Analysis of D’Avezac’s Sereer wordlist  
John Merrill, June 2018

The earliest written record of the Sereer languages is in the form of a ~1000 word list from the latter half of the 17th century. This list, while lacking in linguistic rigor, is notable for providing information on an earlier stage of the language than any other source. The variety recorded is very similar to the modern Nominka dialect, but contains some notable archaisms especially with regard to the noun class system. In the associated spreadsheet, I have done my best to analyze this wordlist. This document serves both to explain the organization of the spreadsheet, as well as to discuss the linguistic features exhibited in the wordlist.

Nature of the document:

The wordlist as we find it was edited and published by Armand D’Avezac in the Mémoires de la Société ethnologique of Paris in 1845. D’Avezac gives a two-page introduction to the wordlist explaining its provenance and presentation. He was made aware of a forgotten tome containing these wordlists in the French Royal Library. It was written by a member (or members) of the French Compagnie Royale du Sénégal during one of the early French expeditions to the area. Gamble (1992) speculates that the list dates from circa 1670. The list is anonymous, but Segerer and Flavier’s Reflex database speculates that the author may be someone by the name of La Courbe. The notebook contains wordlists for Wolof, Mandinka, Fula, Sereer, Soninke, Bainunk (likely Gujaaxer), and Joola Felup (aka Ejamat). There were apparently also lists for Papel, Bijogo, Nalu, and “Sapi” (perhaps Saafi?) which were lost. The lists were originally given one language after another, but D’Avezac has arranged them in columns so they can be easily compared. The list of French words differs slightly from one language to another (e.g. livre for some languages, but lièvre for others), which seems to have caused confusion in some cases. D’Avezac notes that the handwriting is rather difficult to decipher, but he has done his best to interpret it. This has indeed resulted in a number of mistranscriptions, often <z> for <r>, <d> for <g>, and <s> for <a>. D’Avezac has also modernized the spelling of the French words, though he may have in some cases identified a word incorrectly. D’Avezac himself was not a linguist, and did not have any knowledge of Sereer. I do not believe that the original manuscript is available today. There is no way to be sure of whether the wordlists were all collected by the same person, but I suspect that they were. The idiosyncrasies in spelling (e.g. <qu> and <gu> for palatal stops) are consistent between languages, and taken together are unlikely to represent the work of multiple authors.

Miscommunication in the original elicitation environment has resulted in many sometimes comical mistranslations from French to Sereer. It is clear that the language of elicitation was French, whether directly or more likely through a Portuguese interpreter. Often, the French word is confused for a homophonous or near-homophonous word, for example:

<table>
<thead>
<tr>
<th>elicited:</th>
<th>interpreted as:</th>
</tr>
</thead>
<tbody>
<tr>
<td>infâme</td>
<td>une femme</td>
</tr>
<tr>
<td>une fois</td>
<td>un foie</td>
</tr>
<tr>
<td>un ver</td>
<td>un verre</td>
</tr>
<tr>
<td>ensemble</td>
<td>semble</td>
</tr>
<tr>
<td>terre</td>
<td>taire</td>
</tr>
<tr>
<td>marché</td>
<td>marcher</td>
</tr>
<tr>
<td>plus</td>
<td>pluie/plut</td>
</tr>
<tr>
<td>voix</td>
<td>vois</td>
</tr>
</tbody>
</table>
More abstract concepts were generally not conveyed particularly accurately, often resulting in only tangentially related translations:

<table>
<thead>
<tr>
<th>French</th>
<th>Sereer</th>
<th>Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>amaigrir</td>
<td>a neera</td>
<td>‘he’s seasick’</td>
</tr>
<tr>
<td>piller</td>
<td>a bufa</td>
<td>‘he gathers’</td>
</tr>
<tr>
<td>fou, fol</td>
<td>o laaw</td>
<td>‘soul’</td>
</tr>
<tr>
<td>auprès</td>
<td>xiceen</td>
<td>‘pull it!’</td>
</tr>
</tbody>
</table>

In some cases, we can imagine some failed attempt at pantomime resulting in the miscommunication.

<table>
<thead>
<tr>
<th>French</th>
<th>Sereer</th>
<th>Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>collier</td>
<td>pambe ‘goats’</td>
<td>leading an animal by the collar</td>
</tr>
<tr>
<td>nulle</td>
<td>reti maana ‘go over there!’</td>
<td>pointing about with some attempt at negation</td>
</tr>
<tr>
<td>loi</td>
<td>ñoxor ‘fight’</td>
<td>enacting punishment, or pounding a gavel</td>
</tr>
<tr>
<td>croître</td>
<td>a nita ‘he measures’</td>
<td>moving two hands apart in increments</td>
</tr>
</tbody>
</table>

In yet other cases, while the Sereer is rather clear, it has no relation at all to the French, seemingly the result of a complete miscommunication. Unfortunately there are many cases where I have been unable to even guess at what the Sereer might be. While some of these may represent words I am unfamiliar with, I suspect most are the result of miscommunication and imprecise transcription.

