A Diachronic Comparison of the Vai Script of Liberia (1834–2005)

Olena Tykhostup¹ and Piers Kelly²

¹ Friedrich Schiller University, Jena, DE
² The Mint research group, Max Planck Institute for the Science of Human History, Jena, DE

Corresponding author: Olena Tykhostup (tykhostup@gmail.com)

This dataset provides the first comprehensive diachronic comparison of the Vai script of Liberia, as derived from sixteen sources dated between 1834 and 2005. The Vai syllabary was invented by non-literate speakers of a Mande language and is of interest to scholars in the fields of writing and cultural transmission as an emergent writing system. Script samples that entered the dataset were retrieved from a wide variety of published manuscripts. The compiled dataset tracks the evolution of the Vai script via archival records and has reuse potential across various fields of research.

Keywords: Vai script; writing systems

Context
The Vai writing system is a syllabary representing syllables and morphemes of Vai, a Mande language spoken by approximately 167,000 people in Liberia and neighbouring Sierra Leone [1]. What makes the Vai script especially interesting for scholars of both writing and cultural transmission is that it was invented by non-literate and has been continuously transmitted to the present day. It is unknown how many people are presently literate in the script, but for the period 1973–1978, Scribner and Cole [2] estimated that 20.3 percent of the adult male population in their fieldsite could read and write in Vai. In its present standard form, the Vai script comprises 205 individual graphemes.¹

Overview
Created by between six and eight non-literate Vai-speakers in Liberia in about 1833, the Vai script represents the world’s best-documented emergent writing system. The surprising success of the script preceded, and to some extent inspired, the creation of multiple new indigenous writing systems across the West African region [3, 4, 5, 6, 7, 8] a phenomenon that continues into the twenty-first century [9].

Beyond Africa, scholars have long been intrigued by the Vai script, the circumstances of its invention, and its potential to illustrate cultural processes [notably 10, 11]. Later thinkers speculated that the evolutionary trajectory of the Vai script must have recapitulated the evolution of writing itself (for a summary see Kelly [12]).

A few thinkers have sought to substantiate these ideas by comparing the Vai syllabary across different time periods [13, 14, 15, 3]. However, their comparison charts selected only a handful of historical sources, or they compared a small subset of the syllabary. Svend Holsoe, late professor emeritus of anthropology at the University of Delaware and scholar of Liberia, began compiling a more expansive comparative chart but was unable to complete the work before he passed away [Charles Riley, pers. comm.]. To the best of our knowledge the present dataset includes every dateable manuscript source that is currently available in public archives.

Methods
One of the challenges in compiling the dataset has been in the interpretation of early sources produced by non-Vai visitors. These men did not always understand the phonology of the Vai language and thus assigned graphemes to incorrect syllables. Specific instances of this are noted in the data description below. As for the form of the graphemes, both P.E.H. Hair (cited in [3]) and Gail Stewart [16] expressed scepticism that the early sources produced by Europeans were faithful to the Vai script as it was used at the time. However, our chart shows that syllabaries elicited by outsiders are remarkably congruent with original text sources written by the Vai scribes themselves in the same time periods (see commentary below).

Our dataset was compiled by sourcing unedited digital images of Vai manuscripts, identifying individual graphemes and placing them in a table. The table contains the following fields to facilitate searching: manuscript name and date, unicode number, sound value (transliteration into the IPA), unicode transliteration...
Dataset Description

Script samples that are presented in the table were retrieved from a wide variety of published manuscripts. All sources we have included are reliably dated, except for the Ndole manuscript from the Houghton Library, Harvard: the archival metadata gives an estimate of ca. 1845. We exclude sources of ambiguous origin as well as those copied or adapted from other existing sources. Very short texts are likewise excluded. For instance, the final dataset does not feature the seven-character house inscription copied by Edwin Norris in Cape Mount in 1848 and the accompanying ‘Specimen of Ms.’ [18], nor the brief letter of 26 graphemes collected by Oscar Baumann [19]. We have also excluded the original texts compiled in Ellis [20] since they cannot be reliably dated.

