

Medefaidrin: resources documenting the birth and death language life-cycle

Dafydd Gibbon¹, Moses Ekpenyong², Eno-Abasi Urua²

¹Universität Bielefeld, Germany; ²Univerity of Uyo, Nigeria

gibbon@uni-bielefeld.de, ekpenyong_moses@yahoo.com, amandinyene@yahoo.com

Abstract

Language resources are typically defined and created for application in speech technology contexts, but the documentation of languages which are unlikely ever to be provided with enabling technologies nevertheless plays an important role in defining the heritage of a speech community and in the provision of basic insights into the language oriented components of human cognition. This is particularly true of endangered languages. The present case study concerns the documentation both of the birth and of the endangerment within a rather short space of time of a 'spirit language', Medefaidrin, created and used as a vehicular language by a religious community in South-Eastern Nigeria. The documentation shows phonological, orthographic, morphological, syntactic and textual typological features of Medefaidrin which indicate that typological properties of English were a model for the creation of the language, rather than typological properties of the enclaving language, Ibibio. The documentation is designed as part of the West African Language Archive (WALA), following OLAC metadata standards.

1. Language birth and death

The documentation of language decline and death has reached a high level of awareness during the past decade. The documentation of language birth is an unusual notion, because the birth of languages generally takes at least a generation, from language contact, borrowing and code-switching, through pidgin languages used by adults to creoles (with native speakers). Another language birth process, which takes much longer, is the creation of a new language, such as French, Italian or Spanish, through numerous intermediate stages from a typologically different ancestor language, in this case Latin.

But there are many other factors involved in language birth, one of which is the 'invention' of an artificial language. The most well-known artificial languages are those of logic, mathematics and computation on the one hand, and international languages motivated by romantic internationalism on the other, such as Esperanto. A third category of artificial language is the religious *spirit language*, ranging from *glossolalia* ('speaking with tongues') to the vehicular spirit languages. The case to be presented here concerns the Medefaidrin language, developed and used as a vehicular language in South-Eastern Nigeria. The particular interest of this language lies in its manifestation of language contact in a rapid and almost complete language life-cycle from birth through active use to moribundity in the space of some 80 years. Similar rapid cycles have apparently not been documented.

The documentation of language endangerment and language death, and the creation of linguistic and technological resources for languages in these situations have become strongly focussed in documentary linguistics and increasingly also in the human language technologies in recent years. The language life-cycle of *birth-development-endangerment-death-rebirth-...* has rarely been focussed in its entirety (Figure 1). This contribution reports on the documentation of the birth, development, and endangerment phases of Medefaidrin.

The immediate objectives of the documentation initiative are: first, to save rapidly deteriorating paper media and to provide a basis for understanding the factors determining the 'genesis' and 'exodus' (terms due to Henderson 1982) of this language, and their outcomes; second, to render the data sustainable and reusable for hermeneutic and structural linguistic analysis and for teaching purposes, ultimately in the form of an XML encoded archive. We describe the basic resource creation steps of sustainable document representation by means of images and transcriptions. Further applications oriented documentation and development procedures with script-independent or script-adaptive OCR are currently not planned.

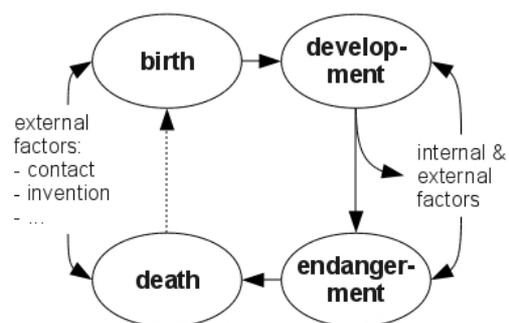


Figure 1: Language birth, (iterative) development, endangerment, death, re-birth cycle.

