This book was conceived during the closing event of the DiM project, developed within the framework of the Erasmus plus KA204 - Strategic Partnerships for Adult Education programme. It is the fruit of a collective effort, aimed at grasping the complex and changeable set of phenomena connecting language and identity. Its fourteen chapters intend to offer food for thought on some of the currently most debated questions for linguists in the global village, and are divided into three thematic sections: 1) multilingualism, minority languages and the eternal dichotomy between orality and writing; 2) lexicography and L2 teaching; 3) the role of linguistics in particularly complex multilingual contexts. The book was published thanks to a grant obtained in 2018 by Regione Friuli Venezia Giulia.

Ilaria Micheli, coordinator of the scientific committee of the ATRA series, is Associate Professor in African Linguistics at the University of Trieste. Flavia Aiello is Associate Professor in Swahili at the University of Naples “L’Orientale”. Amelia Pensabene is a teacher of English at the CPIA Avellino, where she was coordinator for the DiM project. Maddalena Toscano is a former researcher and teacher of Swahili language at the University of Naples “L’Orientale”.

Language and Identity
Theories and experiences
in lexicography
and linguistic policies
in a global world
Edited by
Ilaria Micheli,
Flavia Aiello,
Maddalena Toscano,
Amelia Pensabene

Euro 18,00
This book has been edited and published in the framework of the RICFVG 2018 programme funded by Regione Autonoma Friuli Venezia Giulia

EUT Edizioni Università di Trieste, Trieste 2021.

ISBN 978-88-5511-268-0 (online)

EUT Edizioni Università di Trieste,
via Weiss 21 – 34128 Trieste
http://eut.units.it
https://www.facebook.com/EUTEdizioniUniversitàTrieste
Language and Identity
Theories and experiences
in lexicography
and linguistic policies
in a global world

Edited by
Ilaria Micheli
Flavia Aiello
Maddalena Toscano
Amelia Pensabene
# Table of contents

<table>
<thead>
<tr>
<th>Introduction - Ilaria Micheli</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PART 1: GENERAL ISSUES</strong></td>
<td>17</td>
</tr>
<tr>
<td>Linguistic identity in and out of Africa - Karsten Legère</td>
<td>18</td>
</tr>
<tr>
<td>Multilingualism in India and the significance of multilingual dictionaries - Suhnu Ram Sharma</td>
<td>35</td>
</tr>
<tr>
<td>The natural supremacy of spoken language. Orality and writing in Africa - Ilaria Micheli</td>
<td>43</td>
</tr>
<tr>
<td>From IPA to wildcards: a critical look at some African Latin orthographies - Mauro Tosco</td>
<td>56</td>
</tr>
<tr>
<td><strong>PART 2: LEXICOGRAPHY AND DIDACTICS OF L2</strong></td>
<td>85</td>
</tr>
<tr>
<td>L’italiano come seconda lingua nei CPIA (Centri Provinciali per l’istruzione degli Adulti): questioni e prospettive - Emilio Porcaro</td>
<td>86</td>
</tr>
</tbody>
</table>
A bottom-up experience: the DiM project
- Angela Mormone, Mariastella Battista, Amelia (Lia) Pensabene 98

Lexicography and language learning of Swahili L2 at UNIOR:
the Swahili-Italian online dictionary project - Flavia Aiello,
Maddalena Toscano, Rosanna Tramutoli 113

Unveiling oral and writing skills of low-literate learners
of L2 Italian: from research to teaching - Marta Maffia 133

PART 3: RELATIONSHIPS BETWEEN ORALITY AND WRITING
IN MULTILINGUAL CONTEXTS, LINGUISTIC POLICIES
AND CASE STUDIES 145

Farsi language and Iranian immigration in Europe
- Adriano Valerio Rossi 146

Lingua e storia in Africa: considerazioni sul caso del tigrino
(Eritrea ed Etiopia) - Gianfrancesco Lusini 157

Transcription and orthography in two endangered
languages of Ethiopia: Ts'amakko and Ongota - Graziano Savà 174

Wolof language and literature: an introduction - Emiliano Minerba 202

The future of minority languages in Nigeria - Gian Claudio Batic 217

Linguistic ambiguity in Timor-Leste: local languages
between pride and shame - Carolina Boldoni 231
From IPA to wildcards: A critical look at some African Latin orthographies

MAURO TOSCO
University of Turin

ABSTRACT

The article presents and discusses a few African Latin orthographies. The scope of the work is set out in section 1, while section 2 discusses a few orthographies featuring IPA symbols and diacritics. They were often the work of linguists and missionaries and were conceived for mother-tongue alphabetization and in order to translate and publish religious literature. They are scarcely useful in everyday casual writing, and especially so on a keyboard (where only a restricted set of symbols is to all practical purposes available). They are contrasted in section 3 with the use of digraphs and, most of all, with “wildcards:” symbols of the Latin, basic (unmodified) alphabet that are taken to use, often in an idiosyncratic manner, in order to represent phonemes that do not have a direct, built-in representation. The discussion is wrapped up in section 4, where the limits on the use of wildcards are evidenced and the practical limitations of many African orthographies reiterated.

