

---

**(Un)classifying Shabo:  
phylogenetic methods and results**

**Tyler Schnoebelen**

---

Proceedings of Conference on  
**Language Documentation & Linguistic Theory 2**

Edited by Peter K. Austin, Oliver Bond, Monik Charette,  
David Nathan & Peter Sells

13-14 November 2009 School of Oriental and African Studies, University of London

Hans Rausing Endangered Languages Project  
Department of Linguistics  
School of Oriental and African Studies  
Thornhaugh Street, Russell Square  
London WC1H 0XG  
United Kingdom

Department of Linguistics:  
Tel: +44-20-7898-4640  
Fax: +44-20-7898-4679  
linguistics@soas.ac.uk  
<http://www.soas.ac.uk/academics/departments/linguistics>

Hans Rausing Endangered Languages Project:  
Tel: +44-20-7898-4578  
Fax: +44-20-7898-4349  
elap@soas.ac.uk  
<http://www.hrelp.org>

© 2009 Tyler Schnoebelen

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, on any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the author(s) of that part of the publication, except as permitted by UK copyright law.

ISBN: 978-0-7286-0392-9

This publication can be cited as:

Tyler Schnoebelen. 2009. (Un)classifying Shabo: phylogenetic methods and results. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS.

or:

Tyler Schnoebelen. 2009. (Un)classifying Shabo: phylogenetic methods and results. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS. [www.hrelp.org/eprints/ldlt2\\_28.pdf](http://www.hrelp.org/eprints/ldlt2_28.pdf)

# **(Un)classifying Shabo: phylogenetic methods and results**

TYLER SCHNOEBELEN

*Stanford University*

## 1. INTRODUCTION

Shabo is one of our greatest classification puzzles. The handful of researchers who have worked on it have tended to place it within the Nilo-Saharan phylum, though where it should fall within that phylum has been left a mystery. This paper aims to pursue the Nilo-Saharan hypothesis by reviewing previous evidence and adding new data from a month spent living among the Shabo in a remote village in the coffee-growing mountains of southwest Ethiopia.

Beyond documenting a highly endangered language (there are perhaps 400 speakers of Shabo), this paper offers an innovative use of the phylogenetic tools that biologists have developed for their classification needs. I demonstrate the potential power of such methods in building classifications and exploring the increasing amounts of typological data available to us. These tools can be used heuristically to develop hypotheses: could Shabo be related to Fur, Nubian, or Ngiti? In this case, no convincing ‘phylogenetic signal’ emerges. And the evidence makes a Nilo-Saharan hypothesis especially unlikely. Based on our current understanding, it is better to call Shabo an isolate like Basque or Burushaski (see also Hombert & Philippson, this volume).

## 2. PREVIOUS CLASSIFICATIONS

Because Shabo isn’t very well described, it is possible to give a comprehensive overview of what people have said about its classification. Despite the fact that the Shabo live between the major phyla of Nilo-Saharan and Afroasiatic, no one seems to seriously consider the possibility that it is an Afroasiatic language - though connections to Omotic are enough to need some explaining. The major proposals are that it is (a) Nilo-Saharan, (b) its own phylum, (c) unclassified. Within Nilo-Saharan, the best guesses are that it is related to the Koman languages (Gule, Kwama, Opuuo, Uduk), but has borrowed most extensively from a neighboring Surmic language, Majang, among whose speakers the Shabo live.<sup>1</sup>

The Nilo-Saharan hypothesis for Shabo seems to come down to three pieces of evidence: a plural morpheme, *-k*, an accusative marker (also *-k*), and many NS resemblances in the basic vocabulary. There are problems with all three of these.<sup>2</sup> Perhaps the most important point is one that Ehret (1995) raised: there are too

---

<sup>1</sup> See Bender (1977, 1983), Teferra and Unseth (1989), Teferra (1991, 1995), Fleming (1991), and Ehret (1995).

