

A forgotten figure in Siouan and Caddoan linguistics: Samuel Stehman Haldeman (1812-1880).ⁱ

0. Abstract

In the light of Bob Rankin's Dhegiha work this paper examines some of the earliest recorded material on Kanza and Osage, collected and transcribed by the naturalist Samuel Stehman Haldeman in an alphabet of his own devising (Haldeman 1859, 1860). Although his transcriptions fail to capture many crucial phonetic and phonemic distinctions, they are useful as records of earlier and more conservative forms of these languages.

1 Introduction.

Robert Rankin's examination of earlier sources on Native American languages which have rarely been the subject of fuller description impel us to look at the work of other early collectors of data on Siouan and Caddoan languages. We may mention for instance his paper on Max von Wied's brief vocabulary of Kaw, Kanza or Kansa (Rankin 1994), Nor should we overlook his splendid salvage work on Kanza (the name I will use henceforth in this paper) and Quapaw, and his pivotal role in the organization of the Siouan-Caddoan Conferences.

One researcher is almost overlooked nowadays (despite a *Memoir* by Lesley 1881 which hymns his activities while getting its dedicatee's name wrong). The naturalist, sawmill manager and avocational linguist Samuel Stehman Haldeman (1812-1880) was mostly known to the linguists in the 19th century for his 'Analytic Orthography' (Haldeman 1859, also produced in book form as Haldeman 1860). This was a prizewinning attempt to construct a universal phonetic alphabet, based on Latin letters (and following some precepts of

Classical Ciceronian Latin pronunciation, for instance <C> for /k/ and <V> for /w/) but enhanced with some created symbols. It also added a number of diacritics, for documenting phonetic data in the world's languages, especially from previously undescribed languages of North America and elsewhere.

This work represented a determined effort to describe and notate speech sounds, for which it was awarded the Trevyllian Prize in London against eighteen contenders. And there its reception ground to a halt. The alphabetic system, based on what Haldeman assumed were Classical Latin letter-values, was well-adapted to indicate certain aspects of vowel quality and quantity and basic consonantal distinctions. But his pioneering work is one of several such pre-International Phonetic Association schemes proposed in the 19th century, of which Lepsius (1863) is the most famous and influential, and it is cumbersome. Specifically it is handicapped by the fact that because it required a large number of special fonts and diacritics it was difficult to reproduce, with the result that nobody save Haldeman ever adopted it.

Haldeman's system is elaborate but just how successfully or consistently he applied his own transcriptional system is moot. For instance in his Chinese data (actually Cantonese), he fails to indicate any tones for the numerals for Guangzhou Cantonese, although he makes an effort to do so for the nearby Macanese variety of Cantonese. Haldeman uses the phonetic terminology of his time, with *surd*, *sonant*, *lenis*, *asper*, employed where modern phoneticians talk about *voiceless*, *voiced*, *unaspirated* and *aspirated* sounds, and with *sigmal*, *lingual*, *cerebral*, *guttural* and *faucal* used for modern *alveolar*, *dental*, *retroflex*, *velar* and *uvular* respectively. He also talks about "pure" (non-nasalized) and nasal sounds, and arrays consonants according to their degree of 'interruption' (plosives are the most "interrupted" consonants in this scheme). He also divided consonants into *mutēs*

(plosives and nasals) and *liquids* (other consonants). Haldeman (1860: 83, 369) recognizes thirty four vowel qualities, which he arranges in a dense A-shaped diagram, and he indicates vowel quantity with macrons and breves. Unlike Daniel Jones (Jones 1909) he does not propose a scheme in which the distinction between back [ɑ] and front [a] is crucial.

The concentration in this work is on the Dhegiha-language data, though observations from Haldeman's work on Caddo and Wichita will be added where relevant. (Unfortunately I lack just sufficient modern lexical data to give as full an analysis with modern examples of the Caddo and Wichita data as I would wish.) Data from Haldeman's work are taken from Haldeman (1860), a corrected and book-length edition of Haldeman (1859). Haldeman divided his work into sections (often extremely short and usually corresponding to paragraphs), in addition to the book being paginated. Both modes of reference will be used here.

