Transcription and orthography in two endangered languages of Ethiopia: Ts'amakko and Ongota

GRAZIANO SAVÀ
Università degli studi di Napoli "L'Orientale"

ABSTRACT

This paper touches upon the process of mapping spoken languages onto writing systems. Case studies relating to two endangered languages of Ethiopia, Ts'amakko and Ongota are presented. The discussion concerns two kinds of mapping, transcription for descriptive purposes and orthography for literary and literacy purposes. It is shown that transcription is more scientific and precise but less readable than orthography, that is more user-friendly for the wider public and the community of speakers.

KEYWORDS

Ts'amakko; Ongota; transcription; orthography; endangered.

1. TS'AMAKKO AND ONGOTA

Ts'amakko (also Tsamay and similar) and Ongota (also ʕongota, Birale and similar) are found in the Bena-Tsemai district (*woräda*) of the South Omo Zone, which is one of the administrative sections of the Southern Nations Nationalities and Peoples Federal State of Ethiopia. Ts'amakko is a Cushitic (therefore Afroasiatic) language and belongs to the East branch of this group. The classification of Ongota is uncertain (see below).

Ts'amakko and Ongota are two endangered languages. They are part of the 28 languages that are declared in danger or extinct by the UNESCO in the Atlas of the World's Languages in Danger (http://www.unesco.org/languages-atlas/index.php). In particular, Ts'amakko is considered as definitely endangered, while Ongota is considered critically endangered. In terms of number of speakers Ts'amakko, according to the Atlas, counted 8621 speakers in 1998, while the Ongota speakers, were 8 in 2007. The official numbers by the Ethiopian government are controversial. The 2007 Population and Housing Census of Ethiopia (Central Statistical Agency of Ethiopia 2012) reports 17.390 Ts'amakko speakers in the South Omo Zone. However, according to the previous census of 1994 the Ts'amakko mother tongue speakers were 7.820 (Central Statistical Authority of Ethiopia 1996). Therefore, it is hard to believe that in ten years the number of Ts'amakko speakers was more than doubled. The statistics for Ongota are also to be considered with great care. According to the last 2007 census, the number of Ongota mother tongue speakers was 469 out of a population of 897. However, the last counting of the present writer in 2018 was a bit more than 100 for the ethnic group members with 7 speakers.

What is a speaker of Ongota should be better defined and should be done in relation to the Ts'amakko language. In fact, Ongota is so endangered because the people switched to Ts'amakko in everyday conversation. Those who still know Ongota are only a group of elders that use it scantly in situations in which they do not want to be understood, as a sort of secret language. Therefore, in everyday life, Ongota has a very little role.

The start of language switch was probably due to the decision of the present last speakers of Ongota to stop teaching it to their children and to teach them Ts'amakko instead. They already were bilingual due to the contacts with neighbouring Ts'amakko. Indeed, the Ongota live within the Ts'amakko territory, along the Wäyto River, inside the valley of this river. The Ongota were hunter-gatherers, fishermen and bee-keepers, but due to the contact with pastoralist Ts'amakko, they began agriculture and small-scale cattle breeding.

The alliance with the Ts'amakko seems to be one of several in the history of the Ongota. And each time there were if not a real language shift, at least heavy bilingualism and language influence. The result is that it is hard to find proper classification of Ongota, which presently is not classified, and which could be an isolate.

2. TS'AMAKKO DESCRIPTION, DOCUMENTATION TRANSCRIPTION AND ORTHOGRAPHY

Ts'amakko was described in *A Grammar of Ts'amakko* prepared by the writer as the accomplishment of a PhD at Leiden University (Savà 2005). The only noteworthy previous description, with comparative aims, is an article by Hayward (1989). The present paper focuses on the transcription used in Savà (2005).

The present paper will also deal with the Ts'amakko transcription system adopted in a documentation project which was part of a wider project on some Afroasiatic languages called *CorpAfroAs,* a *Corpus for Spoken Afroasiatic Languages: Prosodic and Morphosyntactic Analysis* (funded by the French Agence Nationale de la Recherche. Mettouchi, Vanhove & Caubet 2015).

An orthographic system was adopted for the preparation of a storybook in Ts'amakko. This is also described in the following paragraphs.

Finally, an orthographic system was designed and adopted for a literacy project in Ts'amakko carried out by the Summer Institute of Linguistics. This system will also be discussed.

2.1 GRAPHEMES AND PHONOLOGICAL INVENTORY OF CONSONANTS IN A GRAMMAR OF TS'AMAKKO

To start with, here is the chart of the Ts'amakko consonant phonemes in which phonemes are represented by IPA characters:

	Bilabial		Alveolar		Palatal		Velar		Uvular		Pharyng.		Laryng.	
	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v
Obstruents														
Stops	p	b	t	d	tʃ		k	g						ʔ
Fricatives			s	z	ʃ	ʒ			χ		ħ	ʕ	h	
Glottalic		ɓ	ts'	ɗ	tʃʼ			ɠ	q'					
Non-obstruents														
Glides		w				j								
Lateral				l										
Trill				r										
Nasals		m		n		(ɲ)								

176

In the following chart the phonological transcription is "normalized" by spelling conventions that replace some IPA characters:

/tʃ/ = \<c\>

/tʃ'/ = \<c'\>

/ʃ/ = \<š\>

/ʒ/ = \<ž\>

/χ/ = \<x\>

/j/ = \<y\>

Below is the chart showing the transcription of Ts'amakko consonantal phonemes as they appear in Savà (2005)[1]:

	Bilabial		Alveolar		Palatal		Velar		Uvular		Pharyng.		Laryng.	
	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v
Obstruents														
Stops	p	b	t	d	c		k	g					ʔ	
Fricatives			s	z	š	ž			x		ħ	ʕ	h	
Glottalic		ɓ	ts'	ɗ	c'			ɠ	q'					
Non-obstruents														
Glides		w				y								
Lateral				l										
Trill				r										
Nasals		m		n		(ɲ)								

The representation of some phonemes still requires the use of diacritics or special IPA characters. As can be seen, the choice was to indicate palatalization of the fricative sibilants with the haček (ˇ), while it was deemed superfluous to add an haček to the other sibilants \<c\> and \<c'\>. The apostrophe \<'\> is used to mark ejectives, while for implosives the special IPA charac-

---

[1] As conventionally common, the // contain phonemes, the [ ] contain phonetic realisation and \< \> contain graphemes.

ters for these sounds are adopted. Specific IPA characters represent also the voiceless pharyngeal fricative /ħ/, the voiced pharyngeal fricative /ʕ/ and the voiced laryngeal stop /ʔ/, while the voiceless uvular fricative is simply <x> a simplification of IPA [χ]. The palatal nasal <ɲ> appears rarely only in loanwords.

All consonant phonemes can appear geminated. Gemination is represented by doubling the character.

2.1.1. CONSONANTAL GRAPHEMES AND ALLOPHONIC REALISATION

Several phonemes have allophonic realisations due to their position in the word or for free variation:

- all stops and glottalised obstruents are partially released word-finally. The only exception is /p/, which appears as [f], as we will see below;
- the voiced palatal fricative /ž/ is affricated as [dʒ] in postconsonantal and geminated position;
- the voiceless uvular fricative /x/ is trilled [χ̃] before high vowels;
- the voiceless pharyngeal fricative /ħ/ can also be trilled [ħ̃] and it is pronounced with a particularly powerful airflow;
- the voiced pharyngeal fricative /ʕ/ is often glottalised [ʕʔ] when in word-initial position and geminate;
- the voiceless alveolar ejective affricate [ts'] can be optionally pronounced as a fricative [s'];
- the ejectives /ɓ/ and /ɠ/ can be freely devoiced into [ɓ̥] and [ɠ̥] respectively;
- the position of the tongue in the articulation of the alveolar implosive /ɗ/ is apical [ɗ̺] or laminal [ɗ̻];
- the uvular ejective /q'/ is quite unstable. It can be optionally pronounced either as an affricate ejective [qχ'] or as a voiceless implosive [ʠ] or a voiced implosive [ɢ];
- another problematic phoneme is what in the grammar is transcribed as <p> but that, in fact, corresponds to a phoneme that is articulated between bilabial stop and labio-dental fricative [f] or something in between. The variants partially depend on the position in the word. In word-initial position /p/ is pronounced as a plain stop [p], an aspirated stop [pʰ] or a bilabial fricative [ɸ]. Intervocalically /p/ appears as [f] or [ɸ]. In most of the cases both allophones are possible. In pre-consonantal position /p/ is realised as /f/. As for postconsonantal position, the /p/ can only be preceded by /m/, /r/ or /l/. When preceded by /m/

it is pronounced as [pʰ]. When preceded by /r/ or /l/ it is pronounced as [f]. Final /p/ appears always as [f]. When geminated, /p/ is articulated as [pʰ:] morpheme internally and [pʰ:] or [f:] across morpheme boundaries. This brief account of the realisation of /p/ does not aim to be exhaustive but to show that this phoneme in Ts'amakko is problematic in deciding how to represent it in transcription. The decision in the grammar at the level of analysis is to leave the phonemic status of this phoneme as for its stricture parameter. At the level of transcription, it is decided to write it as <p> in initial, postnasal and geminated root-internal position and to write it as <f> in intervocalic, pre-consonantal, postconsonantal (except nasal), final and geminated across morpheme boundary position.

In all these cases, the grapheme corresponds to the representation of the phoneme regardless of the allophonic realisation. As we saw, an exception is /p/ that is actually represented by two graphemes according to its position.

Different decisions are taken if the phoneme changes due to some morphological processes. In these cases, phonological rules apply that cause a modification of the phoneme. In some cases, this modification is represented by the relevant grapheme, in some cases not. The decision depends on the fact that modification may correspond to an element already represented graphically in the transcription inventory. If there is no corresponding grapheme, the grapheme representing the phoneme is used. If the modification corresponds to an already existing grapheme, this is used. See some examples:

- the /t/ assimilates the voicing of a preceding /b/, /d/ or /g/. It becomes [d] and is transcribed as such: <d>;
- the implosives /ɓ/, /ɗ/ and /ɠ/ are devoiced before /t/. /ɓ/ and /ɠ/ become ejectives [p'] and [k'] respectively but are represented graphically as <ɓ> and <ɠ> because these ejectives are not present as phonemes and have no relevant character in the transcription system. /ɗ/, instead, not only is devoiced, but it looses glottalisation becoming [t]. In this position, therefore, it is represented as <t> since this grapheme is present in the transcription inventory;
- for the same reason, when /ɗ/ changes to [n], in contact with the suffix -ni of first person plural unmarked paradigm[2], it is transcribed as <n>;
- the alveolar nasal /n/ assimilates the place of articulation following a velar, becoming [ŋ], and a bilabial, becoming [m]. However, /n/ remains <n> because <ŋ> is not used else in the transcription system and in order to simplify the transcription avoiding the use of the IPA

---

2   The Unmarked is one of the two main verbal paradigms in Ts'amakko. It is characterised by the fact that it is neutral in terms of aspect as it can be used to describe both perfective and imperfective actions.

character ŋ. A parallel solution was adopted in the case of the allophone [m] even if there is a grapheme <m> in the system;

– a different point regards the use of <ʔ> word initially. Since, according to Ts'amakko syllable structure, there are no syllables with no onset but all the syllables must be minimally CV, those words that appear vowel-initial actually carry an initial /ʔ/. This is therefore constantly transcribed initially.

2.1.2. VOWELS

The situation of the vowels is quite simple. Ts'amakko has five cardinal vowels, /a/, /e/ /i/, /o/ and /u/ with their long counterparts /aa/, /ee/, /ii/, /oo/ and /uu/ graphically represented in the same way. Lengthening as a consequence of a phonological rule occurs when a nominal is followed by the case clitic =*ma* (Savà 2020) or, optionally, by a locative case suffix. Vowel lengthening also indicates that a name is a possessor. See examples:

manne "house"
maanne=*ma* "to the house"

pašo "field"
paš-ilo or *paaš-ilo* "in the field"

beze "male name"
ħaarko beeze "Beze's hand"

2.1.3. TONE

The tone system of Ts'amakko is based on a high tone and a low tone and has limited function load. It is not the case to make a full explanation of it in the present paper, as it suffices to state that, with some exception, tone is not marked in the Grammar of Ts'amakko. This is because the grammar provides information about the tonal characteristics of almost all word classes and verb paradigms and tone placement can be predicted from this information. An exception are those nominals that do not follow the default patterns HL or HHL. In these cases, tone is marked. The graphic rendering of tone is acute accent <´> for high tone and grave accent <`> for low tone.

2.1.4. A TRANSCRIPTION ALPHABET FOR A GRAMMAR OF TS'AMAKKO

For the sake of the present article, the consonants and the vowels as they are transcribed in *The Grammar of Ts'amakko* are put in the alphabetical order as follows:

a
b
ɓ
c
c'
d
ɗ
e
f
g
ɠ
h
ħ
i
k
l
m
n
ɲ
o
p
q'
r
s
š
t
ts'
u
w
x
y
z
ž
ʔ
ʕ

That makes 35 characters. Consider that all characters can also appear doubled to represent consonantal gemination and vocalic lengthening. Moreover, there is no distinction between majuscules and minuscules.

2.1.5. SAMPLE TEXT

In order to better exemplify the transcription system discussed above, here is a sample text extracted and adapted from *A Grammar of Ts'amakko*. Notice that punctuation marks are not used, except for the question mark, and each line corresponds to a sentence or a coherent sequence of sentences. Moreover, in *a Grammar of Ts'amakko* the texts are interlinearised with grammatical and lexical glosses, that are left over in the text below. An English translation follows:

maakke garrilo ʔaaka maakke gubalatte
q'arra garro ʔarɗo bitami na gubale maarte bitamti
maarte gubalatte gurti ba woq'ošiba ɗali
maarte ɗaltinnay garro kiyanay kaayu ka ɗali kiyi ba ʔarɗilo turɗitte sorto ki
ʔaɗɗa gubale gore buskaɓti
gore zingatte saʔate lákkíyay bukaɓe ba gubale saʔate salaħ ki xafti
gore kiyankinnay ʔine ka buskaɓti ba ʔato ʔaakkama baɗay? kiye
gaaħko ɠonɗami ba deelloma ʔolaɗi ʔise ka kiyiti
gore kesse bukaɓi gaaħko moo ɠonɗamu ba ka delay gubalatte kiye
ʔise kiyannay ʕarɗo moo ki ɗalaɗa nunnu bukaɓanku? kiyiti
tannu ʔise ka bolɠomisi
tannu garro kiya nay boɠolkonu q'ole c'oxinɗa
ʔombottanne kúnkóyay ʔaxxe c'oxonkibba ganda xumɓi c'oxankibba nunnu
ʔawkose boɠolko nata ʔaɠima šeeɠonki
boɠolko ʔombotto ʕugisi nay ʔombotto dookko ʔakka ʕuggisi nay ʔombotto
dookko ʔakka ʕuggisi nay ʔombotanne q'àrú šeeɠe xumɓi ʕugisi kiyankinnay
ʔombotanne ketta xumɓi rawti nay bolte takka ki ʕaq'i
bolte tettakka gaaraʕtema beʕi
ɗawra boɠolko ħaarkoyay ɠabbikka kʉp bayi ba q'eeɗɗa loq'a boɠoltekka
kʉp bayiti nay garaʕte nata ɓoʔti
ʔasa tannu gubalatte mala dayi? gubale ka parti
maakke garrilo ʔaaka maakke gubalatte kettay