<sup>2</sup> For a longer discussion of past evidence and future prospects, see the full version of this paper at [http://www.stanford.edu/~tylers/notes/qp/Unclassifying\\_Shabo\\_QP\\_Schnoebelen\\_Jun-3-09.pdf](http://www.stanford.edu/~tylers/notes/qp/Unclassifying_Shabo_QP_Schnoebelen_Jun-3-09.pdf).

many similarities to too many different languages. No one has actually shown any sort of correspondence sets or consistencies.<sup>3</sup>

Despite the weakness of the evidence, we may still not want to reject the Nilo-Saharan hypothesis. In an attempt to move classification efforts forward and perhaps even help re-imagine the internal structure of NS, I move on to structural features - phonological/morphological/syntactic. These have not typically been mentioned in the NS literature but turn out to be relatively good markers of genetic relationships.

### 3. CLASSIFICATION WITH STRUCTURAL FEATURES

Wichmann and Holman (forthcoming) take data from the World Atlas of Language Structures ('WALS' - Haspelmath et al 2005) and calculate which features are the most stable across the groups that we are most confident are related to one another. I take their most stable features (e.g., 'Is there an initial velar nasal', 'What's the word order', 'Is there an inclusive/exclusive distinction?') and use them to show which languages Shabo has an affinity to. The WALS data is somewhat sparse, so I have added my fieldwork data on Shabo, Shekkacho, and Majang; through a literature review, I have also added data points for more than a dozen other languages to help fill in the picture, effectively adding hundreds of new datapoints.<sup>4</sup>

The two main classifications of Nilo-Saharan are by Bender (for example, 2000) and Ehret (2001). The two systems are very different, but both agree that there are two 'complex' subgroups where the majority of languages go - one is 'Eastern Sudanic', the other 'Central Sudanic'.

If Shabo is Nilo-Saharan, it should be related to *something*. I proceed, then, by first imagining that Shabo is an Eastern Sudanic language. If so, which language is it most structurally similar to? Next, I imagine that Shabo is a Central Sudanic language and compare it with other Central Sudanic languages - where would it fit? Finally, if I consider it relative to the 'other' Nilo-Saharan languages, which affinities do we see?

It is not feasible to compare the amount of language data we are working with manually.<sup>5</sup> But phylogenetics programs give us ways to find patterns in the data and propose relationships. SplitsTree4 (Huson 1998) and MrBayes (Huelsenbeck

---

<sup>3</sup> I have also looked at how Shabo may fit in with the patterns of shared innovations that Bender (1989/1991) has identified. There is very little compelling evidence, certainly nothing consistent. At this point, it could all be explained away by coincidence or contact.

<sup>4</sup> I collected data for 66 features that Wichmann and Holman had identified as being stable. After removing features that had no variance, few entries, or that were highly correlated, I ended up with 47 features. To see these features and a guide to coding them, see [http://www.stanford.edu/~tylers/notes/qp/Scoring\\_stable\\_features\\_6-4-09.pdf](http://www.stanford.edu/~tylers/notes/qp/Scoring_stable_features_6-4-09.pdf).

<sup>5</sup> Given just 50 languages, there are more possible tree topologies than there are atoms in the known universe.

and Ronquist 2001) are two of the programs used by both biologists and computationally-oriented historical linguists.<sup>6</sup> Because of the relative novelty of these methods in linguistics, I use these tools heuristically. That is, they generate hypotheses about where to look, which I then consider in greater detail.

### 3.1. *An introduction through Eastern Sudanic*

Eastern Sudanic is one of the best understood branches of Nilo-Saharan. Even those who question the existence of a Nilo-Saharan family take these languages to be related to one another (see Sands 2009, for example).<sup>7</sup>

The most traditional visualization of genetic relationships is the family tree, which makes very strong claims - each branch of a tree indicates a speciation event where some subset of taxa (languages, bacteria, plants, or mammals) broke off from others.

But often there is uncertainty - the data suggests several different possible trees. In this case, it is useful to visualize relationships as a network of interconnections. In these representations, the length of lines between two taxa tell us how related they are likely to be; the amount of 'webbing' tells us how confident we should be - more webbing indicates more noise or uncertainty.

If we put the data for all the Eastern Sudanic languages (plus Shabo) into SplitsTree4, we get the network of relationships in Figure 4. The amount of webbing here tells us that there are a number of different possibilities given the data.<sup>8</sup> SplitsTree4 is capable of running a variety of algorithms to compute relationships; the one used here is the Neighbor-Net algorithm (Bryant and Moulton 2004).

The first thing to recognize is that Neighbor-Net is a 'distance-based method'. It is therefore subject to critiques that it uses less information than other methods because it is using distances, not the data itself. As mentioned earlier, if the algorithm cannot find an optimal tree, it leaves the issue unresolved (the webbing in the middle).