The documentation is being integrated into the West African Language Archive (WALA) described by Gibbon et al. (2004) and the metadata will be documented according to OLAC standards (Bird & Simons 2001). The multimodal and text technological resource documentation procedures follow Carstensen et al. (2009), Gibbon et al. (2000) and Lobin (2009). First we outline the special features of the Medefaidrin language. Then we describe the resource documentation techniques used in the Medefaidrin resource documentation project.

2. The Medefaidrin language

Medefaidrin is a Christian ‘spirit language’ of the Oberi Okaiame (‘church freely given’) Ibibio community in South-Eastern Nigeria. From a sociolinguistic point of view, Medefaidrin is an artificial special purpose language. The language emerged around 1927, and is particularly interesting as a purpose-designed vehicular language with language contact features in relation to English and Ibibio.

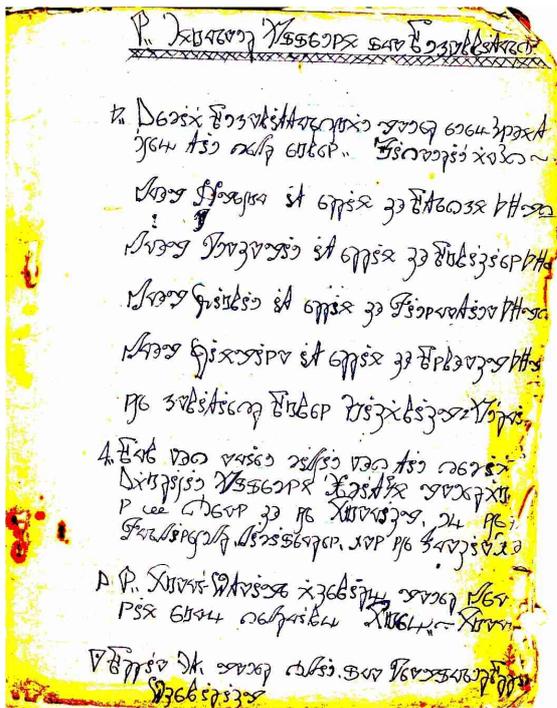
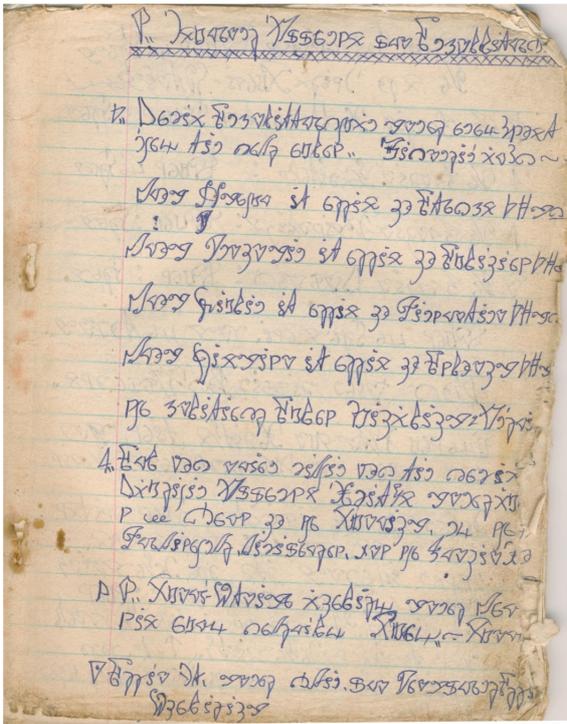


Figure 2: Scan of notebook written in Medefaidrin script, unfiltered (top) and filtered for colouration, brightness, and contrast (bottom).

Medefaidrin was inspired and rapidly developed by Ibibio speakers in a very short space of time around 1927. The language was consciously developed and used by adults, and not filtered by the universal constraints of the child language acquisition process. The language functioned from the start essentially as a ‘secret language’ identifying the Oberi Okaiame religious community, and enabling this community to practise their religion and conduct their daily lives without interference from the colonial administrators. Unfortunately this turned out to be a self-defeating strategy which led to increased persecution. The functionality of the language includes religious ceremonies and formal communication with official institutions, in community schools, grammar and mathematics are taught in Medefaidrin.