Figure 1. A digest of consonantal symbols in Haldeman (1860: 121, section 576).

[FIGURE ABOUT HERE]

2 Haldeman the Americanist and his work on Dhegiha.

In addition to a number of versions of the Lord's Prayer (including those in Cherokee and Wyandot) Haldeman (1860) provided data in the form of 75 cardinal numeral sets from 1-10 from a wide range of languages of Europe, Asia and North America (plus Grebo from Liberia). A number of these were Algonquian, Muskogean and Iroquoian languages in addition to numerals from Makah, Chinook, Comanche, Jicarilla Apache and the Yuman language Iipay 'Aa. Among the languages on which Haldeman tried out his spelling system are the Dhegiha Siouan language Kanza (for which he also provided some Santee parallels for certain forms from Riggs 1852). He also provided data from the Caddoan languages Caddo (which Haldeman referred to as 'Nadaco') and Wichita (referred to as 'Waco' but identical with the Wichita recorded in the 20th and 21st centuries, for instance in Rood (1975). In each case the data presented are cardinal numerals from 1-10 and some additional lexicon (over 40 such items in the case of Caddo and 10 from Wichita), evidently recorded by Haldeman from native speakers and not previously listed elsewhere. Haldeman also collected the numerals from 1-10 in Osage.

As Rankin (1994) showed, Max von Wied (Wied 1839-1841) had described sounds in various Dhegiha languages (Wied documented Omaha, Ponca, Kanza and Osage) quite well within the limitations of his annotated Franco-German spelling system. This means that though his work is superb for its time he missed many crucial details and failed to record other details consistently. Haldeman's system was theoretically more precise as far as it went (although there is little consistent coverage of tones and essentially none of consonants which are ingressive, velaric, or other kinds of clicks). But it was deployed less consistently and less accurately. His records of Kanza and Osage do show an ability to indicate primary

stress using acute accents, respectively, while grave accents are used to indicate a variety of vowel qualities, short vowels are marked with breves, long vowels with postvocalic dots, and bowel nasalization is represented with the ogonek or Polish hook placed at the bottom of the line after the vowel.

Working on small amounts of material (often only the numerals from 1-10) from a large number of languages, Haldeman recognized that some sounds were problematic in terms of his descriptive criteria, as his discussion of the two ejectives Caddo /t'/ and Wichita /k'/ shows (Haldeman 1860: 99, section 448 and 131, section 574). But he did not make the leap (as García 1760 had done for Coahuilteco) by discovering that what made these sounds distinctive from other speech sounds but similar to one another was their common possession of ejective quality, with the corollary that ejectives should be represented consistently. As a result he was unable to indicate the ejective quality of the final consonant in Caddo *wists'i* “one”. In fact, Haldeman’s attempts at transcribing Caddo (for instance Haldeman’s <*vàtət*'> for *waadat* “earth”, in which he fails to hear that the medial stop is voiced: Haldeman 1860: 135, section 633) are scattershot enough to be unreliable. Even so he recognized that the Caddo word for “cheek” used a dental or alveolar rather than a velar nasal.

His array of consonantal types was defective in other respects. Although Dhegiha languages contain ejectives, the small samples from Kanza and Osage which Haldeman cites happen not to include any of these sounds. Haldeman would probably have been unable to indicate these, as they are not provided for in his consonantal chart, and his encounters with them in Caddo and Wichita left him uncertain as to the nature and phonetic structure of the ejectives which he encountered there. He also lacked a consistent way of indicating the glottal stop, either initially, medially or finally, which is a special problem when recording Caddo data.

Nor did Haldeman's system capture the three degrees of phonemic vowel length which are present in Wichita (although Haldeman 1860: 80, sections 353-355) provides the wherewithal to do this.