From the network on Figure 4 it is pretty clear that Shabo is an outlier. Phylogenetic programs will try to relate any set of taxa that we give them. So if we included Basque or Japanese with the Eastern Sudanic languages, it would try to figure out where they best belong, too. None of the other languages are as far

---

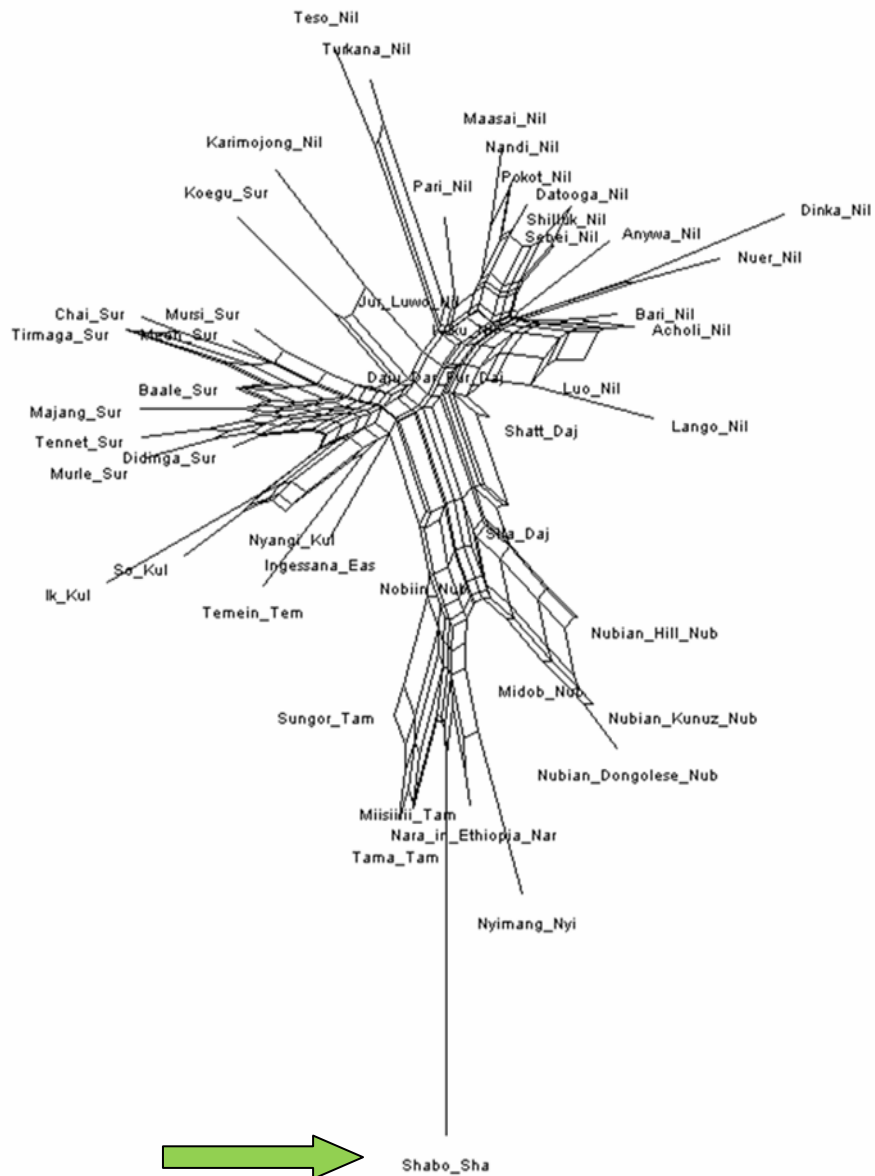
<sup>6</sup> Wichmann and Saunders (2007) compare these and a number of other methods. They like NeighborNet for visual analysis; their preference for overall accuracy is Bayesian analysis. Another excellent source is Nichols and Warnow (2008). Saunders (2005) is also a good reference. For a step-by-step how-to manual for using phylogenetic software with linguistic data, see [http://www.stanford.edu/~tylers/notes/qp/Linguistic\\_phylogenetics\\_4-23-09.pdf](http://www.stanford.edu/~tylers/notes/qp/Linguistic_phylogenetics_4-23-09.pdf).

