glossonyms mostly date back to the time of European colonialism and are often based on the wrong assumption that socially or ethnically defined groups speak distinct languages (cf. Fabian 1986; Van de Velde 1999). Very often one and the same language was spoken by different social groups, such as Tutsi and Hutu in the precolonial kingdoms of Rwanda and Burundi (whose languages can actually also be considered as regiolectal varieties), while large precolonial polities, such as the Kongo kingdom, were multilingual (cf. Mufwene 2001: 178; Bostoen & de Schryver 2018b). As a matter of fact, multilingualism was and continues to be the rule rather than the exception in (Bantu-speaking) Africa, as Schadeberg recognizes: "Bantu speakers have long lived in a multilingual continuum, where many speakers master not just their own variety of speech but also those of their neighbours. Linguistic differentiation and convergence are actively pursued, one serving to establish distinct group identities, the other one to forge alliances and to foster good neighborship. The almost wilful selective adoption of new features is facilitated by structural similarities between Bantu languages" (2003: 158).

Such a situation of "prolonged socio-economic intercourse amongst small-scale (genetically related) linguistic communities" is exactly what Harrison (2003: 230 ff.) considers to be a great adversary to the Comparative Method, because it makes contact-induced transfers between languages undetectable, especially if there has been a massive diffusion of lexical items across a multilingual domain. With reference to the linguistic landscape of New Caledonia, Grace

---

[21] On the other hand, speech communities in the WCB homeland such as Ngwi which also show dorsal fricatives manifest low degrees of admixture with documented western rainforest hunter-gatherer groups.

(1996) argues in the same vein that regularity of sound change depends on particular sociocultural monolingual conditions, which are rather specific to Europe in recent history, i.e., a perfect overlap between the linguistic communities and linguistically defined languages or dialects. In linguistic communities where multilingualism is the norm, a linguistic repertoire includes resources from several languages. In pre-colonial times, there were probably no communities defined on the basis of one language in New Caledonia. Even though one language was spoken in more than one village, there was probably no sense of community among those villages speaking that same language. Grace hypothesizes that "the aberrancy (which is to say the relative intractability by the Comparative Method) of the languages of southern New Caledonia comes precisely from such a lack of correspondence between the languages themselves and the associated linguistically relevant communities" (1996: 175). Absolutely regular sound change would rather be surprising in a context in which "linguistically relevant communities (networks of intercommunicating speakers) fail to correspond to what according to strictly linguistic criteria are individual languages (or dialects)" (Grace 1996: 175).

When multilingualism is the norm, languages are constantly in contact and boundaries between languages – if existent at all – are necessarily highly fluid. Under such circumstances, it is rather hard to distinguish between lexical diffusion as a contact-induced process within a speech community and substrate influence as the outcome of language contact between distinct speech communties. The opposition between language-internal and language-external becomes rather fuzzy and communal norms with regard to pronunciation and ongoing sound change are rather liberal. In the absence of a notion of "correct" or in-group-specific articulation, synchronic phonological variability eventually resulting in irregular sound change is to be expected. Under such conditions, the Comparative Method loses effectiveness, especially when it comes to defining subgroups based diagnostic shared innovations. However, it does not become entirely disfunctional. It contributes to identifying regular patterns in irregular sound change and undoing them for the sake of linguistic reconstruction. One such example are reconstructed forms manifesting phonological variability in the BLR database. Reconstructions with the same meaning but slightly different forms, such as BLR 2568 *pígò 'kidney' vs. BLR 2587 *píjò 'kidney' vs. BLR 6234 *píkò 'kidney', are known in historical Bantu studies as "osculant" (Guthrie 1962, 1967). Sometimes a more thorough knowledge of certain sound changes allows one to reduce phonologically osculant reconstructions to a single proto-form (cf. Bostoen 2001; Ricquier & Bostoen 2008). However, very often osculant reconstructions need to be maintained as they are, because sound correspondences between Bantu languages turn out to be not fully regular and systematic. In such cases, osculance points either to a prehistoric language contact event (Bostoen & Bastin 2016) or to a sound change that started off but did not affect the whole lexicon of an ancestral Bantu language (see Section 4.1). The fact that certain sound changes were not fully systematic across all relevant lexical items explains why sound correspondences between present-day Bantu languages can often not be reduced to a single proto-form from which all present-day reflexes can be derived through Neogrammarian-like sound change.

4.4 Spread-over-spread events in Bantu language history

Current-day Bantu languages are commonly seen as descending directly from the ancestral languages of the first Bantu-speaking settlers in a given region. Likewise, language phylogenies of current-day Bantu languages are interpreted as reflecting the original migration of Bantu speech communities (cf. Currie et al. 2013; Grollemund et al. 2015). Although migrating Bantu speakers definitely spread their languages from the Bantu homeland in the Cameroonian-Nigerian borderland to Eastern and Southern Africa, their present-day geographic distribution does not necessarily reflect the original migration of Bantu speech communities. Analyses of human Y-chromosomal DNA from modern-day Bantu speakers indicate that several expansion phases succeeded each other, obliterating any possible founder event and camouflaging the genetic signature of the first migration (see, for instance, de Filippo et al. 2011; Pour, Plaster & Bradman 2013). Archaeological data also encourage us to adopt a spread-over-spread model for Bantu language history rather than reconstruct the dispersal of Bantu speakers and their languages as a single, long-term, continuous macroevent. Seidensticker et al. (2021), for instance, argue that two periods of more intense human activity are separated in the entire Congo rainforest by a population collapse between 400 and 600 CE followed by major resettlement centuries later. A similar disconnection between Early and Late Iron Age, especially in terms of pottery production, has been observed in the archaeological record of Bantu-speaking Southern Africa (cf. Mitchell & Whitelaw 2005; Huffman 2007; Mitchell 2013; Schoeman 2013; Loftus, Mitchell & Ramsey 2019). Such spread-over-spread events must have led to recurrent instances of Bantu-internal language shift, i.e., communities that abandon their ancestral Bantu language in favor of another (socially more successful) Bantu language (cf. Bostoen 2018). Today, this sociolinguistic process is known to be one of the principal causes of language death in Central Africa and elsewhere and must also have been common in prehistoric times. Attempts to disentangle successive strata of Bantu language history have been rare (cf. Möhlig 1977, 1979, 1981a, 1981b; Masele & Nurse 2003; Bostoen 2007). Moreover, their implications for the historical linguistic method have never been fully assessed.

We believe that apart from synchronic multilingualism, which has been characteristic of Bantu speech communities through time, diachronic spread-over-spread events as the ones discussed above contributed to the emergence of pervasive MUR patterns in most present-day WCB languages. The fact that irregularity of sound change is so common in WCB languages may indicate that rather than having evolved from a single ancestral language, these varieties might be the complex product of multiple superimposed historical layers (cf. Möhlig 1977, 1979, 1981a, 1981b). In Section 3.3, we showed that PWCB *k has /ʁ/ and /Ø/ as reflexes in Ngwi. Possibly, the words which have a zero reflex (distributionally less common) belong to a now extinct language from an earlier spread-over-spread event. In that language, PB *k and *g evolved to zero. When its speakers shifted to ancestral Ngwi, which had /ʁ/ as the regular reflex of PWCB *k, they may have retained a core of their ancestral vocabulary in which PB *k and *g had become zero. Even if such a scenario is hard – not to say impossible – to substantiate in the absence of historical language evidence, it ties in quite well with the fact that Ø is the predominant reflex of PWCB *k in the other languages of the wider WCB homeland and with what we know about the chronology of sound changes in that area. Although we cannot date the loss of intervocalic consonants, we know that this loss is older than other common diachronic sound changes such as final vowel loss. Likewise, elsewhere in WCB, small portions of words with minority reflexes of PWCB *k could also be relics of ancestral languages that became extinct as the outcome of Bantu-internal language shift.

5. Conclusions

In this article we presented the first case study of multiple unconditioned reflexes (MUR) of one and the same proto-sound (PWCB *k in C_2) in a well established Bantu branch (WCB). This study yields several conclusions that are significant not only for the diachrony of this major Bantu branch, but also for African historical linguistics more generally.

First of all, we proposed an evidence-based alternative solution to a long-standing problem within Bantu/Niger-Congo historical linguistics, i.e., the so-called "double (consonantal) reflexes". We sought an explanation for the fact that certain PB consonants have two or more reflexes without a clear phonological conditioning in several northwestern Bantu languages outside of the two scenarios that have been traditionally debated, i.e., phonemic merger vs. phonemic split. In line with similar studies of pervasive irregularity in sound change in other parts of the world, we argued for a third explanation, i.e., the acceptance of irregular sound correspondences as the expected outcome of the specific sociocultural circumstances in which Bantu speech communities developed after the large-scale expansion of the Bantu languages. Irregularity does not need to be "resolved" by assuming ancient phonemic contrasts or latent conditionings that were lost. Quite the opposite, it should be embraced as an indicator of intensive language contact and linguistic stratification in the linguistic prehistory of Bantu speech communities. Widespread multilingualism in Bantu-speaking Central Africa and the absence of a one-to-one overlap between linguistically and socially defined units make irregular sound change the expected outcome. Through the case study of MUR in WCB, we provided empirical evidence to place Bantu irregular sound change within the broader discussion of irregularity in Neogrammarian sound change across different language families (see most notably Durie & Ross 1996 and chapters therein).[22]

Second, within this deeply-rooted multilingual environment without well-established linguistic "identities", we argued that lexical diffusion or the irregular spread of sound change word by word instead of phoneme by phoneme might be one of the language contact scenarios that (partially) explains the MUR of PWCB *k. We showed that in contrast to the initial merger of PB *g and *k, the subsequent evolutions of PWCB *k rarely affected all targetable items in the lexicon of WCB daughter languages. In the absence of a clear-cut phonological or morphological conditioning, the gradual spread of these sound shifts through the lexicon is a plausible scenario to account for the two or more reflexes of PWCB *k in C_2 in most present-day WCB languages. Lenitions such as fricativization or complete loss involve the mutation of several phonetic features at once and are known to be more sensitive to lexical irregularity in the world's languages. Unfortunately, we miss data on precolonial sociolinguistic factors which may have inhibited the full spread of fricativization and/or deletion of PWCB *k in C_2 through the lexicon of the innovative WCB languages. Nonetheless, the geographic distribution of fricative reflexes of PWCB *k emerging from the case study in Section 3 does suggest that contact between closely related WCB varieties played a role in the diffusion of these sounds.

---

[22] An anonymous reviewer wonders whether with this statement we claim or imply that whenever we observe regular sound change in a given language family throughout the world, we should infer that the regularity is due to absence of multilingualism and/or language contact. The answer is no. For example, multilingualism is pervasive in Austronesian, a language family where the Comparative Method has been successfully applied (Kikusawa 2015). Nevertheless, while shared innovations have led to the identification of many undisputable subgroupings in Austronesian, languages spoken in the west of the Austronesian area posit considerable problems in this respect (Kikusawa 2015: 662–667). This shows that the regularity in sound change can serve for subclassification purposes in multilingual areas, but irregularity might coexist with regularity, at least in some areas of the world.

Third, given that velar and uvular fricatives are quite rare as reflexes of PB *k and *g in Bantu, their concentration in certain zones of the WCB domain (and elsewhere in the northwestern area) suggests that they may be an areal feature. This is especially so in the northwestern part of WCB, where B40–70 languages belonging to different WCB subclades have been entertaining long-standing contacts, but also closer to the WCB homeland, where several languages spoken in the vicinity of the Kasai River have them. We speculated that fricative reflexes of PB *k and *g in WCB (and elsewhere in NWB and CWB languages) might be the outcome of substrate influence from no longer spoken languages of extinct autochthonous hunter-gatherer groups, which already inhabited the region before the first Bantu speakers arrived south of the Equatorial rainforest some two to three millenia ago. This hypothesis finds only partial support in genetic studies on WCB speech communities.

Fourth, we entertained the possibility that MUR might be the result of the stratified non-tree-like history of the Bantu languages. Wilhelm Möhlig's stratification model finds support in recent interdisciplinary research suggesting that the dispersal of Bantu speakers from their homeland in the present-day Nigeria-Cameroon borderland throughout sub-Saharan Africa between approximately 5,000 and 1,500 years ago was a series of multiple migratory events through time. In this scenario, populations and languages in a given area died out or moved elsewhere and left relics which were then absorbed by a new spread event and so forth successively. Each of these incorporated relic languages might have contributed an unconditioned reflex to the incorporating language.

The take-home message we hope to leave with our readers is that the scenarios presented in Section 4 to account for the sound change irregularity observed in the West-Coastal Bantu region are not mutually exclusive. At present, we are unable to say if one of these might have greater explanatory power compared to others and we therefore consider them all as possible contributing factors to the phenomenon of MUR in WCB.

Funding

This research was funded by the ERC Consolidator Grant # 724275 granted to the second author and by the FWO (Fonds Wetenschappelijk Onderzoek Vlaanderen) postdoctoral fellowship n° 12ZV721N granted to the first author.

Acknowledgments

We are grateful to two anonymous reviewers for their critical and stimulating comments which greatly improved the argumentation in this article.

References

Bachmann, Armin R. 1989. Zum Fortis/Lenis-Kontrast in den nordwestlichen Bantusprachen. *Afrikanistische Arbeitspapiere* 19.23–31.