KEYWORDS

Orthographies; Latin script; IPA; diacritics; digraphs.
1. A SHORT INTRODUCTION WITH MANY CAVEATS AND UNSOLICITED EXCUSES

Why are many African orthographies so bad? Is it because so many African languages are phonologically complex? Indeed, they are. But is this the only reason? And, for that matter, in what sense can we say that an orthography is “bad”?

This article will delve into an analysis along language-internal criteria, and the appropriateness of an orthography (which inversely correlates with deviation from a one-to-one correspondence between graphemes and phonemes; Sgall 1987) will not be called into question. Rather, attention will be focused on effectiveness (Cahill and Karan 2008; Cahill 2015). Granting that graphization has been, in Africa and elsewhere, a top-down process and that data on the approval of an orthography on the part of its real stakeholders – all of its potential users – are difficult to get, the analysis will concentrate on the internal characteristics of an orthography that are in all likelihood bound to facilitate or hamper its use.

Just as all languages are equal but some of them are more equal than others because they have been ausbauized (Tosco 2008) into written, official languages, also not all orthographies are equal: “big” languages may go along fine for centuries with awful, impractical and obsolete orthographies. Minority languages cannot: often their speakers are illiterate and must face the usual hurdles of acquiring literacy; in many cases reading and writing skills are offered to the minority language speakers in a locally or internationally big language. Both possibilities are widespread in Africa (and elsewhere). Or the minority language speakers already know the local majority language and its orthography and have thereby access to a sizable amount of material and information available in that language; the additional skills required in learning and using the minority language orthography are mostly justified in ideological and moral terms. Any material produced in the minority language will face very high hurdles in dissemination and will not be likely to be a winning competitor, in both qualitative and quantitative terms.

In Africa and elsewhere many orthographies were designed by foreign linguists, policy makers or missionaries – often with some input from local language consultants – and basically with the aim of publishing in the local language. That the publication itself is a Holy Book or the word of the leading party and its sacred leaders does not change much: the speakers themselves will mostly be readers – i.e., passive users. A technically very accurate orthography will certainly be easy to learn.

Not the same can be said about actively using the orthography – i.e., as writers, and especially so when handwriting is replaced by the use of keyboards (of typewriters in the past and of computers or cellphones nowadays).
It is certainly true that most of these orthographies have never been tested in extensive, daily usage by speakers, but this use has probably never even been considered.

In short, one feels that not all the needs of the potential users were taken into consideration: the degree to which an orthography will be user-friendly will be much different if reading or writing are taken into consideration.

As the title implies, orthographies based upon the Latin alphabet only will be presented and critically discussed, and actually a tiny minority of them. Specifically, I will not be concerned with:

- the vigorous use of the Arabic alphabet, not only in the past but still today (for which Mumin and Versteegh 2014 is not only a must, but also a fascinating reading);
- the use and present spread of indigenous scripts (*in primis* the Ethiopic syllabary in Ethiopia and Eritrea and the recently revived Tifinagh Berber alphasyllabary);
- finally – and regretfully (and simply out of lack of competence and data) – the rise and to a certain extent spread of many “new” alphabets of Africa – such as the Vai syllabary, the N’Ko alphabet or the Mandombe script. A good overview is provided in Kootz and Pasch (2010).

2. THE UNBEARABLE WEIGHT OF LINGUISTICS (AND LINGUISTS)

Both digraphs and IPA-based modifications of Latin letters are found in the orthography of Hausa (Chadic; ISO 639-3: hau¹), probably the African language with the highest number of native speakers. Digraphs are < sh >, and < ts > for /ʃ/ and /(t)s’/, respectively; IPA symbols are < ɓ >, < ɗ > for implosives and < ƙ > for ejective /k’/. Ejectivization is instead not marked in the case of the affricate /ts’/, while < ‘y > stands for a palatal glottal stop /ʔʲ/). Under representation is found in the case of the same symbol < r > used for both an alveolar trill (the latter often expressed by < ī > in linguistic works, with plain < r > being reserved to retroflex flap /ɭ/) and most of all in the case of vowel length, which goes unmarked. As in many, maybe most African orthographies, tones are left unmarked.