<sup>7</sup> It also has perhaps the most complicated subgrouping: Nilotic. There has actually been a fair amount of work done on the internal classification of Nilotic, though there is still some outstanding controversy. Having reviewed proposed proto-forms for Nilotic and its subgroups, I am confident that Shabo will not fit with these languages.

<sup>8</sup> It would likely reduce if we (i) filled in more of the empty cells, and (ii) gave it additional, stable features (adding lexical data will be an important next step).

away - I do not think anyone would propose taking Dinka (on the right) and Teso (at the top) out of Nilo-Saharan. The Kuliak languages (like Ik and So located on the left of the network), on the other hand, have been taken out of Eastern Sudanic, and indeed out of Nilo-Saharan by some.

**Figure 3**  
 Eastern Sudanic+Shabo; here the language names are appended by the subgrouping ('Nil'='Nilotic', 'Sur'='Surmic', etc.)



The information here makes it look like there is a cluster of Taman, Nubian, and Nyimang languages and that Shabo best fits in with them. Keeping that in mind, we switch to a rather different method: Bayesian inference using software called MrBayes. Bayesian methods try to estimate the probability that each tree is the true tree. They do not produce a single tree, but a probability distribution on the set of trees.

MrBayes will consider millions of different hypothetical genetic relationships and evaluate them based on the data that we have given it. In our case, our data is all the patterns of ‘Who has initial velars,’ ‘Who has five cases,’ ‘Who has postpositions,’ etc. One possible hypothesis for the relationships of Indo-European would be: ‘Swedish and Sanskrit are a group separate from Norwegian and Persian’. That hypothesis will probably have a very low probability - that is, the probability of that hypothesis given the data is going to be vanishingly small (unless there is something wrong with our data or with the reconstruction of Indo-European). We prefer hypotheses with high probability and remove ones that posit proto-Swedish and proto-Norwegian.

When we put our ‘Eastern Sudanic plus Shabo’ data into MrBayes, it tells us which relationships showed up repeatedly in the optimal trees (only reporting nodes that exist in more than 50% of the optimal trees - that is, eliminating the pre-convergence trees).<sup>9</sup>

In this case, 97% of the optimal trees grouped Shabo with Nara, Nyimang, and Dongolese Nubian. Nichols and Warnow (2007: 778) discuss a way to judge phylogenetic methods by looking at what they call ‘compatible resolution’, that is, the tree may not have all of the subgroupings we expect, but it should not mix established subgroups (like placing Persian with Norwegian). The other part of the evaluation is to not miss any defined groups: our data, unfortunately, misses most of the subgroups that have been claimed. In Nichols and Warnow’s terms, this is a desirable property, but not as essential. Here are the relationships that MrBayes posits, given our data:

- Dinka and Nuer are closely related (98% of trees). These languages are part of ‘Western Nilotic’ and in fact form their own subgroup (‘Dinka-Nuer’).
- Teso and Turkana are closely related (97% of trees). Not only are these Nilotic languages, they are actually part of the ‘Teso-Turkana’ subgroup of the ‘Lotuxo-Teso’ group within Nilotic.
- Nara, Nubian, Tama, and Nyimang are related (we’ll hold off on Shabo for a moment) (97% of trees). Bender (2000) actually puts these languages in the ‘Northern’ subgroup of Eastern Sudanic, that this group emerges just from the structural data is quite good. Furthermore, we saw in the SplitsTree4 calculations that Tama patterned with these languages—the Taman languages are the other part of ‘Northern Eastern Sudanic’.

---

<sup>9</sup> To reduce some of the noise, I only consider languages that have values for at least 25% of the features (i.e., at least 12 out of 47 features). In the case of Eastern Sudanic, this reduces from 47 languages to 32.

- Nara, Nubian, Tama and Nyimang are related to Majang (75% of trees). This is probably false: Majang is part of the Surmic language family—though it is atypical for that family and usually placed as the sole representative of ‘North Surmic’. The Surmic languages, like the Nilotic ones, are placed by Bender in the ‘Southern Eastern Sudanic’ group, whereas Nara, Nubian, Tama, and Nyimang are part of the ‘Northern Eastern Sudanic’ languages.
- Chai and Meen are closely related (55% of trees). These are both part of ‘Southern Surmic’, and beyond that, classified as ‘Southeast Surmic’.