Medefaidrin has attracted scholarly attention, especially Adams (1947). Even a cursory examination of the data shows that Medefaidrin is definitely not a dialect of Ibibio but a distinct language, with different lexical, phonological, orthographic, morphological and syntactic features. However, further analysis shows that Medefaidrin shares many typological features with English and few features of Ibibio, which could be interpreted as evidence of contact with English. The language is not only spoken but also written; the unique Medefaidrin script shows typological similarities with roman script conventions. Other, including inspirational, sources remain unknown.

A number of artificial languages of this kind were developed throughout West Africa during the late colonial period in the early 20th century, and Medefaidrin is taken as a possibly not untypical case. During a comparable period, artificial languages such as Esperanto and Volapük were developed in Europe, but with inverse functionality: universality rather than protection of a specific culture (cf. Cooper 1991, Coulmas 2002, Dalby 1968). The two functionalities are two sides of the same coin: a reaction to the beginnings of globalisation in the early 20th century.

3. The Medefaidrin data

The available data types were investigated during a number of fieldwork expeditions to Edem Urua and Ididep in the Ibiono & Itu Local Government Areas, Akwa Ibom State, Niger Delta region of Nigeria, where the Oberi Okaiame community is located. In addition to original fieldwork, selective data were obtained from Adams (1947), Essien (nd), Hau (1961) and Udofia (1953).

The primary written data are mainly in the form of handwritten notebooks on grammar and mathematics, and letters. Technical media such as typewriters were not available to the community. The media have been subjected to considerable wear and tear and are in very poor condition. An example of Medefaidrin written data is shown in raw and enhanced scans in Figure 2. An official letter is shown in Figure 3.

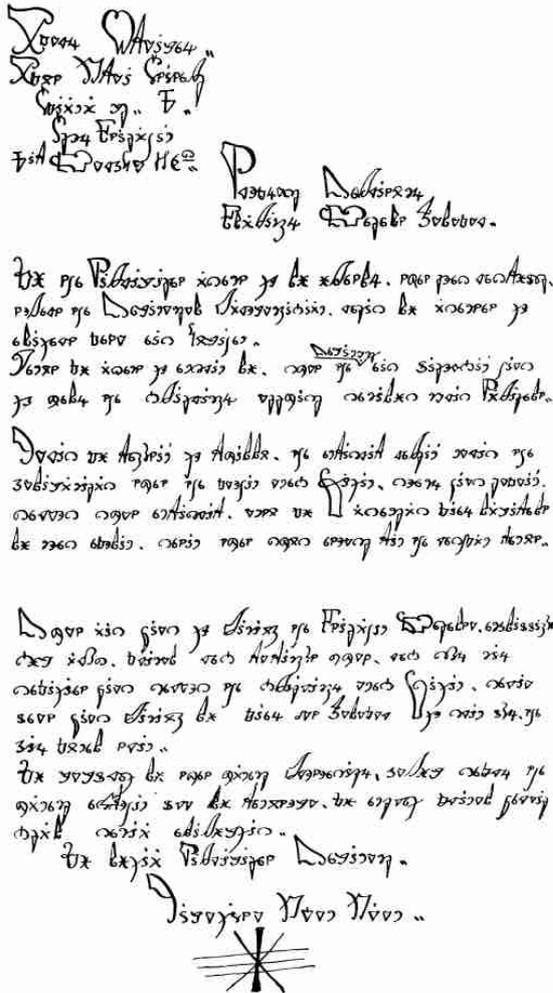


Figure 3: Official letter in handwritten Medefaidrin script.

In addition to the primary data, preliminary linguistic analysis by Adams (1947) produced a character table, a list of numerals up to 100, an elementary sketch grammar. The fieldwork expeditions additionally produced video and audio recordings. Since transcriptions and annotations of the video and audio recordings are not yet available, initial work concentrated on written language documentation.

Since development of automatic optical character recognition is not on our agenda (the paucity of data indicate that this would be gross overkill), the written data are simply scanned and then noise-filtered for readability.