Not surprisingly, leaving aside the still widespread use of Arabic-based Ajami, everyday Hausa written in *boko* (/bōkõ/), i.e., the Latin alphabet, often disregards the “hooked” letters altogether and ‘one still encounters publications, including newspapers, where the plain letters are used’ (Jaggar 2001: 698): ‘the letters are printed without the hooks’ (Newman 2000: 726).

¹ The ISO 639-3 code is provided for all and only African languages after their first mention in the text.
Nor are these problems limited to languages spoken in former British colonies: a cursory look at a few major languages of West Africa, suffices to show the extension of the IPA symbol < ŋ > for a velar nasal – e.g., in both Wolof (Atlantic; wol) and Bambara (Mande; bam). The latter adds to its inventory of graphemes marked by IPA symbols < ŋ > for the palatal nasal as well as < ε > and < ο > for the open-mid vowels; the same array of graphemes is also used in Dyula (Mande; dyu), a major lingua franca of West Africa (whose use and potentialities also as a written language are discussed in Micheli forthcoming).

Other languages face more and more difficult problems in their graphization. In the case of many languages of South Sudan, troubles started at least in 1928, at the time of the Rejaf Language Conference (Tucker 1929), for which Abdelhay, Makoni and Makoni (2016) provide a useful overview. Their attention mostly goes to the ideological aspects of the conference, and their stark critiques – framed in the new orthodoxy of postmodernism and postcolonial studies – focus on the “Orientalist” attitude of the participants (where “Orientalist” is of course an abusive term) and the alleged invention of discrete ethnic groups and languages with the aim of reinforcing and maintaining colonial power. Still, their analysis of Dietrich Westermann’s (1875-1956) orthographic proposals has some merit.

Following his long research in Sudan and West Africa and serving as director of the International Institute of African Languages and Cultures (later the International African Institute) from 1926 until 1939, Westermann proposed a Practical Orthography of African Languages (1928, 1930). Proposed IPA symbols to be used in these “practical orthographies” include among others < ŋ >, < ʃ >, < ʒ >, < ɣ >. Further recommendations include the use of apostrophes for ejectives and/or implosives, of umlaut (diaeresis) for “central vowels.” Also, the notation of dental stops with the digraphs < dh >, < th > is recommended.

An early example of these orthographical choices is Heasty’s (1937) Shilluk dictionary, but the same solutions lie at the basis of the alphabets still used for many languages of South Sudan of different genetic affiliation and to some extent used in education (English only being the official language of independent South Sudan).

Among the main languages, Dinka and Nuer are cases in point: Dinka (din) and other West Nilotic languages have both modal and breathy vowels. Breathiness is marked in IPA by a subscript umlaut (diaeresis), as /α̂/, /ε̂/, etc. The IPA notation is apparently the source of the Dinka superscript umlaut above the vowel, as in < ä >, < ë >, etc.

Although a few digraphs are used for consonants, no generalization of such a solution has been attempted. In stops, a breathy release is marked in IPA by a raised symbol for a voiced glottal fricative, as in /bʱ/: one could imagine a transcription *< ah > for /ḁ/ (instead of < ä >), etc. As vowel length is not phonological, one could even use *< aa >. This leaves the problem of vowel quality: seven vowel qualities are phonemic in Dinka, with open-mid /ɛ/ and /ɔ/ opposed to close-mid /e/ and /o/. Accents are often used in European languages to the effect of marking openness, but other solutions are conceivable, such as digraphs (*< ae >?). The Dinka orthography simply keeps the IPA symbols, and not only sports < ɛ > and < ɔ >, but also breathy < ɛ̈ > and < ɔ̈ >.

Other IPA signs are found in the case of < γ >, < η > and < η >.

The Dinka orthography is interesting for the use of digraphs whose second element is consistently < h >: < dh >, < th > and < nh > for the dental counterparts /d̪/, /t̪/ and /n̪/ to alveolars /d/, /t/ and /n/.

