Toward a subclassification of the !Ui branch of Tuu[1]

Tom Güldemann
Humboldt University Berlin and Max Planck Institute for the Science of Human History Jena

1 Introduction

+ Tuu family as isolated unit in the Kalahari Basin, clear genealogical separation from the two other families Khoe and Kx'a subsumed earlier under "Khoisan" (Güldemann 2014a)
+ Tuu first established as a unit under the label "Southern Bushman" and subjected to survey work by D. Bleek (e.g., 1927, 1929, 1939/40, 1956) > **Table 1**
> !Ui varieties distributed over four of six units in her reference classification

Acro-nym	Label	Location	Researcher(s)	Docu-lect(s)*
[SI]	|xam	Northern part of Cape south of Orange River east and west	Lichtenstein	**4**
			W. Bleek, Lloyd	***|Xam***
SIa	a dialect	Oudtshoorn	Lloyd	5
[SII]	ǁŋ	Langebergen in Griqualand	D. Bleek	**21**
SIIa	ǂkhomani	Northern Gordonia	Doke, Maingard	***N‖ng***
SIIb	ǁkxau	Near Kimberley	Meinhof	20
SIIc	ǁku‖e	Near Theunissen	D. Bleek	17
SIId	seroa	Southern part of Orange Free State, near Bethany	Arbousset	15
			Wuras	16
SIIe	!gã !ne	Transkei	Anders	12
SIII	batwa	Lake Chrissie, eastern Transvaal	D. Bleek	14
SIV	|auni	Country between Nossop and Auhoup, S. Kalahari	D. Bleek	-
SIVa	khatia (xatia)	East of Nossop, S. Kalahari	D. Bleek	-
SIVb	ki|hazi	West of Auhoup, S. Kalahari	Story	-
SV	masarwa	Kakia in the south of Bechuanaland Protectorate	Schultze	-
			D. Bleek	-
SVI	|nu ‖en	Upper Nossop and Auhoup, S.W.A.	D. Bleek	-
SVIa	|nusan	South of Auhoup, S.W.A.	Krönlein	1

Note: [...] = corrected typo, * see !Ui survey in Table 2 below, - Tuu other than !Ui

Table 1: Tuu reference classification of older sources according to Bleek (1956)

[1] I gratefully acknowledge the support of this presentation by JSPS KAKENHI 16H01925 and the help of Hans-Jörg Bibiko (MPI-SHH Jena) in producing Map 1.

+ highly precarious data situation due to: (I) early extinction of most Tuu languages and (II) quantitatively and qualitatively limited records on extinct languages > available data:
a) two surviving language complexes with extensive modern documentation:
Taa in Kalahari of Botswana and Namibia (endangered) (cf. SV + SVI)
N‖ng in southern Kalahari, South Africa (moribund) (cf. SII + SIIa)
b) one extinct language complex with extensive but old documentation:
\|Xam in Karoo, South Africa (cf. SI + SVIa)
c) two extinct languages with less extensive, more recent documentation:
ǂUngkue on Lower Vaal River, South Africa (cf. SIIb)
‖Xegwi in eastern Transvaal, South Africa (cf. SIII)
d) many archival data sets with highly fragmentary information distributed across the entire range of Tuu
> up to now focus on the five units under a)-c) and neglect of archival !Ui data under d)

+ current state of Tuu classification:
- demonstration of genealogical unity of Tuu and external separation vis-à-vis other "Khoisan": Köhler (1975: 316-7), Traill (1975), Hastings (2001), Güldemann (2005)
- bipartite Tuu classification with !Ui as one of two primary branches, as opposed to D. Bleek's original break-up into six units: Westphal (1971), Köhler (1981)
- demonstration of closer relation of Lower Nossob doculects [SIV] to Taa complex [SV, SVI], excluding them from !Ui: Güldemann (2002, 2014b)
> resulting Tuu classification > **Figure 1**