The manuscripts

In this section, we describe the circumstances in which the collected script samples were first documented. This is an essential step aimed at enabling researchers to assess the relevance of the data while addressing their own specific questions.

Gail Stewart has argued that some of the characters from early European sources were corrupted in the process of their documentation. She writes:

The appearance of the early Vai script was familiar to the Liberian Vais who taught me the modern script in the 1950s, but they found it largely illegible and even rather humorous. [...] What amused the modern Vais was in reality a European interpretation of their script: that is, handwriting by Vais had been redrawn by Europeans so that it could appear in print. When the original manuscript of the ‘Book of Rora’ was turned up in 1967 in the Houghton Library of Harvard University, and, in the same year, the two-page Forbes manuscript, also in pre-1850 Vai, was ‘discovered’ in the British Museum, it became obvious that the foreign copyists, with all good intentions, had stylized and distorted the early Vai script to the point of absurdity, and sometimes beyond recognition.

Now we know that the difference between the old script and the modern is not as great as was supposed. [16, p. 1]

The two ‘corrupted’ texts that Stewart specifically mentions are Rora [21] and Koelle [22]. Neither are included in the table since they do not represent instances of primary script documentation. However, our dataset clearly shows that early documentation on the part of Europeans, including Koelle [22], is in fact highly consistent with samples provided by Vai scribes themselves within the same period. The difficulty that Stewart’s colleagues had in interpreting examples of the script is more likely due to the higher degree of synchronic variation witnessed in the 1850s; many allographic variants were to be eliminated, and no doubt forgotten, by the 1950s.

For a fuller historical context of the surviving nineteenth-century Vai manuscripts see Tuchscherer and Hair’s exhaustive commentary [23].


The earliest known sample of the Vai script is conserved at the Houghton Library at Harvard University. It was written by Fan Dawo Kelondo at the request of John Leighton Wilson. The text is only a page long and has not yet been translated, but due its historical importance our table has excerpted every single character. Thus there are thirteen instantiations of ꟊ, ꟍ, while other relatively common characters (such as ꟊ ha, ꟍ ma, and ꟍ na) are not recorded at all. This maximal extraction from the source allows researchers the fullest possible scope for cross-character comparison in subsequent sources.


Stewart’s paper is an analysis of an original handwritten copy of the Book of Ndole, also known as the Book of Rora, authored by Kaali Bala Ndole Wano. Here she helpfully extracts a full syllabary from this original text, which she dates as 1850 or earlier; we are using the Houghton Library date of ca. 1845.


This syllabary was collected by Lieutenant F. E. Forbes in 1849 and communicated in a dispatch to the Admiralty on 23 April of that year. The dispatch was later published in the Journal of the Royal Geographical Society of London. Forbes reports:

A lucky chance took me to a town called “Bohmar,” about 8 miles E. of Cape Mount, and there I met a...
Koelle, S W (1849). Narrative of an expedition into the Vy country of West Africa and the discovery of a system of syllabic writing recently invented by the natives of the Vy tribe. London: Seeleys, Fleet Street; Hatchards, Picadilly; J. Nisbet and Co. Berners Street.

This complete syllabary was produced in a village in Cape Mount by both Mamulu Duwalu Bukēle, one of the principle inventors of the script, and by the German missionary S. W. Koelle, who wrote: 'my landlord [Bukēle] began to copy his book. I, however, had to finish it, and Doalu Bukara [Bukēle] afterwards said to me, 'White people can write better than black people; you must copy my book for me.' I gladly accepted the offer [...] [25, p. 19]. Thus, the Vai characters in this source were produced under the supervision of a Vai scribe even if they were not all written by Bukēle himself.

Koelle's account of the Vai script features occasional inaccuracies when it comes to assigning syllabic values to graphemes. Thus, the graphical signs for ɗ | syllable series are consistently misinterpreted as d| sound values, i.e. ɗ | ɗa is documented by Koelle as ɗa, and • | di entered his report as dī. Similarly, unicode Ɗ | kpa and Ɗ | kpã are both transcribed by Koelle as gba, and unicode ꕠ | ga as ƙa. The adjudication of such conflicting cases in are dataset was based on their graphical similarity to more reliable sources (e.g. the Ndole syllabary reconstructed in Stewart [16]; see above).