As a result of these and other shortcomings, Haldeman's work has received rather little attention from modern phonologists or indeed other linguists. Even Haldeman himself made no use of the system in his work on Pennsylvania German (Haldeman 1872). The discussion in Pilling (1887) and the brief account in Kelly and Local (1989), written incidentally by the academics who taught phonetics to this author, are rare exceptions to this neglect.

Haldeman's data on Osage, comprising merely the cardinal numerals from 1 through 10, and the corresponding forms in Kanza, help us to get a better sense of his transcriptional techniques. Modern Osage data are from Quintero (2009) and Kanza data from Cumberland and Rankin (eds. 2012). Original transcriptional systems have been preserved. We note that the two languages, though very close, are represented differently in regard to orthographic conventions employed to indicate postalveolar sibilants, vowel length and nasalized vowels.

Figure 2: Numerals in Kanza and Osage from Haldeman (1860: 146).

[FIGURE ABOUT HERE]

Figure 2. Numeral data from Kanza and Osage (Haldeman 1869: 146, sections 711, 712).

Figure 3. Lexicon from Kanza, with some Dakota parallels from Riggs (1852) (Haldeman 1860: 135, section 634).

[FIGURE ABOUT HERE]

Dhegiha languages share a number of crosslinguistically marked features in their segmental phonology. These include the differentiation of nasalised from oral vowels, the differentiation of geminate and lengthened stops, of preaspirated and voiceless ejective stops, and the use of a high front rounded vowel. Modern forms are given below (Kansa <u> is /y/ and superscript <n> represents nasalization, indicated in Osage by an ogonek).

3 Modern counterparts of the data.

Below are given modern equivalents in Kanza and Osage for Haldeman's data.

Cardinal Numerals (Haldeman's sections 711 and 712)

	Kanza	Osage
1	<i>míⁿxcí</i>	<i>wí.xce</i>
2	<i>noⁿbá</i>	<i>đo.o.pa</i>
3	<i>yáblíⁿ</i>	<i>đáabrij</i>
4	<i>dóba</i>	<i>tóopa</i>
5	<i>sátaⁿ</i>	<i>sáhtq</i>
6	<i>shápe</i>	<i>šáhpe</i>
7	<i>péyoⁿba</i>	<i>hpéeđqqa</i>
8	<i>kiadóba</i>	<i>hkietóopa</i>
9	<i>sháⁿka</i>	<i>lébrq hce wíjke</i>
10	<i>glébla</i>	<i>lébrq</i>

Additional Kanza lexicon (Haldeman's section 634):

	Kanza
Ear	<i>naⁿtá</i>
Eye	<i>ishtá</i> (note <i>ishtá toho</i> 'iris' and <i>ishtákaⁿha</i> 'eyelid')
Brow	<i>ishtáhiⁿ</i>
Mouth	<i>i</i> (Haldeman's form <i>iha</i> is 'mouth-skin' or 'lips')
Tongue	<i>léze</i>
Nose	<i>pa</i>
Nostril	<i>pa xlóge</i>
Forehead	<i>pe</i>
Fan	<i>ijéayuzúbe</i> (fan hung over baby's face)
Pipe	<i>naⁿnóⁿba</i>
Knife	<i>máⁿhíⁿ</i>
Warm	<i>moshcé</i>
Leggins [sic]	<i>húyuyiⁿge</i>
Shirt	<i>ókiloxla</i>

Note that what are written as single plosives in the modern Kanza orthography are actually geminates, thus <k> is /kk/.

4 Remarks on the forms.

The materials here represent examples of impressionistic phonetic transcriptions, which is what we would expect in a work from the pre-phonemic era. The Kanza and Osage words in Haldeman's material (especially the former) are recorded with comparatively greater detail

than numerical data from some of the other languages are. Indeed the Kanza numerals are recorded with greater detail by Haldeman with respect to accent than they are in the present orthography. But the forms are not necessarily noted with greater accuracy, and neither system indicates the differences between the various voiceless stop series clearly. Tense stops in Kanza in Haldeman's transcription are represented by the use of bold consonantal characters, so that Haldeman's <p> is [p ~ ph], <**p**> is [pp] while <'p> is [hp] in his Kanza work. (This fact is clearer in the version of Haldeman's work published in the Transactions of the American Philosophical Society than in the acid-heavy and aged brown paper of the version of Haldeman 1860 available from archive.org).