4. Document objects

The Medefaidrin texts are not available electronically, nor is there as yet a numerical encoding for the Medefaidrin alphabet, so specific data structures were designed for documentation purposes. The data types and the associated documentation procedures are outlined in the following subsections.

4.1 Handwritten documents

The documents are in poor physical condition (in some cases falling apart) and almost illegible in the

original. Therefore, independently of material preservation steps, the documents were carefully scanned and the scans were archived. Additionally, the scans were enhanced for further processing, using a graphics processor in order to improve readability by changing colouration, brightness and contrast, and also archived. In order to be able to make exact reference to the documents, including both text and marginalia, a vector $V = [document, type, pagenumber, C]$ was defined as a basis for an XML implementation, where C is a subset of cells c , in a two-dimensional matrix M , and where each c is a pair of page location coordinates $[x, y]$. Each object (generally a character) is assigned to a set of such coordinates according to its position on the page. In general the coordinates in C are adjacent, but discontinuous objects may also occur. The coordinates are currently defined in millimetres, but a relative scale 0, ..., 999 is being considered.

4.2 Interpretation of handwritten documents

The oldest surviving Medefaidrin expert is over 80 years old, and other Medefaidrin users are older adults. There is a school for children, but the language is essentially not being passed on successfully. A few translations of fragments into Ibibio are available, but there are none into English. The most urgent need for documentation is therefore to create a sound and sustainable basis for future hermeneutic and structural analysis and at a later date possibly the absence of native speakers.

4.3 Alphabet code chart

The alphabet code chart is comparatively straightforward to document. The coordinate vector technique used for the handwritten documents in general can be used, but since the items are uniquely defined and identifiable as objects, they can simply be numbered sequentially [*document, type, pagenumber, n*]. Time and funding permitting, a font and possibly a tentative Unicode assignment will be designed.

Several of the characters show resemblances to roman characters, but most are uniquely structured. In some cases the phonetic interpretation is not completely clear. As in other codified languages, a distinction is made between the *character name* and the *phonetic interpretation*, e.g. as in English 'k' has the character name /kei/ and the phonetic interpretation /k/. It is not yet clear whether the characters represent phonemes or syllables. A full discussion of the character glyph features is not possible within the framework of the present contribution.

4.4 Numeral chart

A scanned document with the main features of the Medefaidrin numeral system is shown in Figure 5 (Adams 1947). The numbers are given in full from 1 to 40, and then by steps of 20 to 100. Numbers above 100 are not available.

It is not yet clear whether a more complex hierarchical microstructure for homonymous and polysemous entries will be required. Note that pronunciation is also documented in the dictionary, through a link assigned to the cropped scan of the word in Medefaidrin script in the first column.

6. From documentation to description

6.1 Typology of Medefaidrin

The objective of this documentation is to investigate the spirit language Medefaidrin in an attempt to gain new linguistic insights into the language birth and death cycle.

Using transcriptions based on the previously given alphabetic and numerical tables a number of interesting results emerged, which suggest that Medefaidrin has features of a contact language with English and – surprisingly – less so with Ibibio. Above all, the language has the distinctive flavour of a natural language, with recognisable typological properties which are also found in other languages.

6.2 Surface modalities

6.2.1. Phonology

Medefaidrin has a phonology with plausible conventional syllable phonotactics Adams (1947:26). Unlike the enclaving Ibibio, which is a tone language, Medefaidrin is a stress language, though apparently adaptation to Ibibio is taking place. Medefaidrin has voiced fricatives and affricates, /dzjaʃ/, /gizn/, /ruzerd/ and /dzibreant/, which are absent in Ibibio, and also frequent consonant clusters: /seminant/ (Holy Spirit), /atieft/, /dzibreant/, /edikanapt/, /enikrismas/, /ekenskwak/, /edipikn/, which are rare in Ibibio. Further examples of consonant clusters are /fenslet/ (to forgive), /cliffin/ (to know), /xpil/ (to pluck), /osprid/ (quickly), /texran/ (table), /dabt/ (cup).