Bahuchet, Serge. 2012. Changing Language, Remaining Pygmy. *Human Biology* 84.11–43. https://doi.org/10.3378/027.084.0101

Bancel, Pierre. 1988. Double Reflexes in Bantu A.70 Languages. *Pholia* 3.7–16.

Bastin, Yvonne, André Coupez & Michael Mann. 1999. *Continuity and Divergence in the Bantu Languages: Perspectives from a Lexicostatistic Study*. Tervuren: Royal Museum for Central Africa.

Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo C. Schadeberg, eds. 2002. *Bantu Lexical Reconstructions* 3. Tervuren: Royal Museum for Central Africa, online database. https://www.africamuseum.be/nl/research/discover/human_sciences/culture_society/blr

Batini, Chiara, Joao Lopes, Doron M. Behar, Francesc Calafell, Lynn B. Jorde, Lolke van der Veen, Lluis Quintana-Murci, Gabriella Spedini, Giovanni Destro-Bisol & David Comas. 2010. Insights into the Demographic History of African Pygmies from Complete Mitochondrial Genomes. *Molecular Biology and Evolution* 28:2.1099–1110. https://doi.org/10.1093/molbev/msq294

Bissila, Sylvie. 1991. *Description phonologique du ilaale (dialecte teke du Congo)*. Brazzaville: Université Marien Ngouabi, Mémoire de Diplôme d'Etudes Supérieures.

Biton, Alexandre. 1969. *Dictionnaire ndumu-mbede-français et français-ndumu-mbede. Petite flore de la région de Franceville (Gabon). Grammaire ndumu-mbede*. Bar-le-Duc: Imprimerie St-Paul pour l'Archevêché de Libreville.

Blanchon, Jean Alain. 1991. Le pounou (B 43), le mpongwe (B 11a) et l'hypothèse fortis/lenis. *Pholia* 6.49–83.

Blanchon, Jean Alain. 2008. Dictionnaire punu. Manuscript. Lyon, Université Lumière – Lyon 2, DDL – Dynamique Du Langage UMR 5596.

Blust, Robert. 1996. The Neogrammarian Hypothesis and Pandemic Irregularity. *The Comparative Method Reviewed: Regularity and Irregularity in Language Change* ed. by Mark Durie & Malcolm Ross, 135–156. Oxford: Oxford University Press.

Bostoen, Koen. 2001. Osculance in Bantu Reconstructions: A Case Study of the Pair °-kádang-/°-káng- ('fry', 'roast') and Its Historical Implications. *Studies in African Linguistics* 30.121–146. https://doi.org/10.32473/sal.v30i2.107356

Bostoen, Koen. 2007. Bantu Plant Names as Indicators of Linguistic Stratigraphy in the Western Province of Zambia. In *Selected Proceedings of the 37th Annual Conference on African Linguistics* ed. by Doris L. Payne & Jaime Peña, 16–29. Somerville: Cascadilla Press.

Bostoen, Koen. 2008. Bantu Spirantization: Morphologization, Lexicalization and Historical Classification. *Diachronica* 25:3.299–356. https://doi.org/10.1075/dia.25.3.02bos

Bostoen, Koen. 2014. Wild Trees in the Subsistence Economy of Early Bantu Speech Communities: A Historical-Linguistic Approach. *Archaeology of African Plant Use* ed. by C.J. Stevens, S. Nixon, Mary Anne Murray & Dorian Q. Fuller, 129–140. Walnut Creek: Left Coast Press.

Bostoen, Koen. 2018. The Bantu Expansion. *Oxford Research Encyclopedia of African History* ed. by Thomas Spear. Oxford: Oxford University Press. Oxford Research Encyclopedias Online. https://global.oup.com/academic/product/the-oxford-encyclopedia-of-african-historiography-methods-and-sources-9780190698706?q=bostoen&lang=en&cc=us#. https://doi.org/10.1093/acrefore/9780190277734.013.191

Bostoen, Koen. 2019. Reconstructing Proto-Bantu. *The Bantu Languages* ed. by Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson, 308–334. 2nd ed. Oxford: Routledge. https://doi.org/10.4324/9781315755946-10

Bostoen, Koen & Yvonne Bastin. 2016. Bantu Lexical Reconstructions. *Oxford Handbooks Online*. Oxford: Oxford University Press. https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199935345.001.0001/oxfordhb-9780199935345-e-36

Bostoen, Koen, Bernard Clist, Charles Doumenge, Rebecca Grollemund, Jean-Marie Hombert, Joseph Koni Muluwa & Jean Maley. 2015. Middle to Late Holocene Paleoclimatic Change and the Early Bantu Expansion in the Rain Forests of West Central-Africa. *Current Anthropology* 56:3.354–384. https://doi.org/10.1086/681436

Bostoen, Koen & Gilles-Maurice de Schryver. 2018a. Langues et évolution linguistique dans le royaume et l'aire kongo. *Une archéologie des provinces septentrionales du royaume Kongo* ed. by Bernard Clist, Pierre de Maret & Koen Bostoen, 51–55. Oxford: Archaeopress.

Bostoen, Koen & Gilles-Maurice de Schryver. 2018b. Seventeenth-Century Kikongo Is Not the Ancestor of Present-Day Kikongo. *The Kongo Kingdom: The Origins, Dynamics and Cosmopolitan Culture of an African Polity* ed. by Koen Bostoen & Inge Brinkman, 60–102. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108564823.004

Bostoen, Koen & Heidi Goes. 2019. Was Proto-Kikongo a 5 or 7 Vowel Language? Bantu Spirantization and Vowel Merger in the Kikongo Language Cluster. *Linguistique et langues africaines* 5.25–68. https://doi.org/10.4000/lla.759

Bostoen, Koen & Joseph Koni Muluwa. 2011. Vowel Split in Hungan (Bantu H42, Kwilu, DRC): A Contact-Induced Language-Internal Change. *Journal of Historical Linguistics* 1:2.247–268. https://doi.org/10.1075/jhl.1.2.04bos

Bostoen, Koen & Joseph Koni Muluwa. 2014. Umlaut in the Bantu B70/80 languages of the Kwilu (DRC). *Transactions of the Philological Society* 112:2.209–230. https://doi.org/10.1111/1467-968X.12033

Bostoen, Koen & Léon Mundeke. 2011. The Causative/Applicative Syncretism in Mbuun (Bantu B87, DRC): Semantic Split or Phonemic Merger? *Journal of African Languages and Linguistics* 32:2.179–218. https://doi.org/10.1515/jall.2011.007

Botne, Robert. 1992. Double Reflexes in Eastern and Southern Bantu. *Afrikanistische Arbeitspapiere* 29.131–148.

Burssens, Nico. 1999. *Dictionnaire français-buma*. Bandundu: Centre d'Etudes Ethnologiques de Bandundu.

Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511612886

Bybee, Joan. 2002. Word Frequency and Context of Use in the Lexical Diffusion of Phonetically Conditioned Sound Change. *Language Variation and Change* 14.261–290. https://doi.org/10.1017/S0954394502143018

Campbell, Lyle. 1996. On Sound Change and Challenges to Regularity. *The Comparative Method Reviewed: Regularity and Irregularity in Language Change* ed. by Mark Durie & Malcolm Ross, 72–89. Oxford: Oxford University Press.

Campbell, Lyle. 1998. *Historical Linguistics : An Introduction*. Edinburgh: Edinburgh University Press.

Coupez, André, Yvonne Bastin & Evariste Mumba (eds.). 1998. *Reconstructions lexicales bantoues 2 / Bantu Lexical Reconstructions 2*. Electronic database. Tervuren: Musée royal de l'Afrique centrale.

Crane, Thera Marie, Larry M. Hyman & Simon Nsielanga Tukumu. 2011. *A Grammar of Nzadi (B.865) A Bantu Language of the Democratic Republic of Congo with an Appendix of Proto-Bantu – Nzadi Sound Correspondences by Clara Cohen*. Berkeley: University of California Publications.

Cser, András. 2015. Basic Types of Phonological Change. *The Oxford Handbook of Historical Phonology* ed. by Patrick Honeybone & Joseph Salmons, 193–204. Oxford: Oxford University Press.

Currie, Thomas E., Andrew Meade, Myrtille Guillon & Ruth Mace. 2013. Cultural Phylogeography of the Bantu Languages of Sub-Saharan Africa. *Proceedings of the Royal Society B (Biological Sciences)* 280.20130695 20130695, https://doi.org/10.1098/rspb.2013.0695

Daeleman, Jan. 1977. A Comparison of Some Zone B Languages in Bantu. *Africana Linguistica* 7.93–144. https://doi.org/10.3406/aflin.1977.899

Daeleman, Jan. 1983. Les réflexes du proto-bantu en ntandu (dialecte koongo). *Mélanges de culture et de linguistique africaines publiés à la mémoire de Leo Stappers* ed. by Clémentine Faïk-Nzuji Madiya & Erika Sulzmann, 331–397. Berlin: Dietrich Reimer Verlag.

Daeleman, Jan. n.d. Kikoóngo wordlists classified according to their respective tones. Manuscript.

de Filippo, Cesare, Chiara Barbieri, Mark Whitten, Sununguko Wata Mpoloka, Ellen D. Gunnarsdóttir, Koen Bostoen, Terry Nyambe, Klaus Beyer, Henning Schreiber, Peter de Knijff, Donata Luiselli, Mark Stoneking & Brigitte Pakendorf. 2011. Y-Chromosomal Variation in Sub-Saharan Africa: Insights into the History of Niger-Congo Groups. *Molecular Biology and Evolution* 28:3.1255–1269. https://doi.org/10.1093/molbev/msq312

De Grauwe, Jan. 2009. *Lexique yoómbe-français, avec index français-yoómbe (bantu H16c)*. Tervuren: Royal Museum for Central Africa.

de Schryver, Gilles-Maurice, Rebecca Grollemund, Simon Branford & Koen Bostoen. 2015. Introducing a State-of-the-Art Phylogenetic Classification of the Kikongo Language Cluster. *Africana Linguistica* 21.87–162. https://doi.org/10.3406/aflin.2015.1044

De Witte, P. 1955. *Taalstudie bij de Basakata*. Tervuren: Koninklijk Museum van Belgisch-Congo.

De Wulf, Chris, Johannes Goossens & Johan Taeldeman (eds). 2005. *Fonologische atlas van de Nederlandse dialecten (FAND): deel IV: de consonanten*. Gent: Koninklijke Academie voor Nederlandse Taal- en Letterkunde.

Destro-Bisol, Giovanni, Francesco Donati, Valentina Coia, Ilaria Boschi, Fabio Verginelli, Alessandra Caglià, Sergio Tofanelli, Gabriella Spedini & Cristian Capelli. 2004. Variation of Female and Male Lineages in Sub-Saharan Populations: The Importance of Sociocultural Factors. *Molecular Biology and Evolution* 21.1673–1682. https://doi.org/10.1093/molbev/msh186

Dibata Mimpiya, Iliku. 1977. *Le substantif mbuun: étude d'une forme nominale*. Lubumbashi: Université de Lubumbashi, travail de fin d'études de graduat.

Durie, Mark & Malcolm Ross (eds.). 1996. *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*. Oxford: Oxford University Press.

Ebalantshim Masuwan. 1980. *Esquisse grammaticale de la langue idiŋ à Káantsa*. Lubumbashi: Université Nationale du Zaïre, M.A.

Ellington, John Ernest. 1977. Aspects of the Tiene Language. Madison: University of Wisconsin, PhD thesis.

Fabian, Johannes. 1986. *Language and Colonial Power. The Appropriation of Swahili in the Former Belgian Congo 1880–1938*. Berkeley: University of California Press.

Fortes-Lima, Cesar, Carina Schlebusch, Leon Mundeke, Sara Pacchiarotti & Koen Bostoen. 2021. New Genetic Light on Old Linguistic Questions from the West-Coastal Bantu Homeland. Paper presented at "Extracting the Past from the Present", Interdisciplinary Conference on African Precolonial History, March 1–5, 2021, Brussels, Université libre de Bruxelles.

Garrett, Andrew. 2015. Sound Change. *The Routledge Handbook of Historical Linguistics* ed. by Claire Bowern & Bethwyn Evans, 227–248. London: Routledge.

Gerhardt, Ludwig. 1986. The Fortis/Lenis Contrast in Pre-Bantu? Some North-Western Bantu Evidence. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 39:5.540–547. https://doi.org/10.1524/stuf.1986.39.14.540

Goes, Heidi & Koen Bostoen. 2019. Progressive Vowel Height Harmony in Proto-Kikongo and Proto-Bantu. *Journal of African Languages and Linguistics*. 40:3.1–51. https://doi.org/10.1515/jall-2019-0003

Grace, George W. 1996. Regularity of Change in What?. *The Comparative Method Reviewed: Regularity and Irregularity in Language Change* ed. by Mark Durie & Malcolm Ross, 157–179. Oxford: Oxford University Press.

Greenberg, Joseph Harold. 1948. The Tonal System of Proto-Bantu. *Word* 4:3.196–208. https://doi.org/10.1080/00437956.1948.11659343

Grégoire, Claire. 2003. The Bantu Languages of the Forest. *The Bantu Languages* ed. by Derek Nurse & Gérard Philippson, 349–370. London New York: Routledge.

Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu Expansion Shows That Habitat Alters the Route and Pace of Human Dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112:43.13296–13301. https://doi.org/10.1073/pnas.1503793112

Güldemann, Tom. 2011. Proto-Bantu and Proto-Niger-Congo: Macro-Areal Typology and Linguistic Reconstruction. *Geographical Typology and Linguistic Areas with Special Reference to Africa* ed. by Osamu Hieda, Christa König & Hiroshi Nakagawa, 109–141. Amsterdam: John Benjamins. https://doi.org/10.1075/tufs.2.09gul

Guthrie, Malcolm. 1962. A Two-Stage Method of Comparative Bantu Study. *African Language Studies* 3.1–24.

Guthrie, Malcolm. 1967. *Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of the Bantu languages. Volume 1: The Comparative Linguistics of the Bantu Languages*. London: Gregg International.

Guthrie, Malcolm. 1971. *Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of the Bantu languages. Volume 2: Bantu Prehistory, Inventory and Indexes*. London: Gregg International.

Harrison, Shelly P. 2003. On the Limits of the Comparative Method. *The Handbook of Historical Linguistics* ed. by Brian D. Joseph & Richard D. Janda, 213–243. Malden: Blackwell. https://doi.org/10.1002/9780470756393.ch2

Hedinger, Robert. 1987. *The Manenguba Languages (Bantu A.15, Mbo cluster) of Cameroon*. London: School of Oriental and African Studies.

Hombert, Jean-Marie & Médard Mouélé. 1988. Éléments de phonologie diachronique du wanzi (langue bantu du Gabon, groupe B50). *Pholia* 3.183–205.

Huffman, Tom N. 2007. *Handbook to the Iron Age: The Archaeology of Pre-Colonial Farming Societies in Southern Africa*. Pietermaritzburg: University of KwaZulu-Natal Press.

Hyman, Larry M. 2004. How to Become a Kwa Verb. *Journal of West African Languages* 30:2.69–88.

Hyman, Larry M. 2010. Affixation by Place of Articulation: The Case of Tiene. *Rara & Rarissima: Documenting the Fringes of Linguistic Diversity* ed. by Jan Wohlgemuth & Michael Cysouw, 145–184. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110228557.145

Idiata-Mayombo, Daniel-Franck. 1993. Eléments de phonologie diachronique du isangu (B.42). *Pholia* 8.67–108.

Impubi Mukwa, Mfum. 1987. *Le syntagme nominal en yans B85d. Parler de Mikingi Kiakia – Collectivité de Nkara/Niadi*. Kikwit: Institut Supérieur Pédagogique de Kikwit, travail de fin d'études.

Janssens, Baudouin. 1986. Eléments de phonologie et de morphologie historique du basaa (bantou A43a). *Africana Linguistica* 10.149–211. https://doi.org/10.3406/aflin.1986.928

Janssens, Baudouin. 1991. Doubles réflexes apparents en ewondo ou les chassés croisés de la dérivation. *Pholia* 6.155–179.

Janssens, Baudouin. 1993. *Doubles réflexes consonantiques : quatre études sur le bantu de zone A (bubi, nen, bafia, ewondo)*. Bruxelles: Université libre de Bruxelles, thèse de doctorat.

Kanyamibwa, Melchior. 1982. *Confrontation historique du lexique du protobantu avec le cokwe (zone K), le sanga (zone L) et le rwanda (zone J)*. Ruhengeri: Université nationale du Rwanda, mémoire de licence.

Khang Levy, Nyi-m'shum. 1979. *Eléments de grammaire morphologique de la langue lwel*. Lubumbashi: Université Nationale du Zaïre, mémoire de licence.

Kibwenge India'Ane, Passy. 1985. *Langue mpur B85e. Phonologie et morphologie synthématique : parler de Mikingi, collectivité Dwe*. Kikwit: ISP de Kikwit, mémoire de licence.

Kikusawa, Ritsuko. 2015. The Austronesian Language Family. *The Routledge Handbook of Historical Linguistics* ed. by Claire Bowern & Bethwyn Evans, 657–674. London: Routledge.

Kinkade, M. Dale. 1973. The Alveopalatal Shift in Cowlitz Salish. *International Journal of American Linguistics* 39:4.224–231. https://doi.org/10.1086/465270

Koni Muluwa, Joseph. 2010. *Plantes, animaux et champignons en langues bantu. Etude comparée de phytonymes, zoonymes et myconymes en nsong, ngong, mpiin, mbuun et hungan (Bandundu, RD Congo)*. Bruxelles: Université libre de Bruxelles, thèse de doctorat.

Koni Muluwa, Joseph. 2014. *Noms et usages de plantes, animaux et champignons chez les Mbuun, Mpiin, Ngong, Nsong et Hungan en RD Congo*. Tervuren: Musée royal de l'Afrique centrale.

Koni Muluwa, Joseph. 2015a. Lexicon English- French-Ngong (Kwenge dialect). Manuscript. Ghent, Ghent University, DoBes Kwilu Bantu project. http://www.kwilubantu.ugent.be/

Koni Muluwa, Joseph. 2015b. Petit lexique français-anglais-nsambaan (parlers de Kikongo et Kwilumpia) – nsambaan-français-anglais. Manuscript. Gand, Université de Gand, projet DoBes Kwilu Bantu. http://www.kwilubantu.ugent.be/

Koni Muluwa, Joseph & Koen Bostoen. 2011. Umlaut in the Bantu B70/80 Languages of the Kwilu (DRC): Where Did the Final Vowel Go? Paper presented at 41th Colloquium on African Languages and Linguistics, Leiden.

Koni Muluwa, Joseph & Koen Bostoen. 2012. La diphtongaison dans les langues bantu B70–80 (Bandundu, RDC): typologie et classification historique. *Africana Linguistica* 18.355–386. https://doi.org/10.3406/aflin.2012.1015

Koni Muluwa, Joseph & Koen Bostoen. 2015. *Lexique comparé des langues bantu du Kwilu (République démocratique du Congo)*. Cologne: Rüdiger Köppe.

Labov, William. 1981. Resolving the Neogrammarian Controversy. *Language* 57:2.267–308. https://doi.org/10.2307/413692

Labov, William. 1999. *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell.

Labroussi, Catherine. 1999. Vowel Systems and Spirantization in Southwest Tanzania. *Bantu Historical Linguistics: Theoretical and empirical perspectives* ed. by Jean-Marie Hombert & Larry M. Hyman, 335–377. Stanford, CA: CSLI.

Laman, Karl Edward. 1936. *Dictionnaire kikongo-français, avec une étude phonétique décrivant les dialectes les plus importants de la langue dite kikongo*. Bruxelles: Librairie Falk.

Laman, Karl Edward. 1912. *Grammar of the Kongo Language (Kikongo)*. New York: Christian Alliance Publishers.

Laman, Karl Edward & Carl Meinhof. 1928–1929. An Essay in Kongo Phonology. *Zeitschrift für Eingeborenen-Sprachen* 19.12–40.

Lane, Bryan David. 1989. Toward Understanding Multilingualism Among the Mbete of Northern Congo. Arlington: University of Texas, MA thesis.

Li, Paul Jen-Kuei. 1982. Linguistic Variation of Different Age Groups in the Atayalic Dialects. *The Tsing Hua Journal of Chinese Studies* 14.167–191.

Loftus, Emma, Peter J. Mitchell & Christopher Bronk Ramsey. 2019. An Archaeological Radiocarbon Database for Southern Africa. *Antiquity* 93:370.870–885. https://doi.org/10.15184/aqy.2019.75

Loubelo, Fidèle. 1987. *Description phonologique du itsaangi, parler de Madouma-Mossendjo*. Libreville: Université Marien Ngouabi, mémoire de D.E.S.

Maho, Jouni Filip. 2009. NUGL Online: The Online Version of the New Updated Guthrie List, a Referential Classification of the Bantu Languages (4 June 2009). http://goto.glocalnet.net/mahopapers/nuglonline.pdf

Makouta-Mboukou, Jean-Pierre. 1969. *Vocabulaire fumu-français*. Brazzaville: Centre d'Enseignement Superieur de Brazzaville (polycopié).

Makouta-Mboukou, Jean-Pierre. 1976. *Étude descriptive du fumu, dialecte teke de Ngamaba, Brazzaville*. Paris: Paris III, Sorbonne Nouvelle, thèse de doctorat d'état.

Marchal-Nasse, Colette. 1989. *De la phonologie à la morphologie du nzebi, langue bantoue (B52) du Gabon*. Bruxelles: Université libre de Bruxelles, thèse de doctorat.

Masele, Balla F.Y.P. & Derek Nurse. 2003. Stratigraphy and Prehistory: Bantu Zone F. *Language Contacts in Prehistory: Studies in Stratigraphy* ed. by Henning Andersen, 115–134. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.239.09mas

Mavoungou, Paul Achille & Bernard Plumel. 2010. *Dictionnaire yilumbu-français*. Libreville: Éditions Raponda Walker.

Mayanga, Tayeye. 1985. *Grammaire yansi (Rép. du Zaïre)*. Bandundu Centre d'Études Ethnologiques de Bandundu.

Medjo Mvé, Pither. 1989. *Exploration morpho-syntaxique du dialecte le nyani ou contribution à l'étude du verbe ndumu*. Libreville: Université Omar Bongo, mémoire de maîtrise.

Meeussen, Achiel Emiel. 1967. Bantu Grammatical Reconstructions. *Africana Linguistica* 3.79–121. https://doi.org/10.3406/aflin.1967.873

Meeussen, Achiel Emiel. 1969. Bantu Lexical Reconstructions. Manuscript. Tervuren.

Meinhof, Carl. 1899. *Grundriß einer Lautlehre der Bantusprachen nebst Anleitung zur Aufnahme von Bantusprachen – Anhang: Verzeichnis von Bantuwortstämmen*. Leipzig: F.A. Brockhaus.

Meinhof, Carl. 1906. *Grundzüge einer vergleichenden Grammatik der Bantusprachen*. Berlin: Reimer.

Mickala-Manfoumbi, Roger. 1988. *Eléments de description du duma, langue bantu du Gabon (B51)*. Bruxelles: Université libre de Bruxelles, mémoire de licence spéciale.

Miehe, Gudrun. 1989. Zur Genese von fortis und lenis im Volta-Congo. *Frankfurter Afrikanistische Blätter* 1.30–43.

Milroy, James. 1992. *Linguistic Variation and Change: On the Historical Sociolinguistics of English*. Oxford: Blackwell.

Mitchell, Peter. 2013. Early Farming Communities of Southern and South-Central Africa. *Oxford Handbook of African Archaeology* ed. by Peter Mitchell & Paul Lane, 657–670. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199569885.001.0001

Mitchell, Peter & Gavin Whitelaw. 2005. The Archaeology of Southernmost Africa from c. 2000 BP to the Early 1800s: A Review of Recent Research. *The Journal of African History* 46:2.209–241. https://doi.org/10.1017/S0021853705000770

Möhlig, Wilhelm J.G. 1977. Zur frühen Siedlungsgeschichte der Savannen-Bantu aus lauthistorischer Sicht. *Zur Sprachgeschichte und Ethnohistorie in Afrika* ed. by Wilhelm Möhlig, Franz Rottland & Bernd Heine, 166–193. Berlin: D. Reimer.

Möhlig, Wilhelm J.G. 1979. The Bantu Nucleus: Its Conditional Nature and Its Prehistorical Significance. *Sprache und Geschichte in Afrika* 1.109–141.

Möhlig, Wilhelm J.G. 1981a. Stratification in the History of the Bantu Languages. *Sprache und Geschichte in Afrika* 251–317.

Möhlig, Wilhelm Johann Georg. 1981b. Lineare und hybride Lautverschiebungen im Bantu. *Berliner afrikanistische Vorträge (XXI. deutscher Orientalistentage, Berlin, 24.-29. 3. 1980* ed. by Herrmann Jungraithmayr & Gudrun Miehe, 81–102. Berlin: Dietrich Reimer.

Mokrani, Soraya. 2016. *Étude comparée des parlers du groupe bantu kota-kele (B20) du Gabon: À la recherche de nouveaux critères classificatoires*. Lyon: Université Lumière-Lyon 2, thèse de doctorat.

Motingea Mangulu, André. 2004. Divergences en dialectes tiene (bantou B.81): origine disparate ou perte séparée d'identité linguistique?. *Annales Aequatoria* 25.151–202.

Motingea Mangulu, André. 2021. Vers une écologie linguistique des anciens chasseurs-cueilleurs du bassin central congolais: Traces d'une proto-population ancestrale homogène. *Anthropos* 116:1.137–143. https://doi.org/10.5771/0257-9774-2021-1-137

Mouandza, Jean-Daniel. 2001. *Éléments de description du iyaa (parler bantu du Congo-Brazzaville)*. Nice: Université de Nice Sophia Antipolis, thèse de doctorat.

Mouélé, Médard. 1997. *Étude synchronique et diachronique des parlers dúmá (groupe bantu B.50)*. Lyon: Université Lumière Lyon 2, thèse de doctorat.