The tale of the squirrel and the rabbit
One day the squirrel bought an ox and the rabbit bought a heifer. The heifer of the rabbit mated, got pregnant and gave birth. After the heifer gave birth, the squirrel said: "Mine gave birth", and put some placenta into the buttock of the ox. The rabbit made people gather. The people gathered at eight in the morning and the rabbit came at ten. The people said: "You made gather us

and where did you disappear?". She said: "A stone broke and I spent the day sewing it". The people who gathered said to the rabbit: "How does it come that a stone breaks and you sew it?". After that, she was made queen. So, the squirrel said: "Milk cattle for the queen!". They milked ten containers, the whole neighbourhood milked and brought to the place where the queen lived. He made the queen drink the milk containers. He made her drink one milk container and after that, he made her drink another milk container. Eventually, he made her drink all the milk containers they had brought to the place. After she had finished all those milk containers a small drop remained. "It is taboo! A queen does not take anything with the hands. She bends, licks and swallows". When the queen bent, her belly blasted. So, then, what happened to the rabbit? The rabbit died. That was the tale of the squirrel and the rabbit.

2.2. TRANSCRIPTION FOR A DOCUMENTATION OF TS'AMAKKO

A more scientific, IPA-like, transcription was used for Ts'amakko in the context of the project *CorpAfroAs,* a *Corpus for Spoken Afroasiatic Languages: Prosodic and Morphosyntactic Analysis*. In fact, each text was transcribed two times. One transcription is strictly phonetic and the second one is more broadly phonemic. In order for the phonetic transcription to reflect the spoken language more faithfully, various IPA characters, signs and diacritics were used that are not possible to discuss in the present paper. The phonemic transcription is more readable and normalized and is discussed here. The characters are still IPA and the inventory similar to the one of *A Grammar of Ts'amakko*.The only differences are that there is no ejective /c'/ [ʧ'] but only /c/ [ʧ] and the uvular /q'/ is not an ejective.

Transcriptions, annotation and translation were ordered in parallel tiers with the use of the ELAN-Corpa programme. This programme is a version created by Christian Chanard and his collaborators at the French CNRS research center LLACAN (Langues, Langage et Cultures d'Afrique Noire) of the ELAN produced and maintained by the Max Planck Institute in Nijmegen. The difference is that ELAN-Corpa has a special semi-automatic function of morpheme glossing.

From the screenshot below one can see that tier tx@SP1 (text of speaker 1) hosts the broad phonetic transcription and boundaries separate intonation units. The signs "/" and "//" are used to set the boundary of minor and major intonation units respectively, but this does not have an effect on the transcription. The characters used in the transcription are IPA characters that represent all possible nuances of the speech. The tier mot@SP1 (the French word "mot" for "word" is used to name this tier) has still a phonetic transcription with phonetic characters and shows all phonological modifications but it normalizes allophones to phonemes. On this tier boundaries separate

words. In the following tier, mb@SP1 (morpheme boundary of speaker 1) each segment hosts a morpheme, before any phonological modification. The following two tiers are glossing tier, one lexical and grammatical, one syntactic, and the last one is the translation tier. On top of everything there is a tier assigning a reference code to each intonation unit:

	00:00:26.500	00:00:27.000	00:00:27.500	00:00:28.000
ref@SP1 [312]	TSB_NARR_001_023		TSB_NARR_001_024	TSB_NARR_001_025
tx@SP1 [310]	elleχaγe::: /		405	la:ɠɠaŋki //
mot@SP1 [748]	ʔelleka	ke	/	405	la:ɠɠanki	//
mb@SP1 [1095]	ʔelele =ka	ke	/	la:ɠ	~ɠ	-anki	//
ge@SP1 [1094]	togeth CONT PL	/	retur SEM IPFV. //
rx@SP1 [1094]	ADV.M FOC PRO	/	V der.V TAM. //
ft@SP1	Together		they gathered [the animals]

It is to note that consonant length in indicated by doubling the consonantal character. Vocalic length is indicated by the IPA sign [:]. Moreover, the duration of the pauses was marked from 200 ms.

Below is the list of characters of the phonemic transcription and part of a sample text:

a
b
ɓ
tʃ
d
ʤ
ɗ
e
f
g
ɠ
h
ħ
i
j
k
l
m
n
o

p
q'
r
s
ʃ
t
ts'
tʃ
u
w
χ
z
ʔ
ʕ

gelzakko ʔaːka garro
gelzakko ʔaːka garro leʔe ʔellema kaːrink kaːrinko goʃaɗa
gelzakko gelzakko maːre goʃi
garro leʔe goʃi
gelzakko kaʔʔiba baj garro ʔato leʔeja zow
baj gelzo maːreja zow kaji
ʔelleka ke laːɟɟanki
ʔelle laːɟɟennaj q'ajnaka garro leʔeja zeji pugaɗiki ʕaːga
agarro ko pugaɗi kiʕaːga
q'ajnaka likke ʔasa q'ole goʃaɗe goʃe
kuɓa maːreja koɓa leʔejaj bajina garro ko pugaɗi kiʕaːga
baj garro ʔato moː ʤiʔtiba q'ajto χumɓika pugaɗaj bajinaj
gelzakko kijanaj agarro kijana ʔanoʔane…

The squirrel and the baboon
The baboon and the squirrel used to tend the cattle together according to an agreement. The baboon was tending female calves. The squirrel tended cows. The baboon got up and "Friend squirrel. You go with the cows". "Friend baboon. Go with the calves" [the squirrel] said. Together they gathered [the animals]. They gathered [the animals] and the day after the squirrel went with the cows, got satiated and went back home. The squirrel came back home being satiated. The day after right in the same way, they tended cattle. One [went] with the calves. The other one with the cows and the squirrel went back home being satiated. "Friend squirrel, what is that you eat and always get satiated?" He said and the squirrel said Who, me?...