6.2.2. Orthography

The orthography of Medefaidrin is apparently phonemic, though this is not completely settled (Figure 4). The script is unique, though inspection reveals that a number of glyphs are very similar to the Latin alphabet, albeit without the semantics which the glyph has in the context of the printed or handwritten Latin alphabet. The glyphs associated with ‘z’, ‘fe’, ‘dyu’ resemble handwritten Latin ‘z’, ‘s’ and ‘x’ respectively. A distinction between lower case and upper case is made, revealing a possible influence from English, and punctuation corresponds functionally to a subset of English punctuation.

6.3 Morphology

6.3.1. Parts of speech

There is no phonotactic or morphotactic structural difference between nouns and verbs in Medefaidrin, unlike the enclaving language Ibibio (Table 1), in

which vowels begin with one of a small set of vocalic noun class prefixes. Grammatical words tend to be monosyllabic, lexical words disyllabic, but this is not a fixed rule.

The complex set of tonally marked verb prefix and suffix inflections of Ibibio is also completely missing in Medefaidrin.

Word	POS	Gloss
sakdinet	noun	<i>black person (African)</i>
xpradinet	noun	<i>white person (Caucasian)</i>
ranked	noun	<i>orange</i>
primol	verb	<i>like/love</i>
binet	verb	<i>touch</i>
fenslet	verb	<i>forgive</i>
senio	pron	<i>we (1st person plural, subject)</i>
manet	modal	<i>must</i>
sai	modal	<i>can</i>
gias	pron	<i>us (1st person plural, object)</i>
ak	pron	<i>it</i>
swis	dem pron	<i>this</i>
dge	det/art	<i>the (def)</i>
aban	adj	<i>afar</i>
yanod	adj	<i>new</i>
osprid	adv	<i>quickly</i>
dio	prep	<i>on</i>
kin	prep	<i>in</i>
yudia	prep	<i>over</i>
dfe	prep	<i>of</i>
arien	conj	<i>and</i>

Table 1: Examples of Medefaidrin parts of speech (POS).

6.3.2. Inflection

Inflection is by suffixation, e.g. noun plurals are formed by *s/z* suffixation, as in English regular plurals (Table 2).

Singular	Plural	Gloss
drin	drins	<i>word(s)</i>
dyup	dyups	<i>thing(s)</i>
anigrein	anigreins	<i>friend(s)</i>

Table 2: Medefaidrin noun plural inflection.

6.3.3. Word formation

Word formation is similar to English, with MH (modifier-head) precedence (Table 3), but with some inconsistency. The status of the constituent parts as suffixes or independent stems could not be clarified.

6.3.4. Number composition

As already noted, the table reproduced in Figure 5 reveals a pure base 20 pattern: the numbers up to and including 19 are monomorphemic, while 20 to 39 share the prefix “edi-”, and 40, 60 and 80 have

the corresponding roots for 4, 6 and 8, respectively..

The number 20 itself has an additional final component; whether this signifies zero or not is not known. The number 40 has the structure “krei+do” (“krei” from “greido”, 2, and “do” from “edi”, 30), i.e. two times twenty. Similarly, “setado” = “seta+do” = three times twenty = 60 and “aidudo” = “aidu+do” = four times twenty = 80.

ai+ma	<i>north</i>	se+ma	<i>south</i>
piu+paya	<i>east</i>	piu+zed	<i>west</i>
ant+onim	<i>ground</i>	int+onim	<i>floor</i>
a+sped	<i>right</i>	se+sped	<i>left</i>
me+funsy	<i>good morning</i>	meya+funsy	<i>good morning (response)</i>
baisy	<i>spoon</i>	ku+baisy	<i>ladle</i>
em+sor	<i>mortar</i>	em+siy	<i>pestle</i>
piu+paya	<i>east</i>	piu+zed	<i>west</i>

Table 3: Medefaidrin compound word formation.