Mouguiama-Daouda, Patrick & Lolke van der Veen. 2005. B10-B30: conglomérat phylogénétique ou produit d'une hybridation?. *Studies in African Comparative Linguistics, with Special Focus on Bantu and Mande* ed. by Koen Bostoen & Jacky Maniacky, 91–122. Tervuren: Royal Museum for Central Africa.

Mufanga-Dzmar, Kitoko. 1977. *Contribution à la phonologie ding*. Lubumbashi: Université Nationale du Zaïre, travail de fin de cycle de graduat.

Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511612862

Mula, Emy. 1977. *Esquisse de grammaire kiding munken : phonologie et morphologie (parler de Bantsiona)*. Lubumbashi: Université Nationale du Zaïre, mémoire de licence.

Mundeke, Léon. 1979. *Esquisse grammaticale de la langue mbúún (parler de Éliob)*. Lubumbashi: Université Nationale du Zaïre, mémoire de licence.

Mundeke, Léon. 2011. *Etude morphosyntaxique de la langue mbuun (B87) (parler d'Elyob)*. Lubumbashi: Université de Lubumbashi.

Munkyen Okab, Saan Lakin. 1990. *Etude contrastive phonetico-phonologique entre le français et le ding B86*. Paris: Université de la Sorbonne Nouvelle Paris III, thèse de doctorat.

Mwan Mesongolo, Mekani. 1984. *Phonologie et morphologie de kedeŋ-Lesye (B86). Parler de Mateko-Mfwaay*. Lubumbashi: Université Nationale de Zaïre, mémoire de licence.

Narciso Cobe, Francisco. 2010. *Novo Dicionário Português-Kikongo*. Luanda: Mayamba.

Ndongo Ibara, Yvon Pierre. 2000. The Phonology of Tsambitsɔ: Embɔsi dialect. Brazzaville Université Marien Ngouabi, MA thesis.

Ndouli, Guy Blaise. 2001. *Description phonologique du mbere: parler de Tsama*. Brazzaville: Université Marien Ngouabi, mémoire de maîtrise.

Nguimbi-Mabiala, Jean-Noël. 1999. *Phonologie comparative et historique du koongo*. Lyon: Université Lumiere Lyon 2, thèse de doctorat.

Nguma, Wendo. 1986. *Dictionnaire français-yansi (Rép. du Zaïre)*. Bandundu: CEEBA.

Nkouanda, Blaise Pépin. 1997. *Description phonologique du kiga:ngala, parler bantou du Congo (Groupe H)*. Brazzaville: Université Marien Ngouabi, mémoire de maîtrise.

Nsuka-Nkutsi, François. 1980. Quelques réflexes du protobantou en punu. *Eléments de description du punu* ed. by François Nsuka-Nkutsi, 128–178. Lyon: Centre de Recherches Linguistiques et Sémiologiques (CRLS), Université Lumière-Lyon 2.

Nurse, Derek & Balla F. Y. P. Masele. 2003. Stratigraphy and Prehistory: Bantu Zone F. *Language Contacts in Prehistory: Studies in Stratigraphy* ed. by Henning Andersen, 115–134. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.239.09mas

Nurse, Derek & Gérard Philippson. 2003. Towards a Historical Classification of the Bantu Languages. *The Bantu Languages* ed. by Derek Nurse & Gérard Philippson, 164–181. London: Routledge.

Okoudowa, Bruno. 2016. *Morfologia verbal do Lèmbáámá: Língua banta falada no Gabaõ (África Central)*. Saarbrücken: Novas Edições Acadêmicas.

Pacchiarotti, Sara & Koen Bostoen. 2020. The Proto-West-Coastal Bantu Velar Merger. *Africana Linguistica* 26.139–195.

Pacchiarotti, Sara & Koen Bostoen. 2021. Final Vowel Loss in Lower Kasai Bantu (DRC) as a Contact-Induced Change. *Journal of Language Contact* 14.437–474. https://doi.org/10.1163/19552629-14020007

Pacchiarotti, Sara, Natalia Chousou-Polydouri & Koen Bostoen. 2019. Untangling the West-Coastal Bantu Mess: Identification, Geography and Phylogeny of the Bantu B50–80 Languages. *Africana Linguistica* 25.155–229.

Pacchiarotti, Sara, Lorenzo Maselli, and Koen Bostoen. 2021. The Emergence of Interior Vowels and Heterosyllabic Vowel Sequences in Ngwi (Bantu B861, DRC). *Papers in Historical Phonology* 6. 107–146.

Patin, Etienne, Marie Lopez, Rebecca Grollemund, Paul Verdu, Christine Harmant, Hélène Quach, Guillaume Laval, George H. Perry, Luis B. Barreiro, Alain Froment, Evelyne Heyer, Achille Massougbodji, Cesar Fortes-Lima, Florence Migot-Nabias, Gil Bellis, Jean-Michel Dugoujon, Joana B. Pereira, Verónica Fernandes, Luisa Pereira, Lolke Van der Veen, Patrick Mouguiama-Daouda, Carlos D. Bustamante, Jean-Marie Hombert & Lluís Quintana-Murci. 2017. Dispersals and Genetic Adaptation of Bantu-Speaking Populations in Africa and North America. *Science* 356:6337.543–546. https://doi.org/10.1126/science.aal1988

Patin, Etienne, Katherine J. Siddle, Guillaume Laval, Hélène Quach, Christine Harmant, Noémie Becker, Alain Froment, Béatrice Régnault, Laure Lemée, Simon Gravel, Jean-Marie Hombert, Lolke Van der Veen, Nathaniel J. Dominy, George H. Perry, Luis B. Barreiro, Paul Verdu, Evelyne Heyer & Lluís Quintana-Murci. 2014. The Impact of Agricultural Emergence on the Genetic History of African Rainforest Hunter-Gatherers and Agriculturalists. *Nature Communications* 5.3163. https://doi.org/10.1038/ncomms4163

Paulian, Christiane. 1975. *Le kukuya, langue teke du Congo: phonologie- classes nominales.* Paris: Société des Etudes Linguistiques et Anthropologiques.

Philippson, Gérard. 1999. *HH and *HL tone patterns in Bemba and the Bemba tone systems. *Bantu Historical Linguistics: Theoretical and Empirical Perspectives* ed. by Jean-Marie Hombert & Larry M. Hyman, 395–411. Stanford: CSLI Publications.

Philippson, Gérard. 2018. Double Reflexes' Revisited: Implications for the Proto-Bantu Consonant System. Paper presented at International Conference on "Reconstructing Proto-Bantu Grammar", Ghent University. https://research.flw.ugent.be/en/projects/reconstructing-proto-bantu-grammar

Philippson, Gérard & Rebecca Grollemund. 2019. Classifying Bantu Languages. *The Bantu Languages* ed. by Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson, 335–354, 2nd ed Oxford: Routledge. https://doi.org/10.4324/9781315755946-11

Phillips, Betty S. 1984. Word Frequency and the Actuation of Sound Change. *Language* 60:2.320–342. https://doi.org/10.2307/413643

Phillips, Betty S. 2006. *Word Frequency and Lexical Diffusion*. New York: Palgrave Macmillan. https://doi.org/10.1057/9780230286610

Pokoso, Kayabu-Koko. 1986. *Essai d'anthroponymie ngongo (parler Lukula sud)*. Kikwut: Institut Supérieur Pédagogique de Kikwit, mémoire de licence.

Pour, Ansari Naser, Christopher A. Plaster & Neil Bradman. 2013. Evidence from Y-Chromosome Analysis for a Late Exclusively Eastern Expansion of the Bantu-Speaking People. *European Journal of Human Genetics* 21:4.423–429. https://doi.org/10.1038/ejhg.2012.176

Quintana-Murci, Lluís, Hélène Quach, Christine Harmant, Fracesca Luca, Blandine Massonnet, Etienne Patin, Lucas Sica, Patrick Mouguiama-Daouda, David Comas, Shay Tzur, Oleg Balanovsky, Kenneth K. Kidd, Judith R. Kidd, Lolke van der Veen, Jean-Marie Hombert, Antoine Gessain, Paul Verdu, Alain Froment, Serge Bahuchet, Evelyne Heyer, Jean Dausset, Antonio Salas & Doron M. Behar. 2008. Maternal Traces of Deep Common Ancestry and Asymmetric Gene Flow between Pygmy Hunter-Gatherers and Bantu-Speaking Farmers. *Proc Natl Acad Sci U S A* 105:5.1596–1601. https://doi.org/10.1073/pnas.0711467105

Raharimanantsoa, Ruth. 2012a. *Aspects of Phonology in Eboo-Nzikou (Bantu B74)*. Gothenburg: University of Gothenburg, preparatory research course paper.

Raharimanantsoa, Ruth. 2012b. Processes of Vowel Change Due to Second Consonant Loss in Boma-Nzikou (Bantu B74). Paper presented at the 42nd Colloquium on African Languages and Linguistics, University of Leiden.

Raharimanantsoa, Ruth. 2016. Réduction syllabique, rallongement compensatoire et syllabes trimoraïques en engungwel (bantou B72a). *Journal of West African Languages* 43:2.61–98.

Raharimanantsoa, Ruth. 2017. The Prosody of Tense Marking in Teke-Eboo: A Bantu B70 language of Congo-Brazzaville. Gothenburg: Gothenburg University, MA thesis.

Raharimanantsoa, Ruth & Wilfrid Edouard Ntsiba Ngolo. 2015. *Petit guide d'orthographe et de grammaire du Téké du sud-ouest parlé dans le Pool, la Bouenza et la Lékoumou, en République du Congo*. Brazzaville: SIL-Congo.

Ricquier, Birgit & Koen Bostoen. 2008. Resolving Phonological Variability in Bantu Lexical Reconstructions: The Case of 'To Bake in Ashes'. *Africana Linguistica* 14.109–150. https://doi.org/10.3406/aflin.2008.1052

Ross, Malcolm D. 1996. Contact-Induced Change and the Comparative Method: Cases from Papua New Guinea. *The Comparative Method Reviewed: Regularity and Irregularity in Language Change* ed. by Mark Durie & Malcolm Ross, 180–217. Oxford: Oxford University Press.

Rossel, Gerda. 1998. Taxonomic-Linguistic Study of Plantain in Africa. Wageningen: Landbouwuniversiteit Wageningen, PhD thesis.

Rottland, Franz. 1977. Reflexes of Proto-Bantu Phonemes in Yanzi (B85). *Africana Linguistica* 7.375–396. https://doi.org/10.3406/aflin.1977.906

Rurangwa, Innocent Moïse. 1982. *Éléments de description du ngungwel, langue bantoue du Congo*. Bruxelles: Université Libre de Bruxelles, mémoire de licence spéciale.

Ruttenberg, Piet. 2000. *Lexique yaka-francais, français-yaka*. München: Lincom Europa.

Schadeberg, Thilo C. 1995. Spirantization and the 7-to-5 Vowel Merger in Bantu. *Belgian Journal of Linguistics* 9.71–84.

Schadeberg, Thilo C. 2002. Progress in Bantu Lexical Reconstruction. *Journal of African Languages and Linguistics* 23:2.183–195. https://doi.org/10.1515/jall.2002.011

Schadeberg, Thilo C. 2003. Historical Linguistics. *The Bantu Languages* ed. by Derek Nurse & Gérard Philippson, 143–163. London: Routledge.

Schoeman, Alex. 2013. Southern African Late Farming Communities. *Oxford Handbook of African Archaeology* ed. by Peter Mitchell & Paul Lane, 928–941. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199569885.013.0064

Seidensticker, Dirk, Wannes Hubau, Dirk Verschuren, Cesar Fortes-Lima, Pierre de Maret, Carina M. Schlebusch & Koen Bostoen. 2021. Population Collapse in Congo Rainforest from 400 CE Urges Reassessment of the Bantu Expansion. *Science Advances* 7:7.eabd8352. https://doi.org/10.1126/sciadv.abd8352

Stappers, Leo. 1986. *Boma: eine Sprachskizze*. Hamburg: Helmut Buske.

Stewart, John Massie. 1973. The Lenis Stops of the Potou Lagoon Languages and their Significance for Pre-Bantu Reconstruction. *Papers in Ghanaian Linguistics* ed. by M. E. Kropp Dakubu, 1–49. Legon: Institute of African Studies, University of Ghana.

Stewart, John Massie. 1989. Fortis/Lenis and Vowel Length in Proto-Bantu. *Journal of African Languages and Linguistics* 11:1.45–88.

Stewart, John Massie. 1993. The Second Tano Consonant Shift and its Likeness to Grimm's Law. *Journal of West African Languages* 23:1.3–39.

Swartenbroeckx, Pierre. 1973. *Dictionnaire kikongo et kituba-français*. Bandundu: Ceeba.

Swartenbroeckx, Pierre. 1948. Dictionnaire kiyansi ou kiyei, langage des bayansi ou bayey de territoire de Banningville (district du Lac Léopold II) au Congo Belge. Manuscript. Bruxelles.

Teil-Dautrey, Gisèle. 2008. Et si le proto-bantu était aussi une langue … avec ses contraintes et ses déséquilibres. *Diachronica* 25:1.54–110. https://doi.org/10.1075/dia.25.1.04tei

Thomason, Sarah G. 2007. Language Contact and Deliberate Change. *Journal of Language Contact Thema* 1.41–62. https://doi.org/10.1163/000000007792548387

Van Coetsem, Frans. 1988. *Loan Phonology and the Two Transfer Types in Language Contact.* Dordrecht: Foris. https://doi.org/10.1515/9783110884869

Van de Velde, Mark. 1999. The Two Language Maps of the Belgian Congo. *Annales Aequatoria* 20.475–489.