2.3. ORTHOGRAPHY FOR A TS'AMAKKO FOLKTALE BOOK

There is an on-going project for the publication of some Ts'amakko folktales by the present writer and Dr Pavel Mikeš, the present Czech Ambassador in Addis Ababa[3]. The aim is to publish a booklet with a dozen of folktales for the wider public.

A proper Ts'amakko orthography to be taken as reference for similar projects was designed. Compared to the transcription used in *The Grammar of Ts'amakko* and in the *CorpAfroAs* documentation project, more readable solutions were adopted. IPA characters were avoided, except for the <ʕ> of the voiced pharyngeal fricative. As for the rest:

- the character of the voiceless laryngeal stop <ʔ> is represented by an apostrophe <'> and is not marked word initially;
- all glottalised, ejectives and implosives, are marked by the apostrophe besides the main character. This solution avoids using the IPA characters for the implosives;
- <ts'> was simplified to <s'>;
- the IPA character <ħ> for the voiceless pharyngeal fricative is replaced by <hh>.

Moreover:

- the haček is also avoided. <ž> is, therefore, represented by <j>, that, as we saw, corresponds to the affricated realisation of the voiced palatal fricative /ž/ in postconsonantal and geminated position. As for <š>, it is represented by the English-like <sh>;
- the distribution of <p> and <f> is the same as in *The Grammar of Ts'amakko*;
- all characters, consonantal and vocalic, can appear doubled to represent gemination and lengthening.

a	A
b	B
b'	B'

c	C
c'	C'
d	D
d'	D'
e	E
f	F
g	G
g'	G'
h	H
hh	Hh
i	I
j	J
k	K
l	L
m	M
n	N
o	O
p	P
q	Q
r	R
s	S
sh	Sh
t	T
s'	S'
u	U
w	W
x	X
y	Y

z	Z
'	'
ʕ	ʕ

All 34 characters can be doubled to indicate consonantal gemination and vowel lengthening.

2.4. SAMPLE TEXT

Also in this case, a sample text is presented below:

Maakke gudurkilo aaka maakke gaarmilo
Garmo qooshi. Gudurkokka qooshu. Shambo aaka abbayo qooshi. Abbayo aaka shambo qoshebba lo'o bog'inki. Lo'o ji'ankinnay lo'o ji'tilo ʕagankinnay garmo gaarkomi ka'iba
ufund'eka i'una shambose gudurkilo shambo gudurkilo kiyanay "abba ula guddo garmo na" kiyina gudurko kiyanay "ei ussa garmokka kuyo" kiyi. "Saankoka ji'nanki saanko garmokka kuyo" kiyi. Saanko ji'ankinnay joome ka'inki. "Abba ano intayikka ato intawu" kiyi shambo. Qaru garmoka ammake i'i donnay shuku-yina gudurko intayi. Gudurko intayi ooda. Oodanay garmokkana miinatte gud-donu ufund'e ka'inay garmokkanay miinatte guddonu ka'i. Suurema nunnu ridu. Abbayo suuratte lig'i aaka garmo asa abbayo ki bog'i. Shambose gudurkilo sori zowu. Garmo haysamma gudurko bog'i. Maakke garmilo aaka makke gudurkilo hayissayay d'ikkad'i.

The story of the hyena and the lion
The lion was hunting in the same field where the hyena and his son were hunting. The hyenas killed a cow and started eating it. While they were eating it the lion climbed on a tree and the hyenas could see him. The hyena child said "Father, over there is a lion, let's be careful". And the hyena father said "No, that is not a lion, it is a termite hill" and kept on eating the meat. While they were eating the father repeated "I am sure it is not a lion but a termite hill". After they got satiated, they left but the child said "Father I do not want to go first, you go first". So, the hyena father walked first. The lion followed their movements, started walking towards them and stopped wait-ing for them in a small path into the forest. The hyena father entered that path and the lion killed him. The young hyena ran away. So, the lion killed the father hyena. In this way ends the tale of the hyena and the lion.

2.5. TS'AMAKKO ORTHOGRAPHY FROM A SIL (SUMMER INSTITUTE OF LINGUISTICS) LITERACY PROJECT

SIL Ethiopia is involved in a literacy project in the Ts'amakko area. Training is organised in two locations, Luqa, the village where the present writer did his linguistic field research, and Birale, close to the Ts'amakko main town Wäyt'o. No further news on this project could be collected, besides the orthography adopted for the production of primers and other didactic material (provided by the local collaborator Haylu Berhanu Golla).