Branch and subbranch	**Language (complex)**	**Further subclassification**
Taa-Lower Nossob		
Taa	single unit:	West: West !Xoon, (N|u‖'en [SVI])
		East: East !Xoon, 'N|oha, (N|amani), (Kakia [SV]), etc.
Lower Nossob	(|'Auni [SIV])†	
	(|Haasi [SIVb])†	
!Ui		
	N‖ng:	West: N|uu = (ǂKhomani [SIIa] = N|huki), etc.
		East: (Langeberg [SII]), etc.†
	(ǂUngkue [SIIb])†	
	(‖Xegwi [SIII])†	
	(|Xam [SI])†:	Strandberg, Katkop, Achterveld, etc.
	(Other)†:	‖Ũ‖'e [SIIc], Seroa [SIId], !Gã!ne [SIIe]

Notes: † = extinct, (...) = older data source, [...] = D. Bleek's doculect acronym

Figure 1: Preliminary internal classification of Tuu (after Güldemann 2014a, b)

2 Cross-!Ui comparison

2.1 Linguistic data and methodology

+ large amount of data are small unpublished archival corpora that are deficient but still contain some historically diagnostic information that needs to be organized and analyzed
> important concept of "doculect" as a distinct corpus defined by the time and place of documentation, the person recording the data, and the speaker(s) contributing the data:

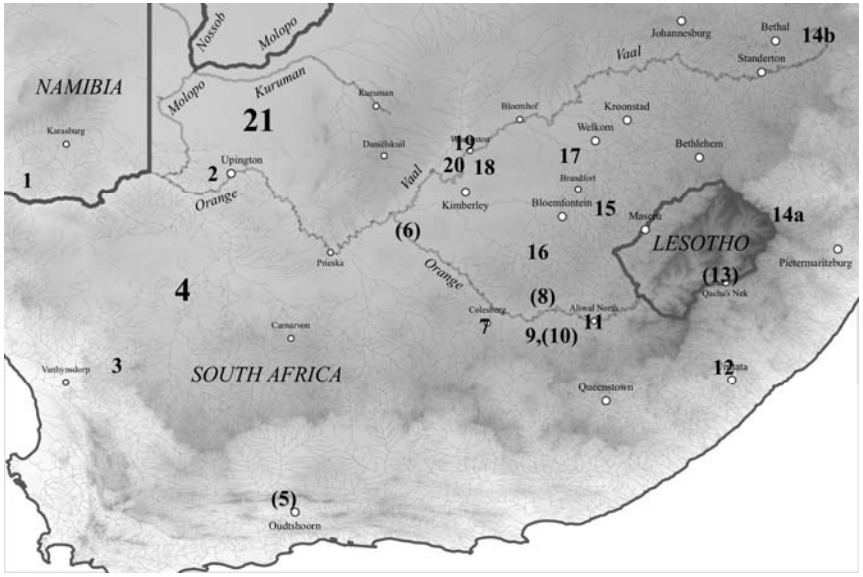
> a linguistic variety as it is documented in a given resource. This term is deliberately agnostic as to whether or not that variety can straightforwardly be associated with a particular 'language' or 'dialect' and, instead, merely focuses on the fact that there is a document either about the relevant variety or directly recording that variety in some way. (Cysouw and Good 2013:342)