The forms are in general readily identifiable from recording of the languages over a century later, as the references from the Kanza and Osage dictionaries (Cumberland and Rankin eds 2012 for Kanza, and Quintero 2009 for Osage) show. The few differences are instructive.

Most interesting in this regard are the numerals, especially NINE and TEN. In Osage NINE is a subtractive compound ("ten lacking one") involving TEN and an allophone of ONE. But Kanza uses the widespread form, possibly reconstructible as *kišqkka*, which is recorded for several Mississippi Valley and Ohio Valley Siouan, Muskogean and Great Lakes Algonquian languages. Modern Osage has simplified the onset of TEN, though Haldeman had what would nowadays be represented as /kar-/ (or maybe /gar-/; his depiction of voicing is not always trustworthy). The form in Dhegiha has irregular reflexes elsewhere in Dhegiha: Omaha-Ponca *gthéba* (omahalanguage.unl.edu/dictionary/), where <th> is /ð/, has lost the liquid found in the second syllable in other Dhegiha languages and in earlier records of Omaha-Ponca (compare Quapaw *kdébnq*: Rankin 1982: 150). The glide which separate the prefix from the form for THREE in Kanza EIGHT has been apprehended by Haldeman as a

front vowel, although the hiatus in the corresponding Osage form has been recognized by Haldeman as such.

Both the Kanza and Osage forms in Haldeman's work include forms of what was originally the enclitic *-xci* "only" at the end of the form for ONE, and this pan-Dhegiha word is a form which was later borrowed into Caddo as *wists'i'*. Note also the initial [di-] in Kanza THREE, now replaced by /j-/ <y->, and the fact that Haldeman did not notice the nasalization of the vowel in the first syllable of Kanza TWO. The form EIGHT in Haldeman's Osage is reflected in the modern language, in modern Kanza and (as a loan, namely *kiyátaw*: Rood 1996:608) in modern Wichita. But the earlier form for EIGHT based on THREE is used in Haldeman's Kanza as a parallel to the form for SEVEN (itself a compound involving TWO). Primary stress and vowel length and nasalization are well represented in Haldeman's work, especially for Osage.

Of the nouns in Haldeman's record of Kanza, most are similar to their modern counterparts. For the rest, if one allows for a modicum of close phonetic detail (for instance the realization of /a/ as a low rounded vowel in NOSE), the quality of transcription is rather high.

EYEBROW may end in a form of *sábe* "black" but this is uncertain, while material which is less easy to identify is attached to the end of EYE and SHIRT. The occasional weakness in Haldeman's powers of perception is seen in the fact that the consonantal sounds in the second part of NOSTRIL are represented in Haldeman's work by his symbol for /ʃ/, while the initial consonant of TONGUE has changed in the 120 or so years between Haldeman's work and Bob Rankin's. WARM seems to include an enclitic, which may be the masculine declarative enclitic *(ey)ao*. Haldeman's remarks about the phonetics of NOSE and TONGUE are

somewhat surprising, as modern Kanza does not permit [h] in coda position and does not use geminate consonants.

5 Conclusion

Data on Haldeman's recording of Dhegiha language data have been presented and the success or otherwise of Haldeman's system in coping with the segmental phonology of these languages, especially the complex consonantal systems, has been evaluated. Haldeman's ability correctly to hear the phonetic features of a language seems to have varied in competence from one language to another. Although his Kanza and Osage data are the most accurately recorded Dhegiha data of their time (and although very little else was available for Kanza when Haldeman's data appeared), his transcription is still far from adequate. This is possibly the result of his imperceptions of certain sounds. Nonetheless the transcriptions list some forms which differ in phonological shape or sense from the modern forms of these words, and as such they have some historical significance.