Compared to the orthography of the folktales presented above, this orthography is characterised by the following features:

– implosives and ejectives are marked by a <h> following the main character. Exceptions are <ts>, that indicate the ejective alveolar affricate, and <c>, that indicates the ejective palatal affricate /ʧ'/. Its non-ejective counterpart is with an <h>, e.g., <ch>. This is a solution adopted in several orthography of languages in the area as the ejective counterpart of this affricate is much more common than the pulmonic one. Due to the higher occurrence of /ʧ'/, the simpler character <c> is used;

– the voiced pharyngeal fricative /ʕ/ is represented by <v>. This is possible because there is no phoneme /v/ in Ts'amakko (as in all the languages of the Ethiopian area);

– the voiceless counterpart of /ʕ/, i.e., /ħ/, is represented by a simple <h>. This is possible because /h/ is not considered among the phonemes of Ts'amakko;

– there is a character combination <ny> for the palatal nasal /ɲ/;

– there is no <p>, only <f>;

– the SIL Ts'amakko orthography includes majuscule letters. In case of double characters, such as, for example, the implosive <bh>, only the first main character is majuscule, i.e., <Bh>;

– questions are indicated with the lengthening of the final vowel.

Here is the list of characters, minuscule and majuscule, in alphabetic order:

a   A

b   B

bh   Bh

c   C

ch   Ch

d	D
dh	Dh
e	E
f	F
g	G
gh	Gh
h	H
i	I
j	J
k	K
l	L
m	M
n	N
ny	Ny
o	O
q	Q
r	R
s	S
sh	Sh
t	T
ts	Ts
u	U
v	V
w	W
x	X
y	Y
z	Z
'	'

All characters can be doubled in order to represent consonant gemination and vowel lengthening. However, in those cases of double character, such as, for example, the implosive <dh>, only the first main character is doubled, i.e., <ddh>.

Unfortunately, it is not possible to provide a sample text.

3. ONGOTA

Besides being a very endangered language and a possible isolate, Ongota is largely understudied. There are only two reference articles as for description: Fleming et al. (1992/93) and Savà and Tosco (2000). The language has also been object of documentation (see below). Part of the documentation work was the creation of a quadrilingual alphabet for the community.

In the following, I will overview the transcription systems adopted in Savà and Tosco (2000) and the documentation of Ongota and the orthographic systems (Latin and Ethiopian) used for the alphabet.

3.1. ONGOTA DESCRIPTIVE ARTICLE

Savà and Tosco 2000 is the result of a fieldwork conducted by the present writer and Mauro Tosco in Jinka with the late Mole Sagane. As for the transcription system adopted, below is the chart of the Ongota consonantal phonemes in IPA characters:

	Bilabial/ labiodental		(Post) Al- veolar		Palato-al- veolar		Velar		Uvular		Pharyng.		Laryng.	
	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v
Plosives Oral		b	t	d			k	g	q				ʔ	
Glottalised				ɗ				ɠ						
Nasal		m		n										
Affricate			ts		tʃ	dʒ								
Fricative	f		s	z	ʃ				χ		ħ	ʕ	h	
Trill				r										
Approximant Central		w				j								
Lateral				l										

In the transcription system used in the article, the following IPA characters were replaced:

/tʃ/ = <c>

/ʃ/ = <š>

/χ/ = <x>

/j/ = <y>

Here is an adaptation of the chart of Ongota consonant phonemes from Savà and Tosco (2000:66). The representation of the phonemes corresponds to the graphic representation:

	Bilabial/ labiodental		(Post) Al- veolar		Palato-al- veolar		Velar		Uvular		Pharyng.		Laryng.	
	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v	-v	+v
Plosives Oral		b	t	d			k	g	q				ʔ	
Glottalised				ɗ				ɠ						
Nasal		m		n										
Affricate			ts		c	j								
Fricative	f		s	z	š				x		ħ	ʕ	h	
Trill				r										
Approximant Central		w				y								
Lateral				l										

The transcription solutions are similar to those adopted in the *Grammar of Ts'amakko*. In part this is also due to the fact that the phonological systems of the two languages are similar and, understandably so, due to the close re-lationship that they have. The palate-alveolar fricative sibilant is marked by the haček (ˇ), while there is no haček on <c>. The apostrophe <'> marks ejectives and implosives are represented by IPA characters. IPA characters are also used for the voiceless pharyngeal fricative /ħ/, the voiced pharyngeal fricative /ʕ/ and the voiced laryngeal stop /ʔ/. The voiceless uvular fricative is simply <x> a simplification of IPA [χ]. As in Ts'amakko, all phonemes can appear geminated. Gemination is represented by doubling the character.

3.1.1. GRAPHEMES AND ALLOPHONIC REALISATION

Some phonemes have optional allophones that are not marked graphically:
  – Voice opposition can be optionally neutralised. See examples of [ħ] ~ [ʕ] and [d] ~ [t]:

gaddaħħuni [gaddaħħuni ~ gaddaʕuni] "big"
gidata [gidata ~ gitata] "you (PL)

  – In the section on Ts'amakko we have seen the complexity of realisation of p/f. In Ongota the only (optional) alternation attested is between [f] ~ [pʰ] word-initially:

oxoni faʔo [faʔo ~ pʰaʔo] "to kindle the fire"

  – The palate-alveolar affricate /j/ optionally becomes a fricative.

janta [dʒanta ~ ʒanta]

It is to notice that Ongota /j/ ([dʒ]) corresponds to Ts'amakko /ž/ ([ʒ]), that has an allophone [dʒ].