A short note of 102 graphemes reproduced by J. S. Payne in the newsletter of the Protestant Episcopal Church. Payne writes in his letter of 6 June 1860: 'Inclosed I send you a note received by [the missionary] Mr. [A. D.] Williams from a chief in the interior, which was translated for him by a Vey youth living near Mr. Williams' residence.' [26, p. 369]. The translation is not provided in the text.


The circumstances and location in which Creswick collected his published syllabary are, unfortunately, quite unclear. Creswick wrote: 'My remarks are confined to a mere recital of what I saw and heard whilst living amongst the Vey Negroes. [...] Certain it is that old men now living, with whom I have conversed, remember the time of its creation; and it is from information received in this manner that I learnt the origin of this, the only native African orthography in existence.' [27, pp. 260–261]. However, of interest is the fact that Creswick documents the re-emergence of Vai schools: 'There are, moreover, schools where the children are regularly taught [the Vai script] by means of small black boards and chalk' [27, p. 261]. This suggests a possible pressure for standardisation of the script at this time. Koelle and Migeod report that Vai was taught in purpose built schools from ca. 1835 but that these were destroyed in war eighteen months later [22, 28].


The first to comment on Vai as an evolving system, Delafosse makes important observations about Vai pedagogy and the high degree of individual variation in writing styles at the end of the nineteenth century, especially among competent scribes. He includes a table that compares the script as recorded by Forbes with the contemporary script. Only his 'contemporary script’ is recorded in our dataset; his source is not explicitly stated but it may be Ghaïsama Sando, who provided a sample of Vai writing in the same book.


The chart produced by Momulu Massaquoi [29], a descendant of Bukēle, was published as part of a campaign to both reform and standardise the Vai script. To this end he invented new characters to fill what he took to be gaps, eliminated 'ambiguous' characters and introduced punctuation marks. This chart was to complement his efforts at introducing Vai into a local school (St. John's in Robertsport). Interestingly, despite his reformatory zeal he retained a handful of acceptable allographs, presumably because they were already too well established in the scribal community.


Of this chart Massaquoi writes: 'I reproduce here the phonetic chart of this language compiled by the writer ten years ago, and published by the Board of Missions of the Protestant Episcopal Church of America in their monthly periodical, the Spirit of Missions’ [30]. However, this version is by no means a facsimile reproduction of the 1899 source: despite being penned by the same author there are numerous subtle variations. It includes a few interesting annotations on the suspected iconic origins of certain graphemes. These icons were discussed at some length by Klingenheben [31].


Like Delafosse before him, Johnston was conscious that the Vai script had evolved considerably since its invention, and that the modern forms were 'simplified and better adapted for cursive writing' [14, p. 1115]. To show this he compiled a chart comparing the forms documented by Forbes and Koelle with 'modern types of the letters as accurately as [he] could obtain them from Vai scholars'

Migeod’s consultants for his syllabary are unfortunately not recorded. He does, however, make a significant observation about the variability of the syllabary in the early twentieth century:

With the creation of new characters based on error it can be seen that the compilation of a standard syllabary is a matter of great difficulty. In fact, it can hardly be said that one exists. Accordingly I cannot claim for the accompanying syllabary that it is either complete or exact. I am frequently coming across new signs, without, unfortunately, always having the means of assigning to them a value. [28, p. 51]


Klingenheben’s essay comments on a number of the earlier sources. We have extracted graphemes from the appendix of characters cited in the article as well as the text of Psalm 23, provided by Zuke Kandakai.


Dalby’s Vai chart is a comparison of the Koelle manuscript 1849 with the version provided by Gail Stewart in 1958, as well as the University of Liberia ‘standard’ syllabary of 1962 (see entry below). For our dataset we have extracted the graphemes provided by Stewart which are described as having been derived from a manuscript chart of the characters most widely accepted by modern users of the Vai script, compiled by [Stewart] in 1958 [3, p. 6].