REFERENCES

- Cumberland, Linda and Robert L. Rankin (eds.). 2012. *Kaá'nze Íe Wayáje. An Annotated Dictionary of Kaw (Kanza) Kaw-English/English-Kaw*. Kaw City, OK: Kaw Nation.
- García, Bartolomé 1760. *Manual para administrar los santos sacramentos de penitencia, eucaristia, extrema-uncion, y matrimonio: dar gracias despues de comulgar, y ayudar a bien morir a los indios*. Queretaro: no publisher cited.
- Haldeman, S. S. 1859. Analytic Orthography; an investigation of the sounds of the voice, and their alphabetic notation. *Transactions of the American Philosophical Society* XI: 259-401.
- Haldeman, S. S., 1860. *Analytic Orthography*. Philadelphia: J B Lippincott.
- Haldeman, S. S. 1872. *Pennsylvania Dutch, a dialect of South German with an infusion of English*. London: Trübner.
- Jones, D. 1909. *The Pronunciation of English*. Cambridge: Cambridge University Press.
- Kelly, J. and J. Local. 1989. *Doing phonology: observing, recording, interpreting*. Manchester University Press.
- Lepsius, K. R. 1863. *Standard Alphabet for Reducing Unwritten Languages and Foreign Graphic Systems to a Uniform Orthography in European Letters*, 2nd edition, London/Berlin.
- Lesley, J. P. 1881. Memoir of Samuel Stedman [sic] Haldeman. *Proceedings of the National Academy of Sciences* 139-172.
- Maximilian zu Wied-Neuwied. 1839-1841. *Reise in das innere Nord-America in den Jahren 1832 bis 1834*, 2 Textbände und 1 Bildband mit Illustrationen von Karl Bodmer, J. Hölscher, Koblenz 1839-41.
- Pilling, James Constantine. 1887. *Bibliography of the Siouan languages*. Bureau of American Ethnology Bulletin 5. Washington: Government Printing Office.
- Quintero, Carolyn. 2009. *Osage dictionary*. Norman: University of Oklahoma Press.

Rankin, Robert L. 1994. Notes on the Kansa Word List of Maximilian, Prince of Wied.

Kansas Working Papers in Linguistics, 19 (2): 201-210.

Rankin, Robert L. 1978. The Unmarking of Quapaw Phonology: A Study of Language Death.

Kansas Working Papers in Linguistics 3: 45-52.

Rankin, Robert L. 1982. A Quapaw Vocabulary. *Kansas Working Papers in Linguistics* 7:

125-52.

Rankin, Robert. 2005. Quapaw. In *Native Languages of the Southeastern United States*,

Hardy, Heather K. and Janine Scancarelli (eds.) Lincoln: University Press. 454-498.

Riggs, M.A. 1852. *An English and Dakota vocabulary*. New York: C. Greathead.

Rood, David S. 1975. *Wichita Grammar*. New York: Garland.

Rood, David S. 1996. Sketch of Wichita, a Caddoan Language. In *Handbook of North*

American Indians, vol 17: Languages, edited by Ives Goddard, 580–608. Washington:

Smithsonian Institution.

WEBLINKS

omahalanguage.unl.edu/dictionary/ Accessed 28 October 2014.

ⁱ I first heard of Bob's Siouan work in the 1980s and grew to know it better in the course of the following decade, thanks to the generosity of him and that of Dave Costa, which I eagerly acknowledge. Bob's example of someone switching linguistic fields and taking the opportunity to document languages before they passed completely out of existence (which he accomplished at a time before language endangerment was regarded as an

important concern) impressed me greatly, as did his use of philological sources. Having presented a paper at the Siouan-Caddoan Conference I was lucky enough to have dinner with him and Giulia Oliverio at the SSILA Meeting in Albuquerque in 1995, where he was kindly, erudite and hilarious company. The study of Native American language is so much the poorer for his passing.