Van Den Eynde, Karel. 1968. *Éléments de grammaire yaka : phonologie et morphologie flexionnelle*. Kinshasa: Université Lovanium, Publications Universitaires.

van der Veen, Lolke. 1991. *Étude comparée des parlers du groupe okani (B.30 Gabon)*. Lyon: Université Lumière Lyon 2, thèse de doctorat nouveau régime.

van der Veen, Lolke. 2003. The B30 Language Group. *The Bantu Languages* ed. by Derek Nurse & Gérard Philippson, 371–391. London: Routledge.

Van Gheel, Joris. 1652. *Vocabularium Latinum, Hispanicum, e Congense. Ad Usum Missionariorû transmittendorû ad Regni Congo Missiones*. Rome: National Central Library, Fundo Minori 1896, MS Varia 274.

Van Leynseele, Hélène & John Massie Stewart. 1980. Harmonie consonantique en pré-nen. *L'expansion bantoue. Actes du Colloque International du CNRS, Viviers (France), 4–16 avril 1977, vol. 2* ed. by Luc Bouqiaux, 387–400. Paris: Société des Etudes Linguistiques et Anthropologiques.

Vansina, Jan. 1995. New Linguistic Evidence and the Bantu Expansion. *Journal of African History* 36:2.173–195. https://doi.org/10.1017/S0021853700034101

Verdu, Paul, Noemi S. A. Becker, Alain Froment, Myriam Georges, Viola Grugni, Lluis Quintana-Murci, Jean-Marie Hombert, Lolke J. Van der Veen, Sylvie Le Bomin, Serge Bahuchet, Evelyne Heyer & Frédéric Austerlitz. 2013. Sociocultural Behavior, Sex-Biased Admixture, and Effective Population Sizes in Central African Pygmies and Non-Pygmies. *Molecular Biology and Evolution* 30:4.918–937. https://doi.org/10.1093/molbev/mss328

Wang, William. 1969. Competing Sound Changes as a Cause of Residue. *Language* 45.9–25. https://doi.org/10.2307/411748

Weiss, Michael. 2015. The Comparative Method. *The Routledge Handbook of Historical Linguistics* ed. by Claire Bowern & Bethwyn Evans, 127–146. London: Routledge.

Winford, Donald. 2003. *An Introduction to Contact Linguistics*. Malden: Blackwell.

Wood, Elizabeth T., Daryn A. Stover, Christopher Ehret, Giovanni Destro-Bisol, Gabriella Spedini, Howard McLeod, Leslie Louie, Mike Bamshad, Beverly I. Strassmann, Himla Soodyall & Michael F. Hammer. 2005. Contrasting Patterns of Y Chromosome and mtDNA Variation in Africa: Evidence for Sex-Biased Demographic Processes. *European Journal of Human Genetics* 13.867–876. https://doi.org/10.1038/sj.ejhg.5201408

Appendix A. Cognate sets involving PB *k and *g not included in Pacchiarotti & Bostoen (2020: 175–189)

The following cognate sets are organized according to the conventions set out in Pacchiarotti and Bostoen (2020: 6–7). They are ordered according to the BLR proto form number (lower to higher) of which we believe they are a reflex. Each cognate set starts with the BLR proto form index number, form and meaning, followed by the synchronic reflexes we could identify. Each reflex is preceded by a bolded alphanumeric code denoting a specific variety and its source as referenced in Appendix 2. Absence of a reflex in one of the languages listed in Appendix 2 could mean that the language has a reflex of another protoform for a given concept or that we lack data for that variety. We specify the meaning of reflexes only if they differ with respect to the meaning(s) assigned to the corresponding protoform in the BLR 2/3 database. We mark high tone as [á], low tone as [à], rising tone as [ǎ], and falling tone as [â]. Absence of tone on a reflex means that we found no tone in the original source. The reflexes of nominal and verbal protoforms may include a noun class prefix. We morphologically segmented this prefix only if we could corroborate that the prefix is not synchronically fossilized on the simple noun stem.

BLR 69 *báká 'knife'	**B42** *dì-bàɣə̀*, **B43** *dì-báɣə̀*, **B501Y** *báká*, **B61Y** *baɣa*, **B62Z** *ò-báɣá*, **B63Y** *baha*, **B77aZ** *Ø-báká*, **H16c** *bááká̀*.
BLR 198 *bík 'to announce (a death)'	**B43** *ù-bíɣə̀* 'foresee', **B44** *ù-bǐɣə̀* 'foresee', **B72b** *bíol*, **B74Y** bíɔ̀ 'cry out', **B77aZ** *bíikà* 'cry out', **B77bX** *u-biɣa*, **B862X** *bík* 'call', **H111** *bìká* 'foresee', **H16b** *bika* 'announce', **H16c** *bíikúlà* 'announce the cause of a disease', **H16g** *bíkà* 'call, name'.
BLR 200 *bíɪk 'to put (away), bury, lay eggs'	**B44** *ù-bíːkə̀*, **H111** *bíiká* 'leave, abandon', **H16a** *bíikà* 'leave, abandon' **H16c** *bíikà* 'leave, abandon', **H16g** *biika* 'leave, abandon'.
BLR 315 *bʊ́g 'plaster, dig foundations'	**B52Z** *ù-bóːkà* 'dig (generic)', **B62Z** *ò-búɣá* 'plaster', **B81X** *ó-bókà* 'plaster, rub on', **B86P** *buk* 'build a roof'.
BLR 320 *bʊ̀gí 'squirrel'	**B44** *dù-búɣù* 'squirrel (sp.)', **B501Y** *mbùkú*, **B52Y** *mbùxú*, **B73bZ** *Ø-mbúkɔ̀* 'squirrel (sp.)', **B73c** *Ø-mbûkú* 'squirrel (sp.)', **H16c** *lù-búùkù*, **H16g** *lù-bùkú* 'squirrel (sp.)', **H31** *Ø-phúkù* 'rat (gen.), forest rat, squirrel'.

BLR 322 *bʊ́gí 'bundle'	**B73c** *Ø-bûyí*, **B81X** *è-bóké*, **B85FX** *é-bǿǿ*, **B86P** *ì-búú*.
BLR 327 *bʊ́ʊk 'to wake up, rise up, go away, fly away'	**B73c** *vű:* 'escape', **B74Y** *ubvúɔ* 'escape', **B77bX** *úbvúúɣà* 'run away from', **H16g** *buuka* 'run away'.
BLR 372 *búg 'break, snap'	**B80zX** *ò-bwúkà*, **B81Y** *bʊka*, **B821** *kò-búʁò*, **B83Z** *bu* 'smash', **B87T** *ká-bóó*, **B861X** *búyè*, **H111** *bùkùnà* 'cut', **H16a** *búkà*, **H16c** *búúkúnà* 'break in two pieces', **H31** *búúkúnà* 'break in two pieces'.
BLR 418 *càk 'to desire, wish, search for'	**B61Y** *ɣo-saɣa* 'want', **B61Z** *ŋò-ʃá:* 'look for', **B62Z** *ò-sàɣà* 'look for', **B63Y** *gi-saha* 'look for, want', **B72b** *sɔ̀ɔ̀ ~ ʃɔ̀ɔ̀* 'look for', **B73bZ** *ɔ́-sáɣà* 'look for', **B73c** *ú-sáá* 'look for', **B74Y** *ùsɔ̀ɔ̀* 'look for, **B77bX** *ù-sòò* 'look for', **B80zX** *ó-sak* 'touch, feel around', **B83Z** *sɔ́ɔ̀* 'look for', **H16b** *ku-saka* 'desire', want'.
BLR 423 *càká 'thicket, bush-country'	**B43** *dì-sǎkù* 'camp used during dry season', **B52Y** *tsákə́* 'bush', **B53Y** *tsákə́* 'bush', **B74Y** *sɔ̀ɔ́* 'leaf', **B77aX** *Ø-sákì* 'leaf', **B81X** *lè-sàká* 'leaf', **H16g** *nsàkú* 'bush'.
BLR 427 *càkan 'to play'	**B41** *ɣu-sanə*, **B43** *ù-sǎ:nə̀*, **B44** *ù-sáɣə́nə̀* 'play, have fun', **B73bZ** *ɔ́-sákáná*, **B85eX** *nsák* 'game', **B85FX** *nsak* 'game', **B864V** *tsàák* 'game', **B865X** *òsàkàn* 'play, have fun', **H111** *sàkànà* 'play', **H16a** *sákànà* 'play, have fun', **H16g** *sàkànà* 'play', **H31** sákáná 'play, have fun'.
BLR 645 *còk 'poke in, put in, prick with a point, hide'	**B43** *ù-tsǒkə̀*, **H111** *sòkà*, **H16b** *sòkà*, **H16g** *sòkà*, H31 sóká 'make something come out of a hole'.
BLR 647 *còká 'axe'	**B52Y** *tsɔ̀kɔ̀* 'traditional axe', **B85FX** *sɔ́k*, **B86P** *sɔ̀k*, **B861X** ì-*ʃùá*, **B862X** *kə̀-ʃwá*, **B865X** *sɔ̂k*.
BLR 715 *còkí 'hair (on head)'	**B43** *dì-tsǔtsə̀*, **B51X** *lí-tsúɣì*, **B80zY** *lè-sùk*, **B85bV** *syɔ̀ḱ*, **B85FX** *lá-tsø*, **B86R** *là-swìí*, **B86Y** *lù-swě*, **B862X** *lə̀-ʃìí*, **B863Y** *ló-tswì*, **B87**, **H111** *nsuki*, **H16a** *ntsúkì*, **H16b** *lusuki*, **H16c** *lù-súùkì*, **H16g** *lù-sùkí*, **H31** *lù-súkì*.
BLR 822 *dáká 'voice' ~ BLR 823 *dáká 'affair, word'	**B72a** *ndɔ́*, **B73bZ** *Ø-ndáá* 'voice, language', **B73bZ** *Ø-ndáɣá* 'affair, problem', **B74Y** *ndɔ́ɔ́* 'affair, problem', **B77aZ** *n-dáká* 'affair, speech', **B80zX** *ndák* 'affair, problem', **B85bT** *ndak* 'palaver', **B85dZ** *ndâ* 'debate, problem', **B86Y** *ndáá* 'debate, problem', **B862X** *n-dá* 'voice', **B865X** *ndáá* 'story, voice'.
BLR 900 *dègè 'weaver bird'	**B42** *ndékì*, **B43** *Ø-ndékì*, **B501Y** *ndɛ̀kɛ̀*, **B51Z** *ndɛ̀kɛ̀*, **B52Y** *ndɛ̀kɛ̀*, **B53Y** *ndɛ̀kɛ̀*, **B73bZ** *Ø-lyɛ́ɣɛ̀*,

	B73c *Ø-lɛ́ɛ́*, **B77aX** *lèkè*, **B85aX** *ndeok*, **B85dZ** *ndí*, **B85FX** *Ø-ndíí*, **B861X** *è-lɛ̂ʁ*.
BLR 1106 *dògò 'witchcraft'	**B43** *n-dǒɣù*, **B501Y** *n-dɔ̀ɣɔ̀*, **B51Z** *ndɔ̀ɣɔ̀*, **B52Z** *ndɔ́xɔ̀*, **B53Y** *lòɣì*, **B74Y** *ndòʊ̀*, **B77aX** *kì-lòkò*, **B821** *ì-lóɣì*,[23] **B822** *ì-lóʁò*, **B861X** *ì-lɔ̂ʁ*, **H16a** *mu-loko*, **H16b** *ǹlòkò*, **H16c** *nlòókò* 'be under the effect of witchcraft', **H16g** *ǹlòkò*, **H31** *bù-lókí*.
BLR 1107 *dógò 'thing'	**B53X** *ì-lóhò*, **B72b** *è-lúɔ̀*, **B73bZ** *ɛ́-lɔ́ɔ̀*, **B74Y** *ì-lóʊ̀*, **B77aZ** *kì-lɔ́kɔ̀*.
BLR 1175 *dʊ̀gʊ́ 'brother or sister (of the same sex), relative, friend'	**B72a** *ndò* 'friend', **B73bZ** *ì-ndúwò* 'friendship, comradeship', **B73c** *Ø-ndúú* 'friend', **B74Y** *ndʊ̀ù* 'friend', **B77aX** *bù-ndùkù* 'friendship', **B77bX** *Ø-ndùú* 'friend', **B85aX** *nduk mpal* 'false friend', *málùù* 'brother/sister in law', **B864W** *bá-ndùg* 'friends',[24] **H16b** *ndúkù* 'friend, comrade', **H16g** *ndùkù* 'friend', **H31** *ndúkú* 'friend'.
BLR 1178 *dʊ̀k 'to name'	**B42** *ndûɣə* 'name (n.)', **B501Y** *lòɣà*, **B51Z** *lòɣà*, **B61Y** *ɣoduɣa* 'tell', **B61Z** *ŋò-dwá* 'tell, say', **B62Z** *ò-dùɣà* 'say', **B63Y** *gi-duxa* 'say', **B73c** *ú-lúú*, **B74Y** *ùlʊ̀ùɔ̀*, **B77bX** *ú-lúà*, **B80zX** *ò-zúkà* 'call', **B83Z** *lo*, **B85dZ** *kò-lùː*, **B86R** *òdwà* 'announce', **B863Y** *kòlúː*, **B864X** *koluk*, **B87W** *kaluː*, **H111** *lùkà*, **H16a** *luka*, **H16b** *lùkà*, **H16g** *lúkà*, **H31** *lúká*.
BLR 1498 *gʊ̀g 'be fitting, be sufficient'	**B43** *ù-ɣǔkù* 'be used to', **B77bX** *u-kuɣa*, **B81Y** *kʊka* 'be able', **B85dZ** *kòkùː* 'be convenient', **B85eW** *ú-kwóó* 'be sufficient', **B861X** *kǔʁ* 'be fitting', **B862X** *kwə̀* 'be sufficient', **B864X** *kókùk* 'befit, suit', **B87T** *kàkùù* 'be sufficient', **H16c** *dùúkà*, **H31** *kúúká*.
BLR 1621 *yʊ̀gʊ́ 'groundnut'	**B81X** *nzòkó* 'peanuts', **B82X** *n-jòʁú* 'groundnut', **B821** *è-dʒùrú*, **B86U** *ndzuu mbil* 'Bambara groundnut', **B861X** *è-yǔʁ* 'peanut', **B864X** *lò-zùk* 'peanut', **B865X** *è-dzùù* 'groundnut'.
BLR 1682 *kàk 'to tie up'	**B43** *ù-ɣǎɣə̀*, **B44** *ùɣáɣə̀* 'obstruct, surround, block', **B52Y** *ù-kàxà* 'be stuck', **B73c** *ú-kǻkúú*, **H16c**