+ first more comprehensive list of relevant !Ui doculects > **Table 2**

No.	Doculect name	Date	Origin [recording location] *
1	N|uusaa by Krönlein	1850s	Lower Orange River [Bethany, Namibia]
2	N|usa by Lloyd	(1880)	Middle Orange River [Cape Town]
3	|Xam by W. Bleek	1866	Achterveld [Cape Town]
4	***|Xam*** **by W. Bleek/Lloyd**	1870s	Karoo [Cape Town]
(5)	!Ui by Anderson	(?)	? [Oudtshoorn]
(6)	!Ui by Smith	1835	S of Douglas and N of Hopetown
7	!Ui by W. Bleek	(1857)	Colesberg [Cape Town]
(8)	!Ui by C. S. Orpen	1877	Bethulie
9	!Ui by W. Bleek	(1857)	Burghersdorp [Cape Town]
(10)	!Ui by Kannemeyer	1890	Burghersdorp
11	!Ui by Lloyd	(1880)	Aliwal North [Cape Town]
12	!Gã!ne by Anders	1920+	Tsolo district
(13)	!Ui by J. M. Orpen	1873	N of Qacha's Nek, Lesotho
14	**‖Xegwi**	1950+	?E of Maluti mountains [Lake Chrissie]
15	!Ui by Arbousset	1836	San places Mokhasi/Puchane
16	!Ui by Wuras	1836+	Bethany
17	‖Ũ‖'e by D. Bleek	(1928)	Theunissen
18	!Ui by Maingard	(1930+)	Boshof
19	‖Kã by D. Bleek	(1920+)	Warrenton
20	ǂUngkue by Meinhof	1929	Warrenton-Windsorton
21	***N‖ng***	2000s	Southern Kalahari

Notes: (n) = excluded doculect with insufficient data, **BOLD** = more than one doculect, *ITALIC* = dialect cluster, (...) = unpublished, * all in South Africa if not noted otherwise

Table 2: An overview of !Ui doculects



Map 1: Geographical distribution of !Ui doculects

+ some individual !Ui doculects can already be joined into larger more coherent clusters:
1) **‖Xegwi** with several doculects based on the same group of speakers in the same location (14b in Map 1) but recorded at different times (1930s-1980s) by different researchers
> represented here under 14 by Honken's (2007) first data collation
2) **|Xam** with more than ten archival doculects (cf. Güldemann 2004)
> represented here under 4 by the central Bleek-Lloyd corpus and under 1-3 by three geographically peripheral doculects (cf. Güldemann 2006, Vosseler 2014)
3) **N‖ng** with more than ten archival doculects (cf. Güldemann 2017)
> represented here under 21 by the modern data collected since the 2000s (notably field notes, Sands et al. 2006)
> classificatory situation in the eastern half of the !Ui area remains entirely unclear

+ grammatical data are commonly assumed to be more diagnostic for historical-comparative analysis but are virtually absent in most doculects of Table 2
> problem hoped to be partly compensated by looking at lexical data that are paradigmatically organized and involve some frozen morphology:
a) personal pronouns > §2.2
b) quantifiers > §2.3
c) basic human and kinship terms > §2.4

2.2 Personal pronouns

+ basic system homogeneous in Tuu as a whole, i.e. including Taa-Lower Nossob branch
> seven non-diagnostic Proto-!Ui items (five also Proto-Tuu, cf. Güldemann 2005)

Person	Singular	Plural
First inclusive		*i
First exclusive	*N	*si
Second	*a	*u
Third	*ha, *hi(N)	

Table 3: The pronoun system of Proto-!Ui

+ one localized innovation attested in the doculects 1, 3-7: oblique series for speech-act participants, notably 1SG *-ke* (cf. Güldemann (2013: 242) for narrow |Xam)
> unclear significance for !Ui branch as a whole due to lack of grammatical data on most other doculects but good support for the distinction of |Xam from the doculects of the N‖ng cluster (21) and ǂUngkue (20), where this feature is evidently absent

2.3 Quantifiers

+ maximum for primary native elements conveying exact cardinal concepts is three:

(1)	Exact cardinal	Non-exact cardinal	Other meaning
a.	‘one’	-	‘alone’
b.	‘two’	-	-
c.	‘three’	?‘more than two’	-
d.	-	‘many’ (count noun)	‘much’ (mass noun), ‘big’

> all attested expressions with higher values are transparently secondary and young in being borrowed or derived from lower simplex forms, for example, rarely present form for 'four' either derived (conceptually from 2 + 2) or borrowed (Khoekhoe *haka*)
> restricted numeral systems as a wider areal feature of the Kalahari Basin (cf. Güldemann and Fehn 2017)