Nouns are generally given as the bare noun, but verbs almost always appear with some additional morphology. Verbs are most often given in the singular imperative with -i (-een with a 3rd sg. object pronoun) or the 3rd person singular default form with a ... -a.

**Spelling conventions:**
The transcription is very inconsistent, and underrepresents a number of contrasts.

- /x/ is sometimes distinguished as <kh, rh> but is often spelled the same as /k/ or /h/.
- <h> can represent /h/, /x/, or Ø.
- Palatal stops are relatively consistently written as <qu>, <gu> for voiceless and voiced respectively. These spellings are also often used for /k/ and /g/. Sometimes /c/ is written as <ky, ki>.
- /j/ is spelled in some words as if it were /j/, and in others as /y/.
- /ñ/ is written as <ny, ni, gn>.
- /ng, nq/ are not distinguished.
- Voiced stops are sometimes written as voiceless, and less commonly vice versa.
- Vowel length is not distinguished.
- Post-vocalic <s> is often silent, e.g. <fasque> for faak ‘yesterday,’ perhaps meant to indicate a long vowel. But this silent <s> is also found with short vowels.
- The list is very inaccurate in distinguishing /u/ from /o/ and /e/ from /i/, especially in prefixes and suffixes. Common spellings for each are:

```
/u/       /o/       /e/       /i/
<ou, u, o> <ou, o>   <ay, ai, e, é> <i, y, ay, é, e>
```
Phonology:
Phonologically, this variety does not seem to differ from the modern Ñominka dialect.
• Very likely that it does not have voiceless implosives (just like modern Ñominka).
  Implosives are with few exceptions transcribed as voiced stops, even in final position.
• Likely that uvular /x/ and glottal /h/ are distinguished.
  /h/ always spelled as <h>, whereas <x> can be <c, kh, k, rh, h>
  This assumes xum ‘tie’ with /x/ (Crêtois’s dictionary gives hum, but I’m suspicious)
• At least some modern Ñominka does not distinguish /ng, nq/ (velar and uvular voiced prenasalized stops). There is probably no way of knowing whether this was the case here. If they were distinguished, the author would have almost certainly spelled them the same.
• Seems that grade III mutation before /ʔ/ yields prothetic /n/, as in some modern dialects.
  <vuimon enda> = wiin a nanda ‘people know’ (root ‘and’)
• Other dialects’ intervocalic /g/ is /h/ in two words: <guiohe> = njogoy ‘lion,’ <goubayhaha> = begax ‘door.’ This is a feature of some modern Ñominka dialects. Other words are not affected, e.g. <aguiga> = a jega ‘he has,’ mod. Ñom. a jeha.

Morphology: noun class:
The following class prefixes are represented:

    = modern
  o-II   o-II oxe
  Ø-I    Ø-I oxe
  go-I   (g)o-I ole
  xa-    xa-II axe   (the only two examples are with immutable consonants)
  ga-II  (g)a-II ale
  ga-III (g)a-III ale
  a-II   a-II ake
  gu-I   (gi)-I le
  Ø-I    (gi)-I le, Ø-I fe
  gu-III (gi)-III ne
  i-III  (gi)-III ne
  u-III  (gi)-III ne
  o-III  o-III onqe
  fo-I   (f)o-I ole
  fa-III fa-III fe

Of the modern classes, only the diminutive plural and augmentatives are not represented.
  gu- is common on many nouns which are found in the modern ne and le classes. u- is common on ne nouns, and there are three examples of i- on ne nouns. About a third of modern ne nouns in the list have gu-, a third have u- or i-, and a third have no prefix. For le, just over half have gu-, and the rest are unprefixed. We can be rather certain that u- and i- are indeed prefixes on some ne nouns, and not simply inserted orthographically before a prenasalized stop. For one, these are found in ‘breast’ and ‘flatfish’ which do not contain prenasalized stops, and furthermore the author writes Wolof and Fula initial prenasalized stops without any prothetic vowel. These prefixes are all probably archaisms (c.f. Fula ngu, ndi, ndu classes), having been replaced by the innovative gi- in modern Ñominka, and lost in other dialects.
As in modern Ñominka (and no other dialect), the nominal prefixes of the ole and ale classes are ga- and go-. There are a few cases where the initial /g/ is omitted. Two of these are cited twice, once with <g> and once without:

<ollau> = <goulaou> = o laaw ‘soul’
<oussako> = o saaku ‘bag’
<onio> = <go nio> = o ñoow ‘life’
<assomman> = a sumaan ‘heat’ (could be ake)
<atodoque> = a toodook ‘rafter’ (could be plural)

Notably there is no /g/ in the one example of no or fo before go- or ga-:

<...nomague> = no maag ‘in the sea’

There are some cases where the class prefix is omitted entirely. Sereer speakers (regardless of literacy) usually think of the prefix as a separate word and sometimes omit it for words given in citation/isolation. This omission is not very common in the list, and it is unlikely that it can account for the sizeable number of unprefixed nouns in the modern ne and le classes.