At the same time, signs provided by the Latin alphabet and available in any standard keyboard but left unused in the Dinka orthography abound: < f >, < h >, < q >, < s >, < v >, < x >, and < z >. One can easily argue that at least < s > and < z > could be good solutions for the dental stops. As < ny > is used for the palatal nasal stop /ɲ/, the absence of the perhaps even more common digraph < ng > for the velar stop /ŋ/ is puzzling.
The same system is basically followed for Nuer (West Nilotic; nus), with the umlaut being replaced for breathiness by underscore:

---

3 This and all the following specimens of South Sudanese orthographies were obtained in Juba in 2013.
Breathiness is not phonological in Murle (Surmic; mur), and this entails the absence of umlaut and underscore; on the other hand, the same IPA signs seen above are again used for both vowels and consonants:
The use of diacritics and IPA signs is quite widespread, being found in many, if not most, orthographies of the area.

Another major language of South Sudan, Bari (East Nilotic; bfa and others), replaces most IPA signs used for Dinka and Nuer with diacritics, with the exception of <ŋ> for the velar nasal. The implosives are marked by a preceding apostrophe: <‘b>, <‘d>, <‘y>. Umlaut is preserved in the case of <ö> and a single digraph <ny> is found. No breathiness contrast operates in Bari.

In this regard, Owen’s (1908) Bari grammar, with its abundance of umlauts, accented letters and digraphs (among which... <ng>) but no special symbol, was certainly better.
Figure 5. An excerpt from a page in Bari (Jujumbu Kendya ko Bari Buk Tomusala 1999: 52)
Mödö (or Jur; Bongo-Bagirmi; bex) is another language of South Sudan. The orthography used in Perrson and Perrson’s (1991) dictionary and grammar resembles Bari in its use of < 'b > and < 'd >, to which < ‘j > for a palatal implosive is added. Again, < η > marks a velar nasal and it further appears in the digraphs < ηg >, < ηb > and < ηm > for prenasalized phonemes.

Umlaut is used for < ï >, and < ë >, and a special sign for < ɔ >.

The extensive use of IPA symbols seems restricted to languages of South Sudan for the historical reasons outlined earlier in this section; digraphs, accents and apostrophes are rather used elsewhere. The orthography proposed for Rendille (East Cushitic; rel) of Kenya makes wide use of digraphs but also of an apostrophe preceding the sign for the modal stop in < ‘b > and < ‘d > for the implosives, and also, strangely enough, in < ‘h > for the pharyngeal /h/. An apostrophe following a digraph is used in < ng’ > for the velar stop (following the orthography of Swahili). The other digraphs are < ch > for the modal affricate /ʧ/, < kh > for the velar fricative /χ/, and < ny > for the palatal nasal. Acute accents mark a high tone. This is all the more disconcerting since Rendille belongs with Somali to the same sub-subgroup of East Cushitic (according to current classifications, they make up, together with Boni, the eastern branch of Omo-Tana, itself a major branching of East Cushitic). As we shall see in more detail below, the orthography of Somali could have provided a solution for a few phonemes, such as < dh > instead of < ‘d >. Although in Somali the corresponding phoneme is postalveolar (/ɖ/) rather than implosive (/ɗ/), it could easily have been adopted and provide a model for * < bh > instead of < ‘b > for bilabial /ɓ/. While the use of the same pattern for * < hh > for the voiceless pharyngeal /ħ/ could have caused problems (in gemination), Somali offered an easy viable alternative in its use of < x >.

In Rendille, < x > is just one among a sizable number of unused signs of the Latin alphabet; the others are < c >, < p >, < q >, < w > and < z >.
Very similar is the alphabet devised by the same missionary body for Dhaasanac (East Cushitic; dsh). Here again we find no IPA symbols and the implosives are represented with an apostrophe preceding the sign for the
voiced stop; as in Rendille, acute accents (in Dhaasanac, on both moras of a long vowel) mark a high tone.

At the same time the alphabet has quite a few peculiarities: the digraph < dh > marks a laminal voiced fricative /ð/, and does not take into account that an alternative pronunciation with its apical counterpart /z/ is well attested in all positions (Tosco 2001: 19).

Figure 7. An excerpt from a page in Dhaasanac (War’gat Markoká 1997: 27)
Both diacritics and special symbols are found in the orthography of Gawwada (East Cushitic; gwd) proposed by the SIL International Literacy Department (with the additional complication that both the Latin alphabet and the Ethiopic syllabary are suggested). The apostrophe here follows the consonantal sign and marks an ejective (thus following the IPA conventions) in < c’>, < k’> and < t’>. IPA symbols are used for the pharyngeals: /ʕ/ and /ħ/. Apart from < sh > and < ny >, the list comprises the use of < h > as second element for the implosives: < bh > and < dh >, but with no value in the case of < ch > for the affricate /ʧ/ (as in English) and of < qh > simply for /q/. The last two also imply that no “bare” < c > nor < q > are used, as well as no < z > nor < v >.