### 3.2. *Shabo as a ‘Northern’ Eastern Sudanic language*

When looking at language groups, we are interested in shared innovations - what evidence is there that links Shabo to the Northern Eastern Sudanic languages and separate from the other Eastern Sudanic languages? The main evidence is:

- No inclusive/exclusive distinctions in verbs or pronouns (at least not in Dongolese Nubian and Shabo)
- Similar word order characteristic: OV, postpositions, Genitive-Noun<sup>10</sup>

But there are also a number of differences - the most striking are:

- Gender distinctions in Shabo pronouns<sup>11</sup>
- Differences in pluralization strategies (recall that Shabo has a separate word for pluralization; Nara, Dongolese Nubian, and Tama use a plural suffix; Nyimang does not mark plurals)
- Shabo’s tense/aspect markers come before and after the verb stem, while these other languages all use suffixes

And of course, we would ideally like to evaluate non-linguistic evidence. In this case, we actually have long records of the Nubians who have been doing hoe agriculture and trade for a long time (Dimmendaal 2007 argues that pastoralism was introduced to these people as early as 5,000 B.C.). This is also true of the other ‘Northern Eastern Sudanic’ languages like Tama and Nara (Bender 1996: 212). Recall that the Shabo have been hunters and gatherers until recently. It is possible that the original Shabo moved into their current location and gave up agriculture, and this fits their mythology, but it is not really clear why they would

---

<sup>10</sup> While both Shabo and Dongolese Nubian have the word order ‘Demonstrative Noun,’ Nara and Nyimang have ‘Noun Demonstrative;’ interestingly, Shabo has ‘Adjective Noun’ while all the other languages (including Dongolese Nubian) have ‘Noun Adjective.’

<sup>11</sup> Shabo’s pronoun system is strikingly different than almost every other African pronoun system.

have needed to give up agriculture unless the climate was quite different from today.<sup>12</sup>

### 3.3. *Shabo and Central Sudanic*

What if Shabo were a Central Sudanic language? While researchers seem to agree *what* the Central Sudanic languages are, the internal classification of the group is still open for debate and considered rather messy.

SplitsTree4 puts Shabo closest to Lendu and Ngiti, but as in Figure 1, Shabo is the language that is farthest away from all the other languages. Other Central Sudanic groups like Bongo-Bagirmi do seem to emerge in the network.<sup>13</sup>

Bayesian methods are unwilling to posit much internal structure with Central Sudanic, only pairing Lendu and Ngiti (which, indeed, should be paired). A closer look at features shows that Shabo does share some word order characteristics with Lendu and Ngiti, but is different in almost every other way. The possibility of Shabo being a Central Sudanic language seems vanishingly small.

### 3.4. *Shabo and the non-Central Sudanic, non-Eastern Sudanic*

Having compared Shabo to the largest subgroups of Nilo-Saharan, we are still left with 18 other languages that could have possible affinities. These languages are a bit of a grab-bag - Bender places various sets of them closer or farther away from Eastern Sudanic and Central Sudanic. Here I have united them under an umbrella that is essentially, ‘Not Eastern Sudanic or Central Sudanic’.<sup>14</sup>

SplitsTree4 places Shabo closest to Fur. It is farthest from Komo, which is interesting since previous researchers have guessed that there was a Koman-Shabo affinity. This is consistent with MrBayes, which also places Fur and Shabo together, although only 53% of the optimal trees contain such a relationship. These languages are over a thousand miles apart, but they share these characteristics:<sup>15</sup>

---

<sup>12</sup> Actually Dimmendaal (2008: 851) suspects that the Ethiopian Highlands where Shabo was spoken were an ancient refugium or retreat zone that was less affected than other areas undergoing desertification.

<sup>13</sup> To see the networks and read more about what makes Shabo an un-Central Sudanic language, see

[http://www.stanford.edu/~tylers/notes/qp/Unclassifying\\_Shabo\\_QP\\_Schnoebelen\\_Jun-3-09.pdf](http://www.stanford.edu/~tylers/notes/qp/Unclassifying_Shabo_QP_Schnoebelen_Jun-3-09.pdf).

<sup>14</sup> I have tried out Bender’s divisions of Core/Satellite/Peripheral and found largely the same trends for the placement of Shabo.