Ibibio and other neighbouring languages have a base 5 system. English (“score” = 20, “four score” = 80) and French (“quatre-vingt dix-neuf” = 4 x 20 + 19 = 99) retain traces of a base 20 system. The English “score” is archaic, but found in standard Bible translations, so it may be suggested (though hard proof is not available) that the numeral system may be influenced by biblical English, with which the originators of Medefaidrin were familiar. Several of the glyphs used for the numerals bear a distinct resemblance to internationally standardised Arabic numeral glyphs. For example, the first three digits are rotations of “4”; five is a double-barred “2”, ten resembles the barred “7” used in continental Europe.

6.4 Syntax

The sentence structure of Medefaidrin is SVO, and modifier-head precedence is MH:

atiu	veid	sueny	ranked	
<i>I</i>	<i>have</i>	<i>many</i>	<i>orange(s)</i>	
atiu	xtran	zu	dopitel	maghrn
<i>I</i>	<i>walk</i>	<i>to</i>	<i>school</i>	<i>everyday</i>
feas	giad	binet	dge	efuel
<i>let</i>	<i>us</i>	<i>touch</i>	<i>the</i>	<i>egg</i>
sebrifais	fensel	mium		
<i>please</i>	<i>forgive</i>	<i>them</i>		
7	kwid	liu	primol	ak
?	<i>do</i>	<i>you</i>	<i>love</i>	<i>it</i>
nipgiin	xpil	dge	ranked	
<i>do not</i>	<i>pluck</i>	<i>the</i>	<i>orange</i>	

6.5 Text structure

Inspection of the texts shows that text structure and layout correspond closely to the conventions used in English (and in other European languages). The following text linguistic features are particularly conspicuous:

- division into paragraphs,
- left-right lines,

- punctuated sentences,
- space-separated words,
- numbered lists,
- tabulation,
- marginalia in Medefaidrin
- marginalia in Ibibio (‘Work be here. Women work.’),
- Ibibio translation of longer passages (e.g. about the duties of female assistants),
- Ibibio commentary,
- conventional business letter structure.

All of these text structural features are characteristic of European text conventions.

7. Conclusions and perspectives

The documentation procedure so far has shown that the Medefaidrin ‘spirit language’ is an artificial special purpose vehicular language which represents a considerable intellectual achievement by members of the Oberi Okaimε community. Initial basic linguistic documentation has revealed a number of unique features, including the vocabulary itself, and the conspicuously innovative script, and the pure base 20 number system.

It is tempting to speculate about the creative processes underlying the creation of the Medefaidrin language and script, and indeed, closer inspection even at this initial level of documentation has revealed not only unique features but also several features which are strikingly English-like: complex syllables which are more English-like than Ibibio-like; stress-based prosody rather than tone-based prosody; SVO word order like English; MH modifier head order in compounds and adjective-noun constructions; a script with several letters which are Latin-like; English text structure and layout conventions. These English-like features suggest that English, perhaps biblical English and the English of charismatic missionary preachers, provided the grammatical model for the development of the language. In view of the inspirational perspective on Medefaidrin in the community itself, this ‘well-founded speculation’ is certain to be controversial; the similarities between Medefaidrin and English are clear, however, from the point of view of linguistic typology.

The language may appear marginal in the quantitative terms of number of speakers, but the specific insights into the typology of a ‘spirit language’ which have been revealed by even this first level of resource creation justify continuation with further state of the art documentation efforts in order to preserve and possibly to help to maintain the language. Work is in progress on completing the documentation of the resources in the standard OLAC recommended format (Bird and Simons 2001), on integration with the WALA archive, and on examining the generalisability of the results in relation to other ‘secret languages’ of the late colonial period in West Africa.

8. Acknowledgements

We gratefully acknowledge the generous assistance of members of the Oberi Okaimé community, insightful audience comments on a presentation during the WOCAL 6 conference in Köln, August 2010, and helpful LREC reviews. The work is partially funded by the World Bank Step B project.