---

[23] In Mpe B821 and Nunu B822, /ʁ/ > [ɣ] when followed or preceded by /i/. The same allophony is observed by Stappers (1986: 3) for the neighboring North Boma B82X.

[24] Ngong B864 might be a "mixed" or creoloid language strongly influenced by surrounding varieties (see Pacchiarotti, Chousou-Polydouri & Bostoen 2019: 191, footnote 13). Ngong varieties such as B864W spoken in the same area as South-Western languages such as Pende L11, Mbala H41, and Suku H32 manifest the voicing of /k/ in C2 position typical of these South-Western varieties.

	kàákà ‘to obstruct’, **H16g** *kàk* ‘to obstruct’, **H31** *yì-kákú* ‘barrier, obstacle on the road’.
BLR 1762 *kék ‘cut’	**B43** *ù-ɣěɣə̀*, **H16g** *kékà*.
BLR 1901 *kók ‘to pull, to drag’	**B43** *ù-ɣóɣə̀*, **B73c** *kɔ̋ɔ̋*, **B80zX** *òkɔ́k*, **B82X** *kɔ́ʀɔ̀*, **B822** *kóʀò*, **B83Z** *kɔ̀ɔ̀*, **B85bV** *kyɔ́ḱ*, **B864X** *kɔ́kɔ̀k*, **H16a** *kókà*, **H16b** *koka*, **H16c** *kóóká*, **H16g** *kókà*, **H31** *kóká*.
BLR 1903 *kókó ‘crust’	**B43** *dìɣóɣə̀* ‘crust (esp. on head)’, **H16c** *khóókò*, **H16g** *kòkó*.
BLR 2111 *kúkàm ‘to kneel’	**B44** *ù-fúkə̀mə̀*, **B43** *dì-fúɣə̀* ‘misery, bad luck’ ~ **B43** *ùfúkə̀mə̀*, **B73c** *ú-tsűkúnù*, **B85dZ** *fúkám*, **B863Y** *fúkám*, **B864X** *fúkám*, **H111** *fùkàmà*, **H16a** *fúkàmà*, **H16c** *fúúkámà*, **H16g** *fúkámà*, **H31** *fúkámá*.
BLR 2433 *pègà ‘shoulder’	**B80zX** *ípyàk*, **B81X** *è-pɛ̀ɛ̀kè*, **B821** *ì-péʀò*, **B85aX** *pyak*, **B861X** *ì-pɛ́àʀ*, **B865X** *ì-pèk*.
BLR 2513 *pìkà ‘slave’	**B42** *mú-βíɣə̀*, **B43** *mù-βǐɣə̀*, **B44** *mù-βíɣə̀*, **B501Y** *βèɣà*, **B51Y** *véɣà*, **B52Z** *mù-vèxà*, **B53Y** *βèhà*, **B61Z** *ò-yéɣà*, **B85bV** *mpík̀*, **B85FX** *mpìk*, **B864X** *móyík*, **B865X** *mpîk*, **H16a** *mvíkà*, **H16c** *mvììkà*, **H31** *m̀-phíká*.
BLR 2634 *pʊ̀gí ‘pot’	**B80zX** *mpwɔ̀k*, **B821** *mpʊ̀ʀí*, **B822** *mpʊ̀ʀó*, **B85bT** *mpwuk* ‘cooking pot’, **B85eX** *mpfye é mɛn*, **B85FX** *mpøø* ‘cooking pot’, **B861X** *Ø-mpʊ̆ʀ*, **B865X** *mpfyě* ‘cooking pot’.
BLR 2638 *pʊ̀ká ‘insect, bee, ant, caterpillar’	**B73bZ** *mɔ́-fúúkɔ̀* ‘caterpillar (sp.)’, **B73c** *mú-pfúúkù* ‘caterpillar (sp.)’, **B862X** *pfʊ̀k* ‘biting ant’, **H16g** *nɣúkà* ‘caterpillar’, **H31** *m-pfúúká* ‘caterpillar’.
BLR 2825 *tég ‘to set (a trap)’	**B53Y** *téɣè*, **B85bT** *kulyak mwɛd*, **B85dZ** *kɔtɪ mwɛ̂r*, **B86U** *kù-lɛ́ɛ́*, **B863Y** *kɔtɪ mbwɛ̂r*, **B864X** *kɔ̀tɛ́k* myɛ̌t, **B87T** *kàtéé*.
BLR 2911 *tígad ‘to remain’	**B43** *ùsyáːlə̀*, **B44** *ùsyáːlə̀*, **B72a** *síl*, **B82X** *kɔ̀-síɣə̀rà*, **B83Z** *syálà*, **H111** *sààlá*, **H16a** *sáálà*, **H16c** *syáálà*, **H16g** *sáálà*, **H31** *sáálá*.
BLR 2967 *tòk ‘to boil up, bubble up’	**B43** *ù-rǒɣə̀*, **B44** *ù-rɔ́ɣə̀*, **B501Y** *tɔ̀ɣɔ̀*, **B52Z** *ú-tɔ̀xɔ̀*, **B73c** *ú-tɔ́ɔ́*, **B74Y** *ù-tɔ́ɔ̀*, **B77bX** *ú-tòò*, **B80zX** *ò-tók*, **B864X** *kɔ̀-tɔ́k*, **H111** *tòkìsà* ‘make boil’, **H16a** *tókèsà*, **H16g** *tòkà*.
BLR 3047 *tóg ‘bale out (water), draw (water)’	**B501Y** *rókâ* ‘bale out water’, **B51Z** *rókà* ‘fish à la nasse’, **B80zX** *ó-tùkùbà*, **B861X** *tôʀ* ‘draw water’.
BLR 3169 *jákà ‘year, cultivation season, harvest’	**B43** *mw-âːɣə̀* ‘shoot from a banana tree’, **B81X** *mù-ákà* ‘year’, **B82X** *mw-áːʀà* ‘year’, **B87T** *mya-ak*

	'manioc tuber', **H16a** *mw-aka* 'year, season', **H16c** *mwáàkà* 'season', **H16g** *mwààká* 'former times'.
BLR 3171 *jákí 'egg'	**B43** *dyǎkì*, **H16c** *dííkì*, **H16a** *dyákì*, **H16b** *diaki*, **H16g** *dyàkí* **H31** *díkí*.
BLR 3291 *jégam 'lean against'	**B44** *ù-wɛ́kəmə*, **B83Z** *yemoː*, **B863Y** *yɛ́kám*, **H111** *yèkàmà*, **H16a** *yékàmà*, **H16b** *yèkàmà*, **H16c** *yèèkámà*, **H16g** *yékámà*, **H31** *yékámá*.
BLR 3441 *jìkad 'dwell, be, sit, stay'	**B42** *utsaɣələ*, **B43** *mù-tsîɣə̀* 'autochthonous person', **B44** *ù-kálə̀*, **B52Y** *tsǎxə̀là*, **B53Z** *utsaala*, **B73bX** *á-dzàálà*, **H111** *zàkàlà*, **H16a** *kálà*, **H16c** *káálà*, **H16g** *kàlà*, **H31** lù-káálà 'surroundings'.
BLR 3444 *jíkò 'fireplace, country'	**B43** *diku*, **B44** *dí-kù*, **B51Z** *dííkù*, **B81X** *dì-íkè*, **B73c** *Ø-díkì* 'fireplace', **H16c** *dì-zíìkú*, **H16g** *zíkù*, **H31** *zíkù*.
BLR 3445 *jíkʊt 'be satiated'	**B43** *ù-ɣúkùrù*, **B44** *ù-wúkúrù*, **B85bT** *kuyuur*, **B85dZ** *kò-wúr*, **B85eX** *ù-yúr*, **B85FX** *ká-wúr*, **B861X** *yúò*, **B863Y** *kó-wúr*, **B864X** *kò-wúr*, **B865X** *òzúùr*, **B87W** *kà-wúr*, **H111** *yùkùtà*, **H16a** *wúkùtà*, **H16c** *yùùkútà*, **H31** *yúkútá*.
BLR 3525 *jóg 'bathe, wash, swim'	**B61Y** *ɣo-yoɣo* 'swim', **B61Z** *ŋò-yɔ́ɣɔ̀* 'wash', **B62Z** *ò-jɔ́ɣɔ́* 'swim, drink', **B77bX** *ù-yóò* 'wash oneself', **B80zX** *ò-zwɔ́kɔ̀b*, **B85aY** *zɔkœb* 'swim', **B861X** *jôʁ* 'wash oneself', **B862X** *jwə* 'swim', **B865X** *o-zwô* 'to bathe (intr.)'.
BLR 3528 *jógà 'fear'	**B61Y** *bvwoɣo*, **B61Z** *ò-vúwɔ́*, **B62Y** *mvuɔɣɔ*, **B72b** *bùɔ́*, **B74Y** *bùɔ́*, **B77aZ** *bwɔ̀kɔ́*, **B77bX** *Ø-ngwo*, **B80zX** *vwok*, **B81X** *bɔ̀-ɔ́kɔ̀*, **B821** *bùʁó*, **B87W** *mwɔy*.
BLR 3536 *jókà 'snake, intestinal worm'	**B41** *Ø-ɲoɣə* 'snake', **B42** *nyɔ̀ɣə̀* 'snake', **B43** *Ø-ɲóɣə̀* 'snake', **B44** *Ø-ɲɔ̀ɣə̀* 'snake', **B861X** *Ø-ndʒúà* 'snake', **B865X** *ò-dzwó* 'snake', **B87T** *nyɔ́k* 'worm', **H111** *n-yókà* 'snake', **H16a** *nyókà* 'snake', **H16b** *nyoka* 'snake', **H16c** *nyóókà* 'snake', **H31** *Ø-nyókà*.
BLR 5333 *pʊkʊ 'burrow, cavern'	**B85eX** *Ø-fû* 'hole', **B86U** *efuu* 'hole', **B861** *ì-pûʁ* 'garbage hole', **B865X** *i-fùú* 'hole'.
BLR 5428 *nuku 'meat without bones'	**B80zX** *mù-nùk* 'meat, fish', **B81X** *Ø-nùká* 'meat', **B821** *mù-ɲúʁù*, **B822** *mù-núʁù*, **B82X** *mù-nùʁù* 'meat', **B85bV** *ɲùk̀* 'meat'.
BLR 5455 *túká 'banana *Musa* sp.'	**B41** – **B42** *mu-tuka* 'species of plantain with big bunch of purple fruits'; **B43** *dì-tǔkə̀* 'ball of banana mash', **B52W** *mu-tuka* 'French/False Horn Giant/ Medium Red-Green Chimaera Subhorizontal',