<sup>15</sup> In earlier sections, I focused on ‘shared innovations’. Because this group is really defined negatively, it does not make quite as much sense to see how Fur and Shabo are alike separate from the other languages. Instead, I report only on similarities and differences.



- Cases are suffixed<sup>16</sup>
- Both Fur and Shabo are consistently dependent-marking (possessor is marked, not the possessed; the direct object itself is marked, not the verb)
- No inclusive/exclusive distinction in verbal inflection or pronouns
- Word order: Dem-N, postpositions, SOV, Gen-N<sup>17</sup>
- No pronominal possessive affixes (e.g., dog-2.sg = ‘your dog’)
- A negative morpheme of the form *-bV*

The major differences are:

- Fur has a number of pluralization strategies, while Shabo has a plural word
- Fur has no gender distinctions in pronouns at all<sup>18</sup>
- Fur has neither ejectives nor implosives; the ejectives in Shabo are likely to be areal, but that is less likely for the implosives
- Fur has verbal suppletion according to tense (cf. *go/went*); Shabo shows no evidence of suppletion by tense or aspect
- Fur has tense/aspect suffixes, while Shabo marks tense/aspect before and after the verb stem

#### 4. CONCLUSION

By this point, we have given the Nilo-Saharan hypothesis every opportunity to succeed, but none of the data is convincing: (1) The morphology of Shabo is very un-Nilo-Saharan, especially once we correct the data that has been used in the past. This is true for all areas of morphology, especially case and verbal inflection, (2) Shabo does not clearly participate in any the 80 innovations that Bender describes, nor his Nilo-Saharan-wide grammemes, (3) lexical support for the Nilo-Saharan hypothesis has been overstated.

The best hope for a Nilo-Saharan connection is probably in Fur, although even this is held together primarily by word order characteristics and the negative morpheme. Ultimately, it is possible that the negative morpheme is merely a coincidence and that word order is, as well. Indeed, it is possible that Shabo’s word order could have been influenced by Omotic languages that are nearby—all of the Nilo-Saharan languages with similar word order to Shabo are a thousand

---

<sup>16</sup> Jakobi (1989) actually shows that Fur only has peripheral cases (genitive and locative), and none of the core cases (like accusative). None of these markers look similar to Shabo’s markers.

<sup>17</sup> Fur is Noun-Adj and N-Num, while Shabo is Adj-N and Num-N.

<sup>18</sup> Shabo not only has gender distinctions in 1/2/3 and in SG/PL, as has been noted before, but I have discovered that they also mark dual for 1/2/3 persons—and there is a separate pronoun for male and female to completely fill out the paradigm.

miles from it, but the Omotic languages that share word order characteristics are next door.<sup>19</sup>

To classify Shabo alongside other languages will require an explanation for the uniqueness of Shabo's pronoun system, cases, and verbal morphology. Most likely this will require the use of the comparative method since archeology and genetic profiling are unlikely to be available. Application of the comparative method would be a significant advance, not just for classifying Shabo, but for understanding the pre-history of Nilo-Saharan and other Central/East African peoples.

Dimmendaal is pessimistic about ever classifying Shabo, due to the historical layering of inherited and borrowed material (Dimmendaal 2008: 844, etc). While I have demonstrated that it is best to call Shabo a language isolate, I hope I have also shone a light towards new methods that may be used to approach classification questions. More methodological fine-tuning is necessary in using biological tools and concepts, but the techniques are promising. Used along with further fieldwork, I am hopeful that we will move closer to uncovering an unwritten past and preserving an endangered language.

#### REFERENCES

- Bender, M. L. (2000). Nilo-Saharan. In B. Heine, & D. Nurse, *African languages, an introduction* ( 43–73). Cambridge, UK: Cambridge University Press.
- Bender, M. L. (1983). Remnant Languages of Ethiopia and Sudan. (H. Marcus, & G. Hudson, Eds.) *Nilo-Saharan Language Studies* , 336-354.
- Bender, M. L. (1991). Subclassification of Nilo-Saharan. In M. L. Bender (Ed.), *Proceedings of the Fourth Nilo-Saharan Conference, Bayreuth 1989* ( 1-36). Hamburg: Helmut Buske.
- Bender, M. L. (1996). *The Nilo-Saharan languages: a comparative essay*. München: LINCOM Europa.
- Bender, M. L. (1977). The Surma Language Group: A Preliminary Report. *Studies in African Linguistics* 7 .
- Bryant, D., & Moulton, V. (2004). NeighborNet: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* , 21, 255-265.
- Dimmendaal, G. (2007). Eastern Sudanic and the Wadi Howar and Wadi el Milk diaspora. *Sprache und Geschichte in Africa* , 18, 37-67.
- Dimmendaal, G. (2008). Language Ecology and Linguistic Diversity on the African Continent. *Language and Linguistics Compass* , 2 (5), 840-858.