Figure 8. A proposed Gawwada alphabet (https://www.alepeople.org/sites/www.alepeople.org/files/ALPHABET%20of%20%CA%95ALE.jpg)

It is noteworthy that many of these orthographies, and in particular those of Kenya and Ethiopia, completely disregard the practical alphabets designed from the seventies for languages which are structurally and phonologically similar: just as no use of the solutions devised for Somali is made for the very similar Rendille, no attention is paid in the case of Dhaasanac to the contemporary Latin orthographies of Ethiopia (where a majority of the Dhaasanac live).

It is to these innovative orthographies that we turn our attention in the next section.
3. EXAPTATION, OR: LEARNING TO USE WHAT YOU HAVE

3.1. PLAYING WITH WILDCARDS

The Horn of Africa is home to at least two success stories among African Latin-based orthographies: two official or national languages in their respective countries and with many million speakers as well as potential users of written texts: these are Somali and Oromo (both East Cushitic; som, orm). Somali came first, with its orthography officialized in 1972.

The long, troubled history of the graphization of Somali and of the Somali language policy has been told many times and is the subject of whole monographs: Caney (1984) mainly deals with linguistic issues—history of the orthography and corpus planning; Labahn (1982) with the orthography as well as language policy in general. Laitin (1977, 1992 – the latter within the larger African context) explores the political side (although overtly biased in favor of state interventionism and nation building; for a critique cf. Tosco 2014). Short historical overviews are provided in Tosco (2010, 2015).

Predictably, Somali uses the digraphs < sh > for /ʃ/ and < kh > for /χ/, as well as < dh > for a postalveolar /ɖ/. Uvular /q/ is marked by < q > and vowel length by redoubling the sign for the vowel. Pitch is not marked (as well as vowel backing/advancement). Glottal stop is only marked when not in word-initial position by an apostrophe: < ’ >.

The main problem was the absence of an established and practical way to mark the pharyngeals /ʕ/ and /ħ/.

The genial solution came with the use of unmodified Latin letters, namely < c > for voiced /ʕ/ and < x > for voiceless /ħ/ (< p >, < v >, and < z > remain unused).

Consciously or not, it was realised that any Latin alphabet is bound to have a few “wildcards”: symbols that simply come for free with the choice of using the Latin alphabet but have no clear phonemic value to start with, and are therefore available to get assigned, in principle, any value.

A few Latin letters are born as wildcards: e.g., < q > and < c > already in Latin marked allophones of /k/. In the conclusions we will argue that, nevertheless, < q > is worse than < c > as a wildcard. As for < x >, in Latin it was used since the beginning for the cluster /ks/ and has been put to many different uses in different orthographies around the world (its value as /ʃ/ in Maltese, Basque and many other languages is a major example).

Other letters become wildcards on a language-specific basis whenever a phoneme usually expressed by that letter does not exist. Of course, being language-specific, the value of a wildcard is also much “lighter” than the established value of another letter: it is therefore particularly prone to substitution.
When establishing an orthography for ‘Afar (Djibouti, Ethiopia, Eritrea; aar), the orthography of neighboring Somali was a possible choice. ‘Afar is spoken in Djibouti, Eritrea and Ethiopia, but in Djibouti only competition with Somali and the ideological need to obfuscate similarities led instead to an orthography where the peculiar choices of Somali were shuffled: pharyngeals < c > and < x > of Somali became < q > and < c >, while the digraph < dh > for the postalveolar /ɖ/ (a rather obvious choice) became /x/. As ‘Afar (or, in the new orthography, Qafar) has no uvular stop, no new symbol for /q/ was needed. Remarkably, all the other signs of Somali were kept.

The “Djibouti” orthography is used in the ‘Afar regional state of Ethiopia alongside the Ethiopian syllabary.

In Eritrea, after independence (1991; de jure 1993), the languages of Eritrea have been provided with a unified, national Latin orthography from which they depart only for phonemes peculiar to single languages (Semitic languages Tigrinya, Tigre and, of course, Arabic are written, respectively, in the Ethiopic syllabary and in Arabic script). For the ‘Afar minority of Eritrea and the very similar Saho (East Cushitic; ssy) the Somali choices of the 1970’s have been implemented (plausibly in order to sever the links with the ‘Afar in other countries). In the end, three nation states have implemented two different orthographies for similar languages, with one and the same language (‘Afar becoming either Qafar in Djibouti or Cafar in Eritrea) having two different orthographies in different countries (three counting the Ethiopic syllabary).