	B80zX *kɛ-tshúka* ‘bunch of fruits’, **B85dZ** *mó-tsuː* ‘bunch of bananas’, **B85FX** *ɛ́-tswa* ‘bunch (of bananas)’, **B86E** *e-súk* ‘bunch’; **B86W** *i-tswa* ‘bunch’; **B861X** *è tsûʁ* ‘banana bunch’; **B865X** *e-twâ* ‘banana cluster’, **H16a** *m-fuka* ‘plantain stalk’, **H16g** *ǹfúka* ‘trunk of banana plant’, **H31** *m-fúkì* ‘bunch of bananas’.
BLR 6106 *cɪk ‘to play music, to beat drum’	**B52Z** *ù-sîîxà* ‘to play an instrument’, **B61Y** *ɣosiɣa* ‘sing’, **B73c** *ú-sı̋ı̋* ‘to play an instrument’, **H111** *kùsíkà* ‘to play an instrument’, **H16a** *síkà* ‘to play an instrument, music’, **H16c** *sííká* ‘to play an instrument’, **H16g** *síkà* ‘to play an instrument’.
BLR 6108 *cìkà ‘girl, woman’	**B85dZ** *mòʃíː* ‘virgin’, **B85eW** *ɔ̋-syâ* ‘young woman’, **B85FX** *màʃáː* ‘virgin’, **B86U** *mu-syɛ* ‘virgin’, **B861X** *ò-səʁ̂* ‘young, unmarried woman’, **B863Y** *múʃáː* ‘virgin’.
BLR 6196 *tígʊé ‘young orphan’	**B80zX** *kɛ̀-tɛ́kɛ́*, **B822** *è-tíkè*, **B85dZ** *é-tìː*, **B86R** *kè-tý*, **B861X** *è-tsɤ́ʁ*, **B862X** *kə̀-tʃí*, **B87W** *è-tíː*.
BLR 7413 *cókì ‘saliva’	**B821** *nsóɣì*, **B82X** *ǹtʃóɣì*, **B861X** *à-súì*.
BLR 9582 *dák ‘to walk’	**B73bZ** *láɣà* ‘to climb’, **B82X** *làʁà* ‘to walk fast’, **B821** *kò-láʁà*, **B822** *láʁà*, **B85aY** *lakær* ‘to climb’, **B861** *lăʁ*.
BLR 9590 *dákò ‘house (for men)’	**B41** *Ø-ndaɣu* ‘house’, **B42** *Ø-ndaɣu* ‘house’, **B43** *Ø-ndáɣù* ‘house’, **B52Z** *mù-lákà* ‘camp’, **B821** *Ø-ndáʁò*, **B862X** *lák* ‘fireplace’.
BLR 9629 *cakɪd ‘clap one’s hands’	**B43** *ùsáːɣə̀* ‘acclaim, celebrate’, **B62Z** *à-ntsáɣí* ‘palms of hands’, **B72b** *ȁntsóò* ‘acclamation’, **B73c** *sa̋a̋* ‘joke’, **B74Y** *àntsoʊ* ‘acclamation’, **B77aZ** *mà-ntsákà* ‘acclamation’, **B85aY** *nsak* ‘hand clapping’, **B85eW** *sák* ‘happiness’, **B86Y** *mù-sák* ‘happiness’, **B861X** *Ø-ntsâʁ* ‘thanking’, **B862X** *nsàk* ‘happiness’, **H16a** *sàkìlà*, **H16g** *sàkìl*.

Appendix B. Language varieties and sources in this study*

	Language variety	WCB sub-branch	Source
1	Shira **B41**	KLC extended_KLC_North-West	(ALGAB)**
2	Sangu **B42**	KLC extended_KLC_North-West	(Idiata-Mayombo 1993)
3	Punu **B43**	KLC extended_KLC_North-West	(Blanchon 2008; Mavoungou & Plumel 2010)
4	Lumbu **B44**	KLC extended_KLC_North-West	(Mavoungou & Plumel 2010)
5	Wanzi (Mayela) **B501Y**	Kasai-Ngounie_Nzebi-Teke West	(Hombert & Mouélé 1988; Mouélé 1997)
6	Duma (Bembikani) **B51Y**	Kasai-Ngounie_Nzebi-Teke West	(Mickala-Manfoumbi 1988)
	Duma (Bembikani/ Lastoursville) **B51Z**	Kasai-Ngounie_Nzebi-Teke West	(Mouélé 1997)
7	Nzebi (?) **B52W**	Kasai-Ngounie_Nzebi-Teke West	(Rossel 1998)
	Nzebi (Mbigou) **B52Y**	Kasai-Ngounie_Nzebi-Teke West	(Mouélé 1997)
	Nzebi (Lébamba) **B52Z**	Kasai-Ngounie_Nzebi-Teke West	(Marchal-Nasse 1989)
8	Tsaangi (Madouma) **B53X**	Kasai-Ngounie_Nzebi-Teke West	(Loubelo 1987)
	Tsaangi (Lekoko) **B53Y**	Kasai-Ngounie_Nzebi-Teke West	(Mouélé 1997)
9	Mbete (Ndjounou) **B61Y** [Congo]	Kasai-Ngounie_ Mbete	(Lane 1989)
	Mbete ('Obaa', Tsaama I) **B61Z** [Congo]	Kasai-Ngounie_ Mbete	(Ndouli 2001)
10	Mbaama (Okondja) **B62Z** [Gabon]	Kasai-Ngounie_ Mbete	(Okoudowa 2016)

	Language variety	WCB sub-branch	Source
	Mbaama (Sibiti) **B62Y** [Congo]	Kasai-Ngounie_ Mbete	(Lane 1989)
11	Nduumo ('Kuya') **B63Y**	Kasai-Ngounie_ Mbete	(Biton 1969)
12	Ngungwel (Gamboma) **B72a**	Kasai-Ngounie	(Rurangwa 1982)
	Ngungwel ('Mpumpu') **B72b**	Kasai-Ngounie	(Raharimanantsoa 2016, pers. comm.)
13	Laali (Mayeye) **B73bZ**	Kasai-Ngounie_Nzebi-Teke West	(Bissila 1991)
14	Yaa (Bihoua) **B73c**	Kasai-Ngounie_Nzebi-Teke West	(Mouandza 2001)
15	Eboo-Nzikou **B74Y**	Kasai-Ngounie	(Raharimanantsoa 2012a, 2012b, 2017, pers. comm.)
16	Kukwa (West Plateau) **B77aX**	Kasai-Ngounie	(Paulian 1975)
	Kukwa **B77aZ**	Kasai-Ngounie	(Daeleman's archive; Raharimanantsoa pers. comm.)***
17	Fumu (Ngamaba) **B77bX**	Kasai-Ngounie	(Makouta-Mboukou 1969, 1976)
18	Boma Yumu (Pentane/ Mondai) **B80zX**	Kasai-Ngounie_Kwa-Kasai North	(Burssens 1999)
	Boma Yumu (Ito) **B80zZ**	Kasai-Ngounie_Kwa-Kasai North	(Burssens 1999)
19	Tiene ('Dya', Mansele) **B81X**	Kasai-Ngounie_Kwa-Kasai North	(Ellington 1977; Bastin et al. 1999)
	Tiene ('Nkɛtɛ') **B81Y**	Kasai-Ngounie_Kwa-Kasai North	(Motingea Mangulu 2004)
20	North Boma (Mushie) **B82X**	Kasai-Ngounie_Kwa-Kasai North	(Stappers 1986)
	North Boma (Mpukumbu) **B82Z**	Kasai-Ngounie_Kwa-Kasai North	(Bastin et al. 1999)
21	Mpe (Bolebe) **B821**	Kasai-Ngounie_Kwa-Kasai North	own fieldwork
22	Nunu (Mushie) **B822**	Kasai-Ngounie_Kwa-Kasai North	own fieldwork

	Language variety	WCB sub-branch	Source
23	Mfinu (Yuo) **B83Y**	Kasai-Ngounie extended	(Bastin et al. 1999; Daeleman's archive)****
	Mfinu **B83Z**	Kasai-Ngounie extended	(Daeleman's archive)
24	West Yans (Mukonkie) **B85aX**	Kwilu-Ngounie	(Swartenbroeckx 1948)
	West Yans (Makwa) **B85aY**	Kwilu-Ngounie	(Mayanga 1985)
	East Yans **B85bR**	Kwilu-Ngounie	(Swartenbroeckx 1948)
	East Yans (Nkara/Niadi) **B85bS**	Kwilu-Ngounie	(Nguma 1986)
	East Yans (Nkara) **B85bT**	Kwilu-Ngounie	(Impubi Mukwa 1987), (Koni Muluwa & Bostoen 2015)
	East Yans (Mantshiene) **B85bV**	Kwilu-Ngounie	(Rottland 1977)
25	East Nsong ('Luniungu', Kipuka) **B85dZ**	KLC extended	(Koni Muluwa 2010; Koni Muluwa & Bostoen 2015)
26	Mpur (Kwebe) **B85eX**	Loange-Atlantic_ Kamtsha-Kwilu	(Koni Muluwa & Bostoen 2015)
	Mpur (Due I) **B85eW**	Loange-Atlantic_ Kamtsha-Kwilu	(Kibwenge India'Ane 1985)
27	Nsambaan (Kwilumpia/ Longo K.K.) **B85FX**	Loange-Atlantic_Kamtsha-Kwilu	(Koni Muluwa 2015b)
28	East Ngwi (Mangai) **B861X**	WCB	own fieldwork
29	East Lwel (Sedzo) **B862X**	WCB	(Khang Levy 1979; Koni Muluwa & Bostoen 2015)
30	Mpiin (Kipuka) **B863Y**	KLC extended	(Koni Muluwa 2010; Koni Muluwa & Bostoen 2015)
31	Ngong (Kwenge) **B864X**	KLC extended	(Koni Muluwa 2010, 2015a; Koni Muluwa & Bostoen 2015)
	Ngong (Lukula) **B864W**	KLC extended	(Pokoso 1986)
32	Nzadi (Indolo) **B865X**	WCB	(Crane, Hyman & Tukumu 2011; Koni Muluwa & Bostoen 2015)

	Language variety	WCB sub-branch	Source
33	West Ding (Mateko) **B86P**	WCB	(Mwan Mesongolo 1984)
	East Ding (Ipamu) **B86Q**	WCB	(Mufanga-Dzmar 1977; Munkyen Okab 1990)
	East Ding (Bantshione) **B86R**	WCB	(Mula 1977)
	East Ding ('Mbentsie', Bambudi) **B86U**	WCB	(Koni Muluwa & Bostoen 2015)
	West Ding (Sedzo) **B86Y**	WCB	(Ebalantshim Masuwan 1980)
34	Mbuun (Idiofa) **B87T**	KLC extended	(Dibata Mimpiya 1977; Mundeke 1979, 2011)
	West Mbuun (Imbongo) **B87W**	KLC extended	(Koni Muluwa 2010, 2014; Koni Muluwa & Bostoen 2015)
35	Sikongo **H16a**	KLC extended_ KLC_South	(Van Gheel 1652; KongoKing fieldwork; Narciso Cobe 2010)
36	Hangala **H111**	KLC extended_KLC_North	(Nkouanda 1997; Nguimbi-Mabiala 1999)
37	Yombe **H16c**	KLC extended_KLC_South-West	(De Grauwe 2009)
38	Ntandu **H16g**	KLC extended_KLC_East	(Daeleman 1983, n.d., Daeleman's archive)
39	Manyanga **H16b**	KLC extended_KLC_Central	(Laman 1912; Laman & Meinhof 1928–1929; Laman 1936)
40	Yaka **H31**	KLC extended_KLC_Kongoid	(Van Den Eynde 1968; Ruttenberg 2000)

* The language varieties and sources listed in Appendix B are identical to those found in Pacchiarotti & Bostoen (2020: 190–194) with the exception of Tyee B73d which we were unfortunately not able to include in this study. Nevertheless, based on the lexical data found in Raharimanantsoa and Ntsiba Ngolo (2015) and additional data kindly provided to us by Ruth Raharimanantsoa (pers. comm.), Tyee appears to have /ɣ/ (with [h] as an allophone in between two high vowels) and zero as reflexes of PWCB *k in C2. Varieties included in this study but not in Pacchiarotti, Chousou-Polydouri & Bostoen (2019) are shaded in gray in Table 1.

** ALGAB stands for the *Atlas Linguistique du GABon* project supervised by Prof. Lolke Van der Veen at Université Lumière Lyon 2 (cf. http://www.ddl.ish-lyon.cnrs.fr/equipes/index.asp?Langue=FR&Equipe=8&Page=Action&ActionNum=48). Basic vocabulary collected for this language atlas was also used in recent phylogenetic studies (cf. de Schryver et al. 2015; Grollemund et al. 2015; Pacchiarotti, Chousou-Polydouri & Bostoen 2019).

*** The reference (Daeleman's archive) stands for the legacy of Jan Daeleman's research data that was trusted to Ghent University in 2018. Jan Daeleman was a Jesuit father who spent most of his life in the DRC and collected data on multiple Bantu languages spoken there.
**** The basic vocabulary which Bastin et al. (1999) used for their seminal lexicostatistical study of the Bantu language family is currently available on the website of the Royal Museum for Central Africa in Tervuren: https://www.africamuseum.be/nl/research/discover/human_sciences/culture_society/lexicostatistic-study-bantu-languages.