Here are some of the possible grapheme modifications due to phonological rules:
  – A nasal preceding a plosive assimilates the point of articulation of the plosive. This modification is graphically represented:

tagamá "sleep (SG)!" *tagánta* "sleep (PL)!"

  – Due to sibilant harmony, the /s/ of the causative suffixes -san, -is and -as becomes a palato-alveolar š if in the verbal stem there is a palate-alveolar consonant. This modification is represented graphically:

ka = cóq "I shot" *ka = cóqšan* (instead of **cóqsan*) "I made shoot"

Although the imperative plural suffix -ta causes devoicing assimilation to a preceding voiced plosive, the resulting modification is not graphically represented, i.e.,

yeqaɖá "hiccup (SG)!" *yeqaɖtá* "hiccup (PL)!" *[yeqattá]*

3.1.2. VOWELS

As for the vowels, the situation is like in Ts'amakko: there are five cardinal short and long vowels. The long vowels are indicated by doubling the vocalic grapheme. There are occurrences of initial vowels. Therefore, the situation is not as in Ts'amakko, where initial vowels are always preceded by the character of the glottalic stop <ʔ>. The initial glottal stop can be optionally heard, but in view of its irregular presence it is never marked.

3.1.3. ACCENT

There is no systematic marking of accent, which is indicated sporadically in the description when it causes contrast between two segmentally identical words or if it has grammatical meaning. It is marked by an acute accent.

An important grammatical situation in which the accent plays a role and is marked is the distinction between past and non-past. In the following example the accent falls on the penultimate syllable in the past verb form and on the last syllable in the non-past verb form:

ka=múxi "I laughed"
ka=muxí "I laugh/will laugh"

3.1.4. A TRANSCRIPTION ALPHABET FOR SAVÀ AND TOSCO (2000)

Below is an alphabet reconstructed from the description of Ongota by Savà and Tosco (2000):

a
b
c
d
ɗ
e
f
g
ɠ
h
ħ
i
j

k
l
m
n
o
q
r
s
š
t
ts
u
w
x
y
z
ʔ
ʕ

Unfortunately, Savà and Tosco (2000) does not contain texts.

3.2. THE TRANSCRIPTION FOR THE DOCUMENTATION OF ONGOTA

A two-year postdoc grant for the documentation of Ongota was awarded in 2007 to the writer by the Hans Rausing Endangered Languages Documentation Programme. The project produced recordings, some of which transcribed, annotated and translated, of Ongota speech samples.

The transcription, annotation and translation methodology is the same as the one used for Ts'amakko in the project *CorpAfroAs,* a *Corpus for Spoken Afroasiatic Languages: Prosodic and Morphosyntactic Analysis* (see above paragraph 2.2).

3.2.1. A TRANSCRIPTION ALPHABET FOR THE DOCUMENTATION OF ONGOTA

The following points are to notice compared to the transcription of Ongota in Savà and Tosco (2000):

– consonant length is marked by doubling the character. Vowel length is indicated by the IPA mark ":";

- initial vowels are always preceded by <ʔ>;
- <j> indicates the voiced palate-alveolar central approximant, instead of <y>;
- <dʒ> indicates the voiced palate-alveolar, instead of <j>;
- according to a new analysis, <ts'>, <tʃ'> and <q'> represent ejective sounds and therefore carry the apostrophe "". Therefore, <tʃ'> corresponds to <c>;
- <ʃ> corresponds to <š> and;
- <χ> is the equivalent of <x>.

The following list of characters can be extracted from the "mot" tier:

a
b
d
dʒ
ɗ
e
f
g
ɠ
h
ħ
i
j
k
l
m
n
o
q'
r
s
ʃ
t
ts'
tʃ'
u
w
χ
z
ʔ
ʕ

A text extracted from the documentation was published in Ethnorêma by the present writer (Savà 2017). The article, and the whole journal, can be freely consulted and downloaded.

3.3. TWO ORTHOGRAPHIES (LATIN AND ETHIOPIAN) FOR AN ONGOTA ALPHABETIZATION PROJECT

A booklet with a didactic alphabet was realised in the context of the ELDP documentation of Ongota. It consists of one page per letter. On the page there is the single letter in Latin and in Ethiopian script, or *fidäl*, a figure and the corresponding word in Ongota and in Ts'amakko, written in Latin and Ethiopian script, plus the word in Amharic and in English. The name of the language is written under each word. See an example:

As for the Latin script, compared to the transcription in Savà and Tosco (2000) diacritics and IPA characters have been avoided. As a consequence:

– the implosives are indicated by a < h > after the main character as in the SIL Ts'amakko orthography;
– the < š > is represented by < sh >;
– the voiceless laryngeal stop < ʔ > is reduced to ', but it does not appear word ini-tially;
– the voiced laryngeal fricative < ʕ > changes to < v >, as in the SIL system for Ts'amakko (see above paragraph 2.5.);
– the voiceless laryngeal fricative < ħ > is represented by hh;
– besides the < c > in Savà and Tosco (2000), there is an ejective marked as < c' >.