<table>
<thead>
<tr>
<th>phoneme</th>
<th>Somali</th>
<th>‘Afar (Djibouti)</th>
<th>‘Afar, Saho (Eritrea)</th>
<th>gloss of examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>c</td>
<td>cad</td>
<td>q</td>
<td>qado</td>
</tr>
<tr>
<td>h</td>
<td>x</td>
<td>xaakin</td>
<td>c</td>
<td>caakim</td>
</tr>
<tr>
<td>d</td>
<td>dh</td>
<td>dhal</td>
<td>x</td>
<td>xale</td>
</tr>
</tbody>
</table>

Figure 9. Somali, ‘Afar and Saho: wildcards and political choices (adapted from Savà and Tosco 2008: 125)

The principle of using wildcards has been further implemented in Ethiopia in connection with the marking of ejectives in Oromo. A brilliant combination of the criteria of simplicity and frequency is used – supplemented in case by adherence to tradition. The overall picture is apparently puzzling but makes actually good sense:
Figure 10. Modal and ejectives in the Oromo orthography

For the rarely used (mostly in loans) bilabials, plain $< p >$ stands for modal /p/ and the digraph $< \text{ph} >$ for its ejective counterpart /p’/: an additional phonetic feature is paralleled by an additional graphic symbol. For the velars, $< k >$ stands for the modal and $< q >$ for the ejective (here following a long Orientalist and Ethiopianist tradition). Wildcards are instead used in alveolar stops and alveopalatal affricates. For the former, $< t >$ stands for /t/ and $< x >$ for ejective /t’/. For affricates, where $< c >$ is a wildcard and the digraph $< \text{ch} >$ a well-established solution for /ʧ/, frequency decides, and while $< \text{ch} >$ is reserved to modal (and less common in Oromo) /ʧ/, simple $< c >$ stands for its ejective counterpart /ʧ’/.  

The web of motivations at play here is certainly complex; the present writer remembers that, when presented and discussed at the (first) International Symposium on Cushitic and Omotic Languages (Köln, 1986), “some modifications were recommended on the basis of phonetic consistency” (Heine 1988: 620). These recommendations – supported by most scholars (and a very young and naïve writer of these lines) – consisted basically in proposing the use of digraphs with $< \text{h} >$ as second element for the ejectives.  

Wisely, the Oromo did not pay attention to intellectuals and “experts” and stuck to their decisions. Since then, the Oromo solution has been highly influential in Ethiopia and has been followed in recent years by other Latin-based orthographies (cf. Savà and Tosco 2008). One could even say that the use of $< x >$ for /t’/ has become a shibboleth of the new Ethiopian alphabets for a geographically and genetically diverse array of languages. While only Koorete (North Omotic; kqy) and Sidamo (East Cushitic; sid) are presented here, the examples could be multiplied.

<table>
<thead>
<tr>
<th>modal IPA</th>
<th>IPA orthography</th>
<th>ejective IPA</th>
<th>IPA orthography</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p/</td>
<td>$&lt; p &gt;$</td>
<td>/p’/</td>
<td>$&lt; \text{ph} &gt;$</td>
</tr>
<tr>
<td>/t/</td>
<td>$&lt; t &gt;$</td>
<td>/t’/</td>
<td>$&lt; x &gt;$</td>
</tr>
<tr>
<td>/k/</td>
<td>$&lt; k &gt;$</td>
<td>/k’/</td>
<td>$&lt; q &gt;$</td>
</tr>
<tr>
<td>/ʧ/</td>
<td>$&lt; \text{ch} &gt;$</td>
<td>/ʧ’/</td>
<td>$&lt; c &gt;$</td>
</tr>
</tbody>
</table>
Figure 11. An excerpt from a page in Koorete (Koorete Erunxi Pishsharo 1992: 58)
When you accept the principle of wildcards, there is no need to stick to the Oromo solution: $< x >$ again, but in this case for the voiced uvular fricative /χ/, has recently been proposed by SIL for Ts’amakko (East Cushitic, Ethiopia; tsb; closely related to Gawwada). Savà (this volume) has taken this proposal over in his proposed orthography of Ongota (unclassified, Ethiopia; bxe).
3.2. BREAKING AWAY FROM TRADITION

Other orthographic uses seem to point in the same direction: a progressive liberation from the bounds imposed by traditional (European) orthographic norms.