Appendix C. Unconditioned reflexes of PWCB *k in C2 in our WCB convenience sample

	REFLEX 1 in C2				REFLEX 2 in C2				REFLEX 3 in C2			
Variety	PWCB *k	Tokens	%	Total %	PWCB *k	Tokens	%	Total %	PWCB *k	Tokens	%	Total %
Shira **B41**	ɣ	*k: 9/12	75	77	k	*k: 2/12	17	15	Ø	*k: 1/12	8	8
		*g: 1/1	100			*g: 0/1	0			*g: 0/1		
Sangu **B42**	ɣ	*k: 36/-42	86	79	k	*k: 5/42	12	19	Ø	*k: 1/42	2	2
		*g: 5/10	50			*g: 5/10	50			*g: 0/10	0	
Punu **B43**	ɣ	*k: 47/58	81	76	k	*k: 10/-58	17	21	Ø	*k: 1/58	2	3
		*g: 23/34	68			*g: 9/34	26			*g: 2/34	6	
Lumbu **B44**	ɣ	*k: 31/43	72	69	k	*k: 12/-43	28	28	Ø	*k: 0/43	0	3
		*g: 9/15	60			*g: 4/15	27			*g: 2/15	13	
Wanzi **B501**	ɣ	*k: 19/26	73	63	k	*k: 5/26	19	33	Ø	*k: 2/26	8	4
		*g: 13/25	52			*g: 12/25	48			*g: 0/25	0	
Duma **B51**	ɣ	*k: 16/25	64	60	k	*k: 8/25	32	38	Ø	*k: 1/25	4	2
		*g: 9/17	53			*g: 8/17	47			*g: 0/17	0	
Nzebi **B52**	x	*k: 28/39	72	70	k	*k: 11/-39	28	30				
		*g: 17/25	68			*g: 8/25	32					
Tsaangi **B53**	h	*k: 18/22	82	77.5	k	*k: 4/22	18	22.5				
		*g: 13/18	72			*g: 5/18	28					
Mbete **B61**	ɣ	*k: 23/31	74	67	Ø	*k: 8/31	26	33				
		*g: 12/21	57			*g: 9/21	43					
Mbaama **B62**	ɣ	*k: 17/18	94	96	k	*k: 1/18	6	4				
		*g: 10/10	100			*g: 0/10	0					
Nduumo **B63**	x	*k: 46/50	92	95	Ø	*k: 3/50	6	4	k	*k: 1/50	2	1
		*g: 29/29	100			*g: 0/29	0			*g: 0/29		
Ngungwel **B72a**	Ø	*k: 20/20	100	100								
		*g: 18/18	100									
Laali **B73b**	Ø	*k: 16/33	48	51	ɣ/h	*k: 13/-33	39	39	k	*k: 4/33	12	10
		*g: 10/18	56			*g: 7/18	38			*g: 1/18	6	
Yaa **B73c**	Ø	*k: 41/48	85	87	k	*k: 7/48	15	13				
		*g: 20/22	91			*g: 2/22	9					

	REFLEX 1 in C2				REFLEX 2 in C2				REFLEX 3 in C2			
Variety	**PWCB *k**	**Tokens**	**%**	**Total %**	**PWCB *k**	**Tokens**	**%**	**Total %**	**PWCB *k**	**Tokens**	**%**	**Total %**
Eboo-Nzikou **B74**	Ø	*k: 37/37	100	100								
		*g: 21/21	100									
Kukwa **B77a**	g	*k: 24/24	100	97	Ø	*k: 0/24	0	3				
		*g: 12/13	92			*g: 1/13	8					
Fumu **B77b**	Ø	*k: 20/-40	50	55	ɣ	*k: 20/40	50	45				
		*g: 15/24	62.5			*g: 9/24	37.5					
Boma Yumu **B80z**	k	*k: 27/29	93	94	Ø	*k: 2/29	7	6				
		*g: 19/20	95			*g: 1/20	5					
Tiene **B81**	k	*k: 30/30	100	96	Ø	*k: 0/30	0	4				
		*g: 20/22	91			*g: 2/22	9					
North Boma **B82**	ʁ	*k: 21/22	95	95	*k>k	C2*k: 1/-22	5	5				
		*g: 18/19	95		*g >Ø	C2*g: 1/19	5					
Mpe **B821**	ʁ	*k: 20/21	95	96	Ø	*k: 1/21	5	4				
		*g: 24/25	96			*g: 1/25	4					
Nunu **B822**	ʁ	*k: 18/18	100	92	Ø	*k: 0/18	0	8				
		*g: 18/20	90			*g: 2/20	10					
Mfinu **B83**	Ø	*k: 25/25	100	100								
		*g: 12/12	100									
Yans **B85**	k	*k: 48/48	100	98.5	Ø	*k: 0/48	0	1.5				
		*g: 22/23	96			*g: 1/23	4					
Nsong **B85d**	Ø	*k: 21/33	64	67	k	*k: 12/33	36	33				
		*g: 11/16	69			*g: 5/16	31					
Mpur **B85e**	Ø	*k: 17/23	74	70	k	*k: 6/23	26	30				
		*g: 11/17	65			*g: 6/17	35					
Nsambaan **B85F**	Ø	*k: 19/26	73	70	k	*k: 7/26	27	30				
		*g: 12/18	67			*g: 6/18	33					
Ngwi **B861**	ʁ	*k: 18/19	95	92	Ø	*k: 1/-19	5	8				
		*g: 15/17	88			*g: 2/17	12					
Lwel **B862**	Ø	*k: 19/25	76	73	k	*k: 6/25	24	27				
		*g: 11/16	69			*g: 5/16	31					
Mpiin **B863**	Ø	*k: 16/24	67	63	k	*k: 8/24	33	37				
		*g: 6/11	54.5			*g: 5/11	45.5					
Ngong **B864**	k	*k: 17/17	100	96	Ø	*k: 0/17	0	4				
		*g: 9/10	90			*g: 1/10	10					
Nzadi **B865**	Ø	*k: 25/31	81	72	k	*k: 6/31	19	28				
		*g: 11/19	58			*g: 8/19	42					
Ding **B86**	Ø	*k: 29/34	85	79	k	*k: 5/34	15	21				
		*g: 13/19	68.5			*g: 6/19	31.5					

	REFLEX 1 in C2				REFLEX 2 in C2				REFLEX 3 in C2			
Variety	PWCB *k	Tokens	%	Total %	PWCB *k	Tokens	%	Total %	PWCB *k	Tokens	%	Total %
Mbuun **B87**	Ø	*k: 28/33	85	78	k	*k: 5/33	15	22				
		*g: 7/12	58			*g: 5/12	42					
Hangala **H111**	k	*k: 40/40	100	96	Ø	*k: 0/40	0	4				
		*g: 7/9	78			*g: 2/9	22					
Sikongo **H16a**	k	*k: 64/64	100	94	Ø	*k: 0/64	0	6				
		*g: 9/14	64			*g: 5/14	36					
Manyanga **H16b**	k	*k: 40/40	100	90	Ø	*k: 0/40	0	10				
		*g: 14/20	70			*g: 5/20	30					
Yombe **H16c**	k	*k: 55/55	100	95	Ø	*k: 0/55	0	5				
		*g: 14/18	78			*g: 4/18	22					
Ntandu **H16g**	k	*k: 95/96	99	95	Ø	*k: 1/96	1	5				
		*g: 17/22	77			*g: 6/22	23					
Yaka **H31**	k	*k: 40/40	100	92	Ø	*k: 0/40	0	8				
		*g: 9/13	70			*g: 4/13	30					

Appendix D. Assessing lexical items which might have escaped frication at some Kasai-Ngounie node*

		Kasai-Ngounie branch									
		Nzebi-Teke West						Mbete			Fumu
		B501	B51	B52	B53	B73b	B73c	B61	B62	B63**	B77b
1.	BLR 518 *cégé* ‘grassland’	*tséɣé*	*?*	*?*	*tséhè*	*Ø-nséɣè*	–	*?*	*?*	*li-tʃege*	*Ø-ntseɣe*
2.	BLR 820 *dáká* ‘language’	*ndaɣa*	*ndaɣa*	*ndáxà*	*?*	*Ø-ndáá*	*ndắắ*	*?*	*?*	*Ø-ndaxa*	*Ø-ndaɣa*
3.	BLR 1044 *dìɪk* ‘bury’	*nzèèkà*	*?*	*?*	*ù-dììh-á*	*ɔ́-dzííɣɛ̀*	*díí*	*ŋo-jéɣa*	*ò-dʒììɣ-à*	*gi-djixa*	*u-dziɣa*
4.	BLR 1100 *dòg* ‘bewitch’	*lɔ̀ɣo*	*lɔ̀ɣo*	*?*	*ù-lóhò*	*ɔ́-lɔ́ɔ̀*	*ù-lɔ́ɔ́*	*ŋo-ló*	*ò-lɔ́ɣ-ɔ̀*	*gi-loxo*	*ú-lòɣò*
5.	BLR 1685 *kààká* ‘ancestor’	*?*	*kaɣa*	*ngáxà*	*kààɣá*	*Ø-nkáɣà*	*?*	*Ø-ŋkáɣá*	*?*	*nkaxa*	*nkàɣà*

		Kasai-Ngounie branch									
		Nzebi-Teke West						Mbete			Fumu
		B501	**B51**	**B52**	**B53**	**B73b**	**B73c**	**B61**	**B62**	**B63****	**B77b**
6.	BLR 2180 _*mìg_ 'try'	_mìɣà_	_?_	_mèɣ-à_	_mèɣ-à_	–	–	_ŋo-míɣa_	_?_	_gi-mixa_	_u-miɣa_
7.	BLR 7216 _*còɪk_ 'hide, cover'	_?_	_swɛ̀ɛ̀ɣɛ̀_	_?_	_?_	_ɔ́-swɛ́ɛ́ɣɛ̀_	_u-swɛ̀ː_	_?_	_?_	_gi-tʃwege_	_u-sweɣe_
8.	BLR 2568 _*pígò_ 'kidney'	_li-piku_	_βíkù_	_?_	_?_	–	_mú-pfĩ́_	_?_	_?_	_li-pfiɣi_	–
9.	BLR 1607 _*jògù_ 'elephant'	_nzòkù_	_nzókù_	_Ø-nzòxà_	_nzòhà_	–	–	_njoː_	_Ø-ndzoo_	_Ø-ndʒoxo_	_Ø-nzòò_
10.	BLR 900 _*dègè_ 'weaver bird'	_ndɛ̀kɛ̀_	_ndɛ̀kɛ̀_	_ndɛ̀kɛ̀_	_ndɛ̀kɛ̀_	_Ø-lyɛ́ɣɛ̀_	_Ø-lɛ́ɛ́_	_?_	_?_	_Ø-ndege_	_?_
11.	BLR 320 _*bʊ̀gí_ 'squirrel'	_mbùkú_	_?_	_mbùkú_	_?_	_Ø-mbúkɔ_	_Ø-mbûkú_	_?_	_?_	_Ø-mbuxu_	–
12.	BLR 1830 _*kígì_ 'eyebrow'	_kíkî_	_kíkì_	_kíkì_	_kíkì_	–	–	_?_	_?_	_o- kigi_	_mi-kiu_
13.	BLR 1355 _*gègò_ '(molar) tooth'	_kèkù_	_kèkù_	_kèkə̀_	_kèkà_	–	_Ø-kékù_	_?_	_?_	_Ø-kegi_	_?_
14.	BLR 3612 _*jʊki_ 'smoke'	_?_	_úúkì_	_mù-ùkì_	_úúkî_	_Ø-lúúwɛ̀_	_lí-lṹkì_	_yuɣi_	_yuɣi_	_Ø-yugi_	–
15.	BLR 3444 _*jíkò_ 'fireplace'	_nziku_	_diiku_	_?_	_?_	_Ø-dzí_	_Ø-dĩ́kì_	_?_	_?_	–	–
16.	BLR 9642 _*káká_ 'foot' > 'hand'	_káká_	_?_	_lè-kákə̀_	_káká_	_lí-kákà_	–	_ŋgaɣa_	_?_	_li-kaxa_	_kee (?)_
17.	BLR 2362 _*páágʊ̀_ 'tree fork'	_?_	_pákú_	_páká_	_?_	–	_lí-pã́ká_	_?_	_?_	_li-pagi_	_?_

	Kasai-Ngounie branch									
	Nzebi-Teke West						Mbete			Fumu
	B501	B51	B52	B53	B73b	B73c	B61	B62	B63**	B77b
18. BLR 711 **cʊ̀k* 'wash'	*tsoka*	*tsoka*	*u-tsoka*	*?*	*ɔ́-swá-áɣ-à*	*swáː*	*?*	*ò-tʃuka*	*gi-tsuka*	*swaɣa*
19. BLR 760 **cúg* 'support'	*sukû*	*súkà*	*?*	*?*	–	–	–	*?*	*gi-tʃuxa*	*?*

* In Appendix D, "–" means there is another root for a given concept in a given variety and "?" means lack of data.
** Biton (1969: 555) writes that <*h*> is realized *ch* as in German *suchen*, i.e., [*x*]. According to our understanding of the Nduumo B63 data in Biton (1969), <*h*> is in complementary distribution with <*g*>, which seems to appear only in front of <*e*>, <*i*> and <*u*>. Following Biton (1969: 555), <*g*> would always be like in French *goût*, i.e., [*g*], but Medjo Mvé (1989) reports that it is realized as [*ɣ*] in between vowels.

Address for correspondence

Sara Pacchiarotti
Department of Languages and Cultures
UGent Centre for Bantu Studies (BantUGent)
Blandijnberg 2
9000 Ghent
Belgium
sara.pacchiarotti@ugent.be

Co-author information

Koen Bostoen
Department of Languages and Cultures
UGent Centre for Bantu Studies (BantUGent)
koen.bostoen@ugent.be

Publication history

Date received: 27 October 2020
Date accepted: 5 September 2021