9. References

- Adams, R. F. G. 1947. Oberi Okaimé: A New African Language and Script. *Africa: Journal of the International African Institute*, Vol. 17, No. 1, (Jan., 1947), pp. 24-34. Edinburgh University Press.
- Bird, Steven and Gary Simons. 2001. The OLAC metadata set and controlled vocabularies In: *Proceedings of the ACL 2001 Workshop on Sharing Tools and Resources*, pp. 7-18.
- Carstensen, Kai-Uwe, Christian Ebert, Cornelia Endriss, Susanne Jekat, Ralf Klabunde and Hagen Langer (eds.). 2009. *Computerlinguistik und Sprachtechnologie. Eine Einführung*. 3rd revised and extended edition.. Heidelberg: Spektrum Akademischer Verlag.
- Cooper, Robert L. 1991. Dreams of scripts: Writing systems as gifts of God. In: Cooper, Robert L. & Bernard Spolsky, eds. *The Influence of Language on Culture and Thought. Essays in Honor of Joshua A. Fishman's Sixty-Fifth Birthday*. Berlin: Mouton de Gruyter.
- Coulmas, Florian. 2002. *Writing Systems*. Cambridge: Cambridge University Press.
- Dalby, David. 1968. The indigenous scripts of West Africa and Surinam: their inspiration and design. *African Language Studies* 9:156-197.
- Gibbon, Dafydd, Inge Mertins & Roger Moore, eds. (2000). *Handbook of Multimodal and Spoken Language Systems*. New York etc.: Kluwer Academic Publishers.
- Gibbon, Dafydd, Firmin Ahoua, Eno-Abasi Urua, Eddy Aimé Gbery and Moses Ekpenyong. 2004. WALA: a multilingual resource repository for West African Languages, in: *Proceedings of LREC 2004*, Lisbon.
- Essien, O. E. nd. Religion, politics and language: a case study of Oberi Okime. (abstract).
- Hau, Kathleen, 1961. Oberi Okaimé script, texts and counting systems. In: *Bulletin de l'Institut Francais d'Afrique Noire. Série B, Sciences humaines*. 23 (1-2), January-April 1961, pp. 291-308. Dakar.
- Henderson, Eugénie J. A. 1982. Tonogenesis: some recent speculations on the development of tone. Volume 80/1:1-130 (November 1982).
- Lobin, Henning. 2009. *Computerlinguistik und Texttechnologie*. München: Wilhelm Fink.
- Udofia, Akpan Akpan. 1953. Dictionary, transcriptions and translations of texts. In: Hau, Kathleen 1953. *Oberi Okaimé Script, Texts and Counting System*. Material relating to the Oberi Okaimé language of southeastern Nigeria.

10. Appendix: lexicon extract.

Scan + audio	POS	Transcr	Ibibio	English	Medefaidrin example	Ibibio example	English example
	Pron	atiu	ami	I	Atiu sai vus ak	ami mmekan inam	I can do it
	Verb	nimazid	wed	write	Nimazid cc colia zu xium	wed deta nO enje	Write a letter to him
	Noun	colia	leta	letter	Hmbien dge colia	mmO deta odo?	Where is the letter?
	Conj.	capium	nyun	and	Lauz capium trek	di dia NkpO	Come and eat
	Verb	trek	dia	eat	Kwid liu trek liuzio ekyu	afo (amaadia NkpO)	Did you eat your food?
	Noun	dukuma	ikpon; mkpon	cocoyam	Dukuma eis colen fra mea	ikpON OfOn je ami	Cocoyam is good for me
	Noun	abinzy	bia	yam	Ruzaid abinzy fra mea	nyam bia nnO mien	Sell yam for me
	Noun	fosra	garri	garri	Nipgin trek fosra etuals	kuudia garri afd ini	Do not eat garri always
	Noun	yisru	usun	foo foo	Nipgin trek yisru axol	kuudia usuN ikpOON	Do not eat foofoo alone

Table 4: Extract from Medefaidrin lexicon (the script scan also links to a token in a Medefaidrin recording).