+ comparative situation almost inverse of that in §2.2 for pronouns: despite the small inventory, not a single Proto-Tuu item and very few reconstructions even on branch levels
> proto-stages without genuine part-of-speech class of numerals (cf. Güldemann in press)

+ two non-diagnostic !Ui items: Proto-!Ui form *!'uu 'two'; recurrent *n!ona* 'three' that was probably borrowed from Khoekhoe multiple times independently

+ two meanings involve diagnostic innovations:

(2)	Meaning	Predominant in !Ui	Doculects	Local innovation	Doculects
a.	'one'	*!oa(i)	1, 3, 4, 11, 14	*‖’oe	19-21
b.	'many'	*|kx'oai	3, 4, 11, 14, 16	*n!ai < 'big, much'	19-21

2.4 Basic human and kinship terms

+ survey of more than 20 basic human and kinship terms (parentheses: meanings that turned out to lack sufficiently recurrent and specific items):

('aunt'), ('boy'), 'brother', ('brother-in-law'), ('girl'), 'child/children as infant', 'child/children as offspring', 'daughter', 'father', 'female' ~ FEMININE, 'grandfather', 'grandmother', ('husband'), 'male' ~ MASCULINE, 'man/men', 'mother', 'name', 'person/people', 'sister', ('sister-in-law'), 'small' ~ DIMINUTIVE, 'son', ('uncle'), ('wife/wives'), 'woman/women'

> includes many items known in Tuu to be associated with morphology (cf. Boden 2014a, 2014b; Boden, Güldemann and Jordan 2014)

+ various types of differential traits that are potentially diagnostic for classification:
different lexemes for a meaning
different morphemes for identical derivation and number inflection
different collocation of lexical and morphological items
morphological change in shared lexemes
sound change in shared lexemes
semantic change in shared lexemes

+ 14 non-diagnostic !Ui items that can be reconstructed to Proto-!Ui:

(3)	Meaning	Proto-!Ui
a.	'child (offspring)' ~ DIMINUTIVE	*⊙aa
b.	'child (offspring) F ~ daughter'	*⊙aa-FEMININE
c.	'child (offspring) M ~ son'	*⊙aa-MASCULINE
d.	FEMININE	*-xae
e.	'grandparent F ~ grandmother'	*!o(b)i.te
f.	'grandparent M ~ grandfather'	*!oi-MASCULINE
g.	MASCULINE	*-Ṽ ~ *-õ
h.	'name'	*|ãe
i.	'person'	*!ui
j.	'people'	*ǂ(')ee
k.	'people' ~ 'men' ~ 'who'	*tuu
l.	'sibling F ~ sister'	*‖aa-FEMININE
m.	'sibling M ~ brother'	*‖aa-MASCULINE
n.	'woman/women'	*|(')aa.ti/*|aa-PL

+ four meanings involve diagnostic innovations:

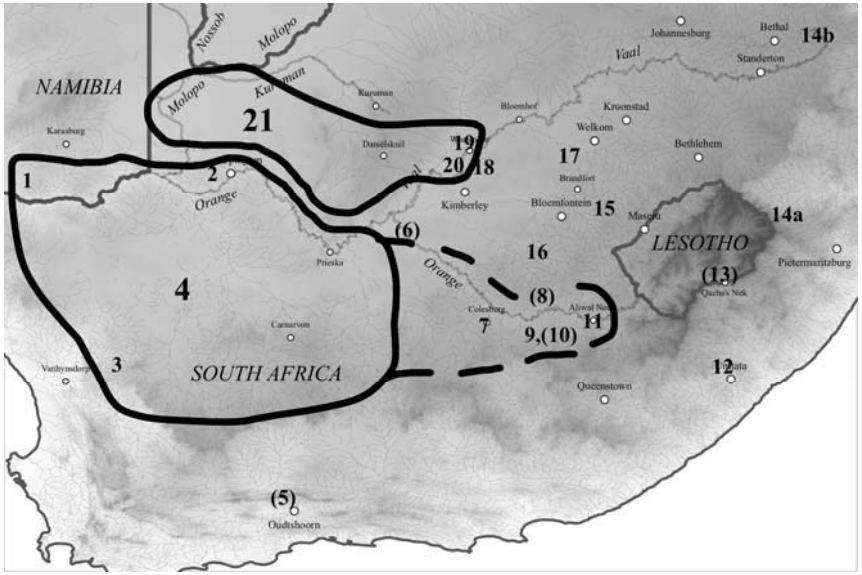
(4)	Meaning	Predominant in !Ui	Doculects	Local innovation	Doculects
a.	'child (infant	*|oba/	12, 16-21	*!(kh)wãa/	1, 4, 11
	or offspring)'	*|oe-PL		*!(')ao-PL	3, 4, 9, 11
b.	'father'	*ãa ~ aa	7, 11-16		
	(?Tuu: *ã$^{\varsigma}$a)	*õa ~ oa	1-9, 17, 18	*ãa.ti	19, 21
c.	'male, man'	*ǂoo	14-21	*goai	1, 3, 4, 9
d.	'mother'	*xãV ~ xaV	11, (16)		
	(?Tuu: *kã$^{\varsigma}$a)	*(k)xõa ~ xoa	1, 2, 4-9, 14, 17	*xãa.ti	19, 21

3 Discussion

3.1 Two larger doculect clusters

a) wider |Xam cluster suggested by three innovations: oblique pronouns (§2.2); *goai 'male, man, husband', *!(kh)wãa/*!(')ao-PL 'child' (§2.4) (cf. already Güldemann 2006)
> appears to include Upper Orange doculects: pronouns 7, lexemes 9 and 11

b) Ghaap-Kalahari cluster comprising N‖ng (21) and adjacent Danster !Ui (19, 20) suggested by four innovations: *‖'oe 'one', *n!ai 'many' (§2,3); *ãa.ti 'father', *xãa.ti 'mother' (§2.4) (cf. Güldemann 2017)



Map 2: Geographical distribution of !Ui with two western doculect clusters

+ circumstantial information on geographical boundary between the two clusters, for example, regarding doculect 6 (which unfortunately contains hardly any data itself):

> The Bushmen here say that should they come together with the Bushmen about Daniel's Kuyl they would meet as friends, but they would not comprehend each other. (Kirby 1940: 282)

> escarpment of Ghaap Plateau as possible boundary between Ghaap-Kalahari cluster in the west and !Ui on the lower plains and along the river courses further south(east), whereby the entire Orange River may have been settled on both sides by the wider |Xam cluster

+ doculects in the area east of the Vaal and Orange Rivers with unclear relationship to the two clusters as well as to each other - largely due to lack of data

3.2 A first internal classification

+ abandonment of D. Bleek's tripartite reference structure of !Ui in terms of SI, SII, and SIII
+ relocation of individual doculects, notably of 1 = SVIa to |Xam (cf. Güldemann 2006)
+ two relatively robust western clusters: |Xam (1-4, ?7, ?9, ?11); Ghaap-Kalahari (19-21)
> larger homogeneity of western clusters vs. eastern diversity suggests family spread from east out of a homeland in the wider area from Lesotho up to the Vaal and Orange Rivers

Ghaap-Kalahari

N‖ng [21]	West: N|uu = (ǂKhomani [SIIa] = N|huki), etc.
	East: (Langeberg [SII]), etc.
(Danster)†	ǂUngkue [20 = SIIb], ‖Kã [19]
(Eastern core)†:	Boshof [18], ‖Õ‖'e [17 = SIIc], Seroa [15+16 = SIId], ‖Xegwi [14 = SIII], !Gã!ne [12 = SIIe]
(Wider |Xam)†:	Aliwal North [11], Burghersdorp [9], Colesberg [7], Strandberg-Katkop [4 = SI], Achterveld [3 = SI], Orange N|usa [2], N|uusaa [1 = SVIa]

Notes: † = extinct, (...) = older data source, [...] = doculect no. and/or D. Bleek's acronym
Figure 2: Preliminary internal classification of !Ui

3.3 Brief outlook

+ necessary further research for testing first results by extension to other lexical fields:
- items promising to involve morphology, notably body part and related terminology (cf. Güldemann 2005, Güldemann and Loughnane 2012)
- generally stable vocabulary (cf. Tadmor 2009)

+ diagnostic isoglosses across classificatory boundaries > research on language contact!