This document represents the first formal attempt to collaborate on a national standardisation project. It involved eleven consultants and a committee of four, headed by Zuke Kandakai [32]. The Standardization Committee also introduced new signs for ‘r’, ‘sh’, ‘sz’ and ‘th’. Contrary to Massaquoi’s 1899 standardisation attempt, no allographs are permitted: there is only one sign per syllable.


Of their reproduced syllabary, Scribner and Cole write: ‘This is a modified version of the modern syllabary prepared at the University of Liberia by a group of indigenous Vai script experts and a foreign linguist’ [2, pp. 313–314]. However, it is clear that the chart is not a direct facsimile reproduction of the 1962 set from the University of Liberia. We assume that the discrepancies (or ‘modifications’) are in line with advice from Scribner and Cole’s informants. Interestingly, certain allographic variants are re-introduced here, suggesting that the University of Liberia standardisation campaign of twenty years earlier was not wholly successful.


The most recent standardisation attempt is witnessed in the Vai unicode proposal of 2005. The Vai script was added to the Unicode Standard in 2008. Everson et al. explain the sources for the unicode version of Vai thus:

The primary sources for the Vai characters in the character set proposed are the 1962 Vai Standard Syllabary (which was a distillation of many sources specifying characters for modern use), modern primers and texts which use the Standard Syllabary (and a few glyph modifications reflecting modern preferences), the 1911 additions of Momolu Massaquoi, and the characters found in The Book of Ndole. Secondary sources, such as Johnston 1906 and Dalby 1967, are used as supplementary material and as checks for some of the archaic characters. [33, p. 2]
Al-Sayyid sign languages [34, 35], and mixed languages like Light Walpiri [36] and Gurindji Kriol [37]. Emergent sign languages have been developed ex nihilo by linguistic communities and are thus independent of any ‘parent’ languages and lineages. Mixed languages are also set apart because they involve a naturalistic re-engineering of existing linguistic structures to generate a new system. Since emergent languages (and to a lesser extent mixed languages) sit outside established language families, studies of these systems have the potential to reveal the spontaneous emergence of structure without the ‘noise’ of inheritance and contact. We contend that the Vai writing system has comparable value in tracing the evolution of graphic codes, a field of study that has so far been limited to laboratory settings.

Lastly, the independence of Vai from known script lineages may also allow help refine research into universals and variation in writing systems, e.g. [38, 39, 40, 41, 42]. We expect, for example, that documented changes to the Vai script will enrich discussions of how graphic codes make the transition from non-linguistic sign systems to full writing.

**Limitations and Provisos**

Although dated manuscripts may yet come to light in informal archives, our table is a compilation of dated Vai sources that have hitherto survived in the documentary record. It does not, and cannot, represent the complete history of the script. The full extent of variation and branching for certain graphemes, or sets of graphemes, may never be known. But while extinct grapheme lineages may be under-represented in our data, we can be confident that the graphemes attested from the 1960s onwards are those that survived selection pressures of earlier generations.

**Note**

1. There are 300 Vai graphemes included in the Vai unicode block, however this block includes obsolete characters.

**Acknowledgements**

Thanks to Charles Riley for advice and for supplying a copy of Stewart (1972), to Mireille Djenno of the Indiana University Liberian Collections (IULC) and to Susan Halpert at the Houghton Library, Harvard.

**Competing Interests**

The authors have no competing interests to declare.

**References**


25. Koelle, SW. Narrative of an expedition into the Vey country of West Africa and the discovery of a system of syllabic writing recently invented by the natives of the Vey tribe. 1849; London: Seeleys, Fleet Street; Hatchards, Piccadilly; J. Nisbet and Co. Berners Street.


33. Everson, M, Riley, C and Rivera, J. Proposal to add the Vai script to the BMP of the UCS. Universal Multiple-Octet Coded Character Set International Organization for Standardization. 2005; ISO/IEC JTC1/SC2/WG2 N2948R: L2/05-159R.


39. Changizi, MA, Zhang, Q, Ye, H and Shinsuke, S. The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. The American Naturalist. 2006; 16(5): E117–E139. DOI: https://doi.org/10.1080/07268600500110456