There are a few examples of noun-adjective phrases, and these show the expected noun class agreement in all but one case (‘sour milk’), which is perhaps attributable to haplology.

<groulgourer> = gu ruul gu rew ‘female pig’
<fosso phessis> = fo soow fo sis ‘fresh milk’
<fosso phodako> = fo soow foðaxu ‘sour milk’

Determiners are not common in the list, but we find oxe, and perhaps ne and ke, identical to their modern forms. Most importantly, there is an instance of re in <nakre> = naak re ‘the cow.’ This is a conservative form of this determiner, preserved in the modern Njagañaaw dialect. Other dialects have l- rather than r-.

Verbal morphology:

Negative -eer seems to always retain its final consonant, whereas it is usually dropped word-finally in the modern language. Otherwise there do not seem to be any differences from the modern language.

Lexicon:

While lexical usage does not vary greatly between dialects, the words found in the list are consistent with the Ñominka dialect, and possibly also Siin, but certainly not Saalum. I only have firsthand experience with the Saalum dialect, so I can’t be sure whether the vocabulary is without a doubt Nominka or if it could be Siin. Some notable words which are not found in Saalum:

Ñominka/Siin: Saalum:
suum yuum ‘honey’
-eene -een ‘this’
hakandoox (rare) ‘lay down head’
A common word found only in Ñominka (found in Renaudier 2012, but not the Siin sources I know) is lemb ‘say/call,’ given four times in the list. Two other words of potential importance are <gniab> ‘bring’ given three times, and <lalgand> ‘show, etc.’ given five times. Renaudier has ñis ‘bring’ and lal ‘show’ just as in Siin and Saalum.

There are a good number of Wolof borrowings in the list— seemingly about the same as in the modern language. However, a number of Wolof words are given that to my knowledge are not employed in modern Sereer (I have grouped these together before the Portuguese borrowings at the end of the list). Perhaps these are characteristic of the Ñominka dialect. Regardless, this list dispels any possible notion that the large number of Wolof borrowings in modern Sereer is a recent phenomenon.

There is perhaps only one French borrowing in the list, for ‘amber,’ which is natural since the French did not establish a significant presence in Senegal until much later. There are however a good number of Portuguese borrowings, and only a few of these are still in use today. At the time, the Portuguese were the only European nation with a notable permanent presence in the area. It is quite likely that the consultant for this list spoke Portuguese or at least was more familiar with the Portuguese colonists than most Sereer, as such a person would naturally make for a more attractive consultant for the French author. The fact that the consultant gives <santo> for ‘image’ suggests that he was exposed to Portuguese missionary efforts, and may have even been a Christian convert (to this day, the Sereer are the only ethnic group with a sizeable Christian population in Senegal north of the Gambia). Thus it is impossible to say whether the degree of Portuguese borrowing seen in the list would have been characteristic of the language as a whole at the time.

**Organization of the spreadsheet:**

Each entry is given as it appears in D’Avezac’s wordlist, with the French in column B and the Sereer in column C. The modern Sereer form is given in column D, followed by an English gloss of this modern form. Note that the English gloss is not a translation of the French prompt, but of the Sereer form, which is often quite different. The modern form given is from the Saalum dialect, since that is dialect for which I have the most reliable data. Of course these are not the forms that are directly represented by the wordlist, since they come from a different time and dialect. The most notable differences between the Saalum forms in column D and what the original author would have heard are:

(i) Different noun class prefixes, as shown in the table in the section on noun class.
(ii) No voiceless implosives.
(iii) /x/ and /h/ distinguished, and perhaps /ng/ and /nq/ not distinguished. Words with /h/ in Siin/Ñominka are noted as such in column H.

Column F gives the part of speech of the form in column D, or its noun class for nouns. Column G is used for certain entries to show the modern Sereer equivalent of the French prompt, simply in order to demonstrate that the Sereer form given could not be related to it (usually in the absence of any other insight). Column H gives various notes, often speculating on mistranscriptions, miscommunications, or possible interpretations of unclear Sereer forms. The entries are arranged so that repeated words and related Sereer roots are grouped together for comparison. Column A allows this sorting order to be easily reverted to. This is not the order that they appear in D’Avezac.