(5)	Mainstream |Xam (4)	Trans-Orange |Xam by Lloyd (2)	(Western) N‖ng (21)
a.	*‖khã$^{\varsigma}$a* 'lion'	*!khǫ́.tye* 'lion'	*!qhoe* 'lion'
b.	-	*kebe* 'four~more than three'	*kebe.ke* 'many'

References

Barnard, Alan and Gertrud Boden (eds.). 2014. Southern African Khoisan kinship systems. Quellen zur Khoisan-Forschung 30. Köln: Rüdiger Köppe.

Bleek, Dorothea F. 1927. The distribution of Bushman languages in South Africa. In Boas, Franz et al. (eds.), Festschrift Meinhof. Glückstadt/ Hamburg: J. J. Augustin, 55-64.

Bleek, Dorothea F. 1929. Comparative vocabularies of Bushman languages. Cambridge: Cambridge University Press.

Bleek, Dorothea F. 1939/40. A short survey of Bushman languages. Zeitschrift für Eingeborenene-Sprachen 30: 52-72.

Bleek, Dorothea F. 1956. A Bushman dictionary. American Oriental Series 41. New Haven, Connecticut: American Oriental Society.

Boden, Gertrud. 2014a. N‖ng kinship terminology – salvage documentation with the last speakers. In Barnard and Boden (eds.), 63-84.

Boden, Gertrud. 2014b. Tuu kinship classifications – a diachronic perspective. In Barnard and Boden (eds.), 161-184.

Boden, Gertrud, Tom Güldemann and Fiona Jordan. 2014. Khoisan sibling terminologies in historical perspective: a combined anthropological, linguistic and phylogenetic comparative approach. In Güldemann and Fehn (eds.), 69-102.

Cysouw, Michael and Jeff Good. 2013. Languoid, doculect, glossonym: formalizing the notion “language”. Language Documentation and Conservation 7: 331-359.

Güldemann, Tom. 2002. Using older Khoisan sources: quantifier expressions in Lower Nosop varieties of Tuu. South African Journal of African Languages 22,3: 187-196.

Güldemann, Tom. 2004. Introduction to "Bushman grammar: a grammatical sketch of the language of the |xam-ka-!k'e" by Dorothea F. Bleek. In Hollmann, Jeremy C. (ed.), Customs and beliefs of the |Xam Bushmen. Johannesburg: Witwatersrand University Press with the Ringing Rocks Press, 385-387.

Güldemann, Tom. 2005. Studies in Tuu (Southern Khoisan). University of Leipzig Papers on Africa, Languages and Literatures 23. Leipzig: Institut für Afrikanistik, Universität Leipzig.

Güldemann, Tom. 2006. The San languages of southern Namibia: linguistic appraisal with special reference to J. G. Krönlein’s N|uusaa data. Anthropological Linguistics 48,4: 369-395.

Güldemann, Tom. 2013. Morphology: |Xam. In Vossen, Rainer (ed.), The Khoesan languages. London: Routledge, 241-249.

Güldemann, Tom. 2014a. “Khoisan” linguistic classification today. In Güldemann and Fehn (eds.), 1-41.

Güldemann, Tom. 2014b. The Lower Nossob varieties of Tuu: !Ui, Taa or neither? In Güldemann and Fehn (eds.), 257-282.

Güldemann, Tom. 2017. Casting a wider net over N‖ng: the older archival resources. Anthropological Linguistics 59,1: 71-104.