Digraphs are traditionally treated as combination of two letters and they are alphabetized under the first element only. Thus, *church* is listed in English dictionary after *campaign* but before *cut*, and the Somali-Italian dictionary (DSI 1985) follows this principle, with, e.g., *shabeel* ‘leopard’ after *saddex* ‘three’ but before *sug* ‘to wait.’

As the number of digraphs and wildcards as well as the use of a Latin orthography increase, the weight of tradition decreases. This is when a digraph becomes a “letter;” an autonomous, single grapheme. Already in the Somali-English dictionary (Zorc 1993) all words beginning with *< dh >*, *< kh >* and *< sh >* (the only digraphs of the Somali orthography) are listed separately, but still after their first element: *< dh >* after *< d >*, *< kh >* after *< k >*, and *< sh >* after *< s >*.

Moreover, in a first stage the first element only is doubled in gemination; in Somali, e.g., *< ddh >*, rather than *< dhdh >*, stands for /ɖɖ/, as in *gabaddha* ‘the girl’ (more commonly actually spelled *gabadha*). Once perceived as single letters, each of the elements of a digraph are instead doubled in gemination, yielding, e.g., Wolaytta *geeshsha* ‘clean, pure’.

The next step follows logically: if, e.g., *< sh >* is no longer *< s >* + *< h >* but a brand-new autonomous symbol, the order of digraphs in the alphabet can and must change. This is what happens in recent Ethiopian dictionaries of languages using the Latin orthography, with the digraphs increasingly found all together at the end of the list, as in the Wolaytta dictionary (Tophphiya 1991), with *< ch >*, *< ph >*, and *< sh >* following in this order after *< z >*, and in Oromo (Mekuria 1998) with *< ch >*, *< dh >*, *< ny >* and *< sh >* (no word begins with *< ph >*).

Still, in capitalization the first letter only is capitalized: tradition is adhered to when it implies a simpler solution.

Even a completely different tradition in listing letters may now be accommodated. The following Table is the alphabetical chart present in a Koorete primer. Following the pattern of the Ethiopic syllabary, vowels are listed as columns and consonants as rows, for a total in Koorete of ten vowels (five short and five long) and thirty consonants. Apart from this general graphical arrangement, the order of consonants is the traditional Western one (but the very last consonant is *< th >*). The digraphs are particularly abundant in Koorete: *< ch >*, *< dh >*, *< jh >*, *< ny >*, *< ph >*, *< sh >*, *< xh >*, *< dz >* and *< th >*. They are listed after (and under) their first element (*< ch >* after *< c >*, *< dh >* after *< d >*, and so on), rather than all together at the end.
The order of vowels, too, follows the traditional Western one, with each long vowel after its short counterpart.

Figure 13. A Koorete alphabetical chart (Koorete Bidzunxo Suma Erunxi Pishsharo 1990: 86-87)
3.3. WILDCARDS GALORE

But why to stop at “natural” wildcards? Any unused symbol may come to good use.

We go here beyond the notion that Latin symbols for consonant clusters (as in the case of < x >) or of original allophones (< c >) are “free”: any sign that happens to be useless in the language is up for grabbing and re-use.

Nara (Nilo-Saharan or isolate; nrb) of Eritrea offers maybe the most radical solution so far, and puts into relief as well a few problems.

Most solutions found in the Latin orthography of the languages of Eritrea are not peculiar at all, others partially are: e.g., a palatal nasal is < gn > in Saho, as in Italian, rather than < ny >. Again, shunning the Oromo (and therefore, in a way, the Ethiopian solution), an ejective /t'/ is not marked by < x > but by < th > in Bilin (Central Cushitic; byn) and Saho (East Cushitic; syy), while < ch > marks an ejective palato-alveolar affricate /ʧ'/ – it was seen above that in Oromo it marks the modal. It was also seen above that Saho < c > marks the voiced pharyngeal /ʕ/ (à la Somali); a diacritic is therefore introduced for the modal affricate /ʧ/: < č >.

The velar nasal is of course a problem for any Latin-based alphabet: its most common rendering is < ng >, but many other solutions have been or are in use, such as Swahili and Xhosa (both Bantu; swa and xho) < ng’ >, < nh > in Galician (Western Romance) and Nawat (or Pipil; Aztecan), or simply < g >, as in Fijian (Austronesian).

Although absent in word-initial position the velar nasal is also phonemic in Piedmontese (Western Romance), where it is also probably more common than the alveolar nasal. The orthographic solution devised for Piedmontese is to use < n > where no ambiguity may arise and have a hyphen follow it in other cases (i.e., between vowels) yielding < n- >. This of course conflicts with hyphenation, but is consonant with the liberal use of hyphens in other points of the orthography (such as in order to separate clitics, following the French model).