---

<sup>19</sup> For an analysis of word order and language contact in Ethiopia, see Tosco (2000). The question to ask, of course is whether these Omotic languages could be genetically related to Shabo. That is beyond the scope of this paper, though previous researchers have failed to find sufficient cognates in the lexical data and I only find a handful ('dog', for example). Still, a harder look at the Omotic languages is clearly the next step.

- Ehret, C. (2001). *A historical-comparative reconstruction of Nilo-Saharan*. Cologne: Rüdiger Köppe.
- Ehret, C. (1995). Do Krongo and Shabo being in Nilo-Saharan? In R. Nicolai, & F. Rottland (Ed.), *Fifth Nilo-Saharan Linguistics Colloquium, Nice, 24-29 août 1992* (389-402). Hamburg: Helmut Buske.
- Fleming, H. (1991). Shabo: presentation of data and preliminary classification. In M. L. Bender (Ed.), *Proceedings of the Fourth Nilo-Saharan Conference Bayreuth, Aug. 30*. Hamburg: Helmut Buske.
- Haspelmath, M., Dryer, M., Gil, D., & Comrie, B. (2005). *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- Huelsenbeck, J., & Ronquist, F. (2001). Bayesian inference of phylogenetic trees. *Bioinformatics*, 17 (8), 754-755.
- Huson, D. (1998). SplitsTree: A program for analyzing and visualizing evolutionary data. *Bioinformatics*, 14 (10), 68-73.
- Huson, D., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23 (2), 254-267.
- Jakobi, A. (1989). *A Fur grammar: phonology, morphophonology, and morphology*. Hamburg: Helmut Buske.
- Jordan, L., Mohammed, H., & Netzley, J. (2007). *Sociolinguistic Survey of Shabo*. Manuscript.
- Nichols, J., & Warnow, T. (2008). Tutorial on Computational Linguistic Phylogeny. *Language and Linguistic Compass*, 2 (5), 760-820.
- Sands, B. (2009). Africa's Linguistic Diversity. *Language and Linguistic Compass*, 3 (2), 559-580.
- Saunders, A. (2005). *Linguistic phylogenetics of the Austronesian family: a performance review of methods adapted from biology*. Retrieved from [http://arpiarsaunders.com/documents\\_download/Saunders\\_AustroThesis2005.pdf](http://arpiarsaunders.com/documents_download/Saunders_AustroThesis2005.pdf)
- Teferra, A. (1991). A sketch of Shabo grammar. In M. L. Bender (Ed.), *Proceedings of the Fourth Nilo-Saharan Conference Bayreuth, Aug. 30. Köln, Köppe Verlag*. Hamburg: Helmut Buske.
- Teferra, A. (1995). Brief phonology of Shabo (Mekeyir). In R. Nicolai, & F. Rottland (Ed.), *Fifth Nilo-Saharan Linguistics Colloquium. Nice, 24-29 août 1992*, (169-193).
- Teferra, A., & Unseth, P. (1989). Toward the classification of Shabo (Mikeyir). In M. L. Bender (Ed.), *Topics in Nilo-Saharan linguistics. Nilo-Saharan 3*, 405-418. Hamburg: Helmut Buske.
- Tosco, M. (2000). Is There an 'Ethiopian Language Area'? *Anthropological Linguistics*,
- Wichmann, S., & Holman, E. (n.d.). *Assessing temporal stability for linguistic typological features*. Prepublication version: <http://email.eva.mpg.de/~wichmann/WichmannHolmanIniSubmit.pdf>.
- Wichmann, S., & Saunders, A. (2007). How to use typological databases in historical linguistic research. *Diachronica*, 24 (2), 373-404.