Güldemann, Tom. in press. Did Proto-Tuu have a paradigm of cardinal numerals? In Beyer, Klaus, Gertrud Boden, Bernhard Köhler and Ulrike Zoch (eds.), 40 Jahre Afrikanistik: Festschrift für Rainer Vossen. Köln: Rüdiger Köppe, 119-134.

Güldemann, Tom and Anne-Maria Fehn. 2017. The Kalahari Basin area as a “Sprachbund” before the Bantu expansion. In Hickey, Raymond (ed.), The Cambridge handbook of areal linguistics. Cambridge: Cambridge University Press, 500-526.

Güldemann, Tom and Anne-Maria Fehn (eds.). 2014. Beyond ‘Khoisan’: historical relations in the Kalahari Basin. Current Issues in Linguistic Theory 330. Amsterdam: John Benjamins.

Güldemann, Tom and Robyn Loughnane. 2012. Are there “Khoisan” roots in body-part vocabulary? On linguistic inheritance and contact in the Kalahari Basin. In Güldemann, Tom et al. (eds.), Methodology in linguistic prehistory. Language Dynamics and Change 2,2: 215–258.

Hastings, Rachel. 2001. Evidence for the genetic unity of Southern Khoisan. In Bell, Arthur and Paul Washburn (eds.), Khoisan: syntax, phonetics, phonology, and contact. Cornell Working Papers in Linguistics 18. Ithaca: Cornell University, 225-246.

Honken, Henry. 2007. Short Grammar and Dictionary of ‖Xegwi. Unpublished ms.

Kirby, Percival R. (ed.). 1939/1940. The diary of Dr. A. Smith 1834-1836, 2 vols. Publications 20/21. Cape Town: Van Riebeeck Society.

Köhler, Oswin. 1975. Geschichte und Probleme der Gliederung der Sprachen Afrikas: Von den Anfängen bis zur Gegenwart. In Baumann, Hermann (ed.), Die Völker Afrikas und ihre traditionellen Kulturen, Teil 1: Allgemeiner Teil und südliches Afrika. Studien zur Kulturkunde 34. Wiesbaden: Franz Steiner, 135-373.

Köhler, Oswin. 1981. Les langues khoisan, section 1: présentation d'ensemble. In Perrot, Jean (ed.), Les langues dans le monde ancien et moderne, première partie: les langues de l'afrique subsaharienne. Paris: Centre National de la Recherche Scientifique, 455-482.

Nordhoff, Sebastian and Harald Hammarström. 2011. Glottolog/langdoc: defining dialects, languages, and language families as collections of resources. In Kauppinen, Tomi, Line C. Pouchard and Carsten Keßler (eds.), Proceedings of the First International Workshop on Linked Science 2011 (LISC2011), Bonn, Germany, 24 October 2011. http://CEUR-WS.org/Vol-783/paper7.pdf.

Sands, Bonny, Amanda Miller, Johanna Brugman, Levi Namaseb, Chris Collins and Mats Exter. 2006. 1400 item N|uu dictionary. Unpublished manuscript.

Tadmor, Uri. 2009. Loanwords in the world's languages: findings and results. In Haspelmath, Martin and Uri Tadmor (eds.), Loanwords in the world's languages: a comparative handbook. Berlin/ New York: De Gruyter Mouton, 55-75.

Traill, Anthony. 1975. Phonetic correspondences in the !Xõ dialects: how a Bushman language changes. In Traill, Anthony (ed.), Bushman and Hottentot linguistic studies. Communications 2. Johannesburg: African Studies Institute, University of the Witwatersrand, 77-102.

Vosseler, Annika. 2014. Eine Analyse des Achterveld |Xam Korpus von Wilhelm Bleek, 1870. B.A. thesis: Humboldt-Universität zu Berlin.

Westphal, Ernst O. J. 1971. The click languages of southern and eastern Africa. In Berry, Jack and Joseph H. Greenberg (eds.), Linguistics in Sub-Saharan Africa. Current Trends in Linguistics 7 (ed. by T. Sebeok). The Hague/ Paris: Mouton, 367-420.