Nara is most illuminating in its use of < v > for the velar nasal. Certainly, the presence of prenasalized voiced stops preempted the use of < ng > – as this digraph was chosen, quite correctly, to represent a prenasalized /ŋg/. Still, < v > is a brave choice, and to the best of my knowledge unique. In their proposed orthography for Ts’amakko of Ethiopia, SIL has used likewise < v >, but for a voiced pharyngeal fricative /ʕ/, and Savà (this volume) proposes to copy this in Ongota.
Furthermore: why to stop at letters? In Ethiopia, Wolaytta (North Omotic; wal) has introduced the digit < 7 > for the glottal stop, as in lee7iyaa ‘thin.’ It is most commonly found reduplicated, as in ha77i ‘now’. As elsewhere, the phonological presence of a glottal stop is not marked in word-initial position.

This solution is not totally isolated, as it is also found in Squamish (Coast Salish) of British Columbia (whether it was consciously copied from Squamish is unknown to the present writer). It is still apparently isolated in Ethiopia, where the apostrophe < ’ > is preferred. The two solutions are shown here through the initial page of the Book of Hosea in Wolaytta (Hosee7a) and Oromo (Hose’aa).

---

4 Examples are from the Wolaytta-Amharic dictionary (Tophphiyaa… 1991). The English translations are the most common meanings of the Amharic entries.
< 7 > is certainly more conspicuous – and therefore less likely to be forgotten in casual writing – than the apostrophe. It is obvious that language-internal considerations – such as frequency and relevance in morphological processes – will have to be taken into account in the selection.

Figure 15. Hosea 1:10-11 in the Wolaytta Bible (Geeshsha Maxaafaa 1996: 881)
Figure 16. Hosea 1 in the Oromo Bible (Macaafa Qulquulluu: 1118)
By definition, to use a basic Latin keyboard only excludes graphic iconization (Sebba 2015), i.e., no “icon” similar to Danish < ø > or Spanish < ñ > may arise (although the use of < 7 > in Wolaytta could come very close to be an “icon”). What is possible instead is the idiosyncratic association of a grapheme to a phoneme, as repeatedly shown in this article, and with Somali < c > and Oromo < x > being maybe the most striking cases.

Second, the case of ‘Afar (Qafar) has shown that it is difficult to outsmart wildcards: in comparison to < c >, < q > is much less of a wildcard, and it has a strong association with a uvular stop or it simply marks a back allophone of /k/ (as it was in Latin and is still generally the case in modern European languages). < c >, on the contrary, can retain its Latin value as a velar stop (as in Romance languages with a non-front vowel following) and have different values with other vowels, or still be an alveolar affricate (as in Croatian, Slovenian, Polish, Czech, Slovak, Hungarian…) or many other things. And it can of course become much else in digraphs or with diacritics. Neither a place nor a manner of articulation is strictly linked to it, synchronically.

The fact that no language, to the best of my knowledge, has followed ‘Afar in using < q > for a pharyngeal fricative or has given it yet another value than /q/ is certainly due to the imperfect ausbauization of ‘Afar: in Djibouti, French and Arabic are the only official languages, and both Somali and ‘Afar are conspicuous for their absence from the linguistic landscape. But I venture to say that there is something inherently awkward in making a wildcard out of a card that is not. In exaptation you turn into use what you have and is available; it is certainly more difficult (but not impossible) to change the use of a more or less functional tool.

Third, it is also apparent that wildcards are second bests. This strategy seems to be always secondary to the use of digraphs; e.g., /ʃ/ is always expressed by < sh > and an implosive /d̪/ or postalveolar /dʱ/ is generally < dh >.

The use of digraphs is particularly shunned for vowels, except in the marking of length and notwithstanding the wide use of vocalic digraphs in many European writings.

Finally, maybe the most important – and saddest – conclusion is that many African orthographies were in a way born old: they are utterly incapable of being brought to use in the most modern technologies – pending financial investments (in developing, implementing and marketing keyboards) that the communities cannot sustain.
REFERENCES


1 The publication year found in Ethiopian books is often in accordance with the local calendar (3nkutataš). It is followed here by the corresponding year in the Gregorian calendar.


Tosco, Mauro. 2015. “Short notes on Somali previous scripts” in: Cabdirashiid M. Ismaaciil, Cabdalla C. Mansuur & Saynab A. Sharci