One may notice that the name of the Ongota language actually has a voiced pharyngeal fricative at the beginning, i.e., ʕongota [ʕoŋgota]. The spelling Ongota is used as a simplification.

As for the *fidäl* script, the typical syllabary system with change of vowel by modification of the basic character carrying the vowel /ä/ is adopted. Reference is made to the use of the script in Amharic. Moreover, the following remarks are in order (for the Latin transcription of words in *fidäl* it applies the system used in Savà and Tosco 2000, see above):

- – initial vowels are represented by the voiceless laryngeal stop character አ with relevant vocalic modification. The voiceless laryngeal stop is represented by the same character in the middle of the word. For example: ኢአ *iiʔa* "hand";
- – the example above also shows that long vowel length is not marked. The same for consonantal gemination (see *hobatto*, below);
- – the implosive /ɗ/ is represented by a vocalized ዬ preceded by the "sixth order" of the same consonantal character. For example: ድዲላ *ɗiila* "flour". The same system is adopted for the other implosive /ɠ/. For example: ግጊናኖ *ɠinano* "mosquito";
- – as for the back fricatives, the system makes use of the Ethiopian characters that are used in Amharic even if they do not show difference in pronunciation. All of them represent indeed the voiceless laryngeal fricative /h/ in Amharic. Starting from /h/, this phoneme is represented by ሀ. For example: ሀባቶ *hobatto* "washing". The voiceless pharyngeal fricative is represented by ሃ. For example: ሃንቻ *ħanca* "tree". The voiceless uvular fricative /χ/ is represented by the other Amharic "h", i.e., ሐ. For example: ሐና *xoona* "sheep";
- – the uvular stop /q/ corresponds to the Amharic velar ejective ቀ. For example: ቆሮማ *qoroma* "rhinoceros";
- – the alveolar affricate /ts/ is indicated by the character of the alveolar ejective ጸ even if it is not ejective. For example: ጸናፉ *tsanafa* "six".

Here is an alphabetic list of Latin and Ethiopian characters extracted from the pages of the booklet:

a አ

b በ

c ቸ

c' ጬ

d	ደ
dh	ድ፡ደ
e	ኤ
f	ፈ
g	ገ
gh	ግገ
h	ሀ
hh	ኅ
i	ኢ
j	ጀ
k	ከ
l	ለ
m	መ
n	ነ
o	አ
q	ቀ
r	ረ
s	ሰ
sh	ሸ
t	ተ
ts	ጸ
u	ኡ
w	ወ
x	ሐ
y	የ
z	ዘ

4. CONCLUSION

The decisions that are relevant to what kind of writing system to adopt for traditionally unwritten languages such as Ts'amakko and Ongota depend on the purposes of the writing system. Scientific transcription tends to use IPA characters and diacritics. However, there is a more phonetic and a more phonemic transcription style, depending on the kind of document that is created. The phonetic one is very detailed, using a large array of IPA characters and diacritics since it has to reflect the spoken language in as many details as possible. It has been mentioned but not treated in the present paper. The phonemic transcription is more regular and readable, even if it tends to use IPA characters and diacritics. Readability is the main principle for the designing of orthographic systems for literacy and literary aims. We have seen that, except for the voiced pharyngeal fricative represented by <ʕ>, no IPA characters are used and no diacritics. Other solutions, mainly the use of double characters, are adopted. In two cases a common character representing a consonant not present in the phonemic inventory of the language has been used for something else. It is the case of Ts'amakko and Ongota <v> for the voiced pharyngeal fricative /ʕ/ in the orthographic systems described.

REFERENCES

Central Statistical Authority of Ethiopia. 1996. *The 1994 Population and Housing Census of Ethiopia: Results for Southern Nations, Nationalities and People's Region.* Addis Ababa: Central Statistical Authority.

Central Statistical Agency of Ethiopia. 2012. *The 2007 Population and Housing Census of Ethiopia, Statistical SNNPR.* Addis Ababa: Central Statistical Agency.

Fleming, Harold, C. et al. 1992/93. "Ongota or Birale: A moribund language of Gemu Gofa (Ethiopia)". *Journal of Afroasiatic Languages* 3(3), pp.181-225.

Hayward, Richard, J. 1989. "Comparative notes on the language of the S'aamakko". *Journal of Afroasiatic Linguistics* 1, pp.1-53.

Mettouchi, Amina, Vanhove, Martine & Caubet Dominique. 2015. *Corpus-based Studies of Lesser-described Languages: The CorpAfroAs Corpus of Spoken AfroAsiatic Languages.* Amsterdam: Benjamins.

Savà, Graziano & Tosco Mauro. 2000. "A sketch of Ongota, a dying language of Southwest Ethiopia". *Studies in African Linguistics* 29(2), pp.59-135.

Savà, Graziano. 2005. *A Grammar of Ts'amakko.* Cologne: Köppe.

Savà, Graziano. 2017. "An Ongota ethnotext on beehive". *Ethnorêma* 13 https://www.ethnorema.it/wp-content/uploads/13-09-In-altre-lingue.pdf

[Site consulted on 25/06/2021].

Savà, Graziano. 2020. "Case clitics in Ts'amakko (East Cushitic, Ethiopia)". *Essays on African languages and linguistics in honour of Maarten Mous.* Leiden: African Studies Centre, pp.159-173.