Compiling of a Berber-French Dictionary (Figuig dialect)

Mohamed Yeou

Department of English & LERIC, Université Chouaib Doukkali
m_yeou@yahoo.com

The paper describes the compilation of a bilingual dictionary Berber (Figuig)-French both in paper and in electronic version. The dictionary is root-based and refers to dialectal forms for comparison. The purpose of the dictionary is to contribute to the documentation of Figuig Berber in order to provide a linguistic resource for the Figuigui community and scholars interested in researching the Berber language.

1. Introduction

1.1. Figuig

The Berber variety documented in this dictionary is spoken in Figuig and belongs to the Zenati branch of the Berber language family. Figuig is a oasis situated in the South East edge of Morocco, around 1000 km from Casablanca, and 460 km from the Mediterranean coast. The number of permanent residents is today around 15 000. Precise estimates of the language speakers is difficult to calculate because much of the population have emigrated away to Europe and to major Moroccan cities. The language has recently been listed in the UNESCO Atlas of the world's endangered languages. There is no doubt that such listing is justifiable given the following facts:

- the social and cultural context of the language has undergone great changes, resulting in lexical attrition;
- the language is used mostly by the parental generation and up;
- the limited numbers of L1 resident speakers, who are constantly subject to emigration;
many speakers have negative attitudes towards the language and think their children would be better served by speaking other languages.

The present project of compiling a Berber-French dictionary is partly motivated by the critical need for language resource materials on which language revitalization and language standardization depend.

1.2. Review of literature

There are four major studies on the Berber of Figuig, as well as a few shorter ones. The earliest source is a small glossary of 31 pages by Basset (1885). However, Basset’s work should be read with much caution. No more major fieldwork was conducted on Figuig Berber until 1994 and 1995, when Marteen Kossman and Fouad Saa defended their theses, respectively. Saa (1995) studies some aspects of the phonology and verbal morphology of Figuig Berber based on the generative framework. The thesis appendix is very interesting as it lists the verbal paradigm of 1296 verbs, along with their derived forms. Kossman’s thesis, which was published as a book in 1997, gives a general description of the grammar of Figuig and provides a 144-page Berber-French glossary in the annex. This grammar is an excellent descriptive analysis and the glossary is very helpful. The final two major original sources are Ben-Abbas (2003) and Sahli (2008). Ben-Abbas (2003) investigates the sociolinguistic aspects of word-borrowing between languages in contact, mainly Arabic and French, while Sahli (2008) gives a brief grammar sketch of the language, together with a Berber-Arabic glossary. The glossary consists of a list of 2250 words without context, exclusively from the dialect spoken in Ksar Laâbidate. Another important study is a collection of folktales by Ben-Amara (2007). Ben-Amara transcribes an interesting number of tales from Figuig, but does not translate them. He gives, however, a list of the words used in these tales with their French gloss.

2. Printed Version of the dictionary

2.1. Compiling of the dictionary database

The dictionary database is compiled using Toolbox, a program produced by SIL International (formerly the Summer Institute of Linguistics). Toolbox uses MDF standard (Coward and Grimes, 2000) for lexicon structure and converts certain files in Standard Format into RTF (to be further processed and printed with MS-Word). Toolbox provides field lexicographers with the ability to integrate various types of data: lexical, grammatical, semantic, etc. It has many options for selecting, sorting, and displaying data. It is very useful for helping researchers
generate a reversed finder list as well as analyze and interlinearize text corpora. A sample of the printed output for a formatted dictionary and a reversed finder list are found in the appendix.

The typical database entry has the following fields:

- \lx Lexeme (is the abstract consonantal root)
- \se Subentry (derived word)
- \wa Variant form
- \vn Variant comment (shows source of variant: name of Ksar or name of author)
- \ps Part of speech
- \sn Sense number
- \rn Reversal (this gives the French word(s) or phrase(s) desired for a reversed French-Berber finder list)
- \dn French definition
- \sc Scientific name (two-part name of a species, especially for plants)
- \ng Grammatical information (mainly for the different verb stems)
- \sy Synonym
- \an Antonym
- \cf Cross-reference (general purpose cross-reference)
- \xv Example in Berber
- \xn Translated example in French
- \sd Semantic domain (for entering semantic fields)
- \nt General Notes (dialectal forms and language name from which the word is borrowed; dialectal forms come from published Berber dictionaries and glossaries)
- \wr See (this is a field which I added to cross-reference a variant item to a main entry where fuller information is found)
A sample entry which shows how the information in database is organized (in Toolbox format) is given in Figure 1. The \lx field shows the main form of the lexeme, which is actually the consonantal root, and \ps gives the part of speech. \se shows the derived words. The \xv field gives the definition of the word in French. \xv, \xn pairs give example sentences and their French translations. \va lists variant transcription of the item, and \vn shows which source the variant came from (name of Ksar). The \nt field lists attested Amazigh dialectal forms.
2.2. Amazigh Lat Keyboard

The transcription of Berber in the dictionary is based on a standardized Latin transliteration system, as used by Berberists. However some of the symbols can be different from what is commonly used, as shown in (1):

<table>
<thead>
<tr>
<th>Our system</th>
<th>Other systems</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>ə</td>
<td>ə</td>
</tr>
<tr>
<td>š</td>
<td>c</td>
<td>ʃ</td>
</tr>
<tr>
<td>ź</td>
<td>j</td>
<td>ʒ</td>
</tr>
<tr>
<td>tš</td>
<td>č</td>
<td>tʃ</td>
</tr>
<tr>
<td>dž</td>
<td>ɡ</td>
<td>tʒ</td>
</tr>
<tr>
<td>v</td>
<td>ġ</td>
<td>v</td>
</tr>
<tr>
<td>x</td>
<td>ḥ</td>
<td>χ</td>
</tr>
<tr>
<td>ḍ</td>
<td>d / t</td>
<td>t</td>
</tr>
<tr>
<td>t</td>
<td>t / tʰ</td>
<td>tʰ</td>
</tr>
</tbody>
</table>

To enter text in Berber, I do not use Windows tools such as Character map because they can be very cumbersome. Instead, I created a keyboard layout designed to my own specification. The virtual keyboard, called Amazigh Lat, is made for regular writing and is compatible with Unicode fonts, such as the most recent versions of Times New Roman (from v. 5.0) and Arial (both supplied in Windows Vista and Windows 7). Amazigh Lat keyboard is used with Tavultesoft Keyman Desktop program, a utility for managing keyboard input methods. Both the program and the keyboard can be downloaded at [http://www.tavultesoft.com](http://www.tavultesoft.com).

2.3. Transcription approach

The transcription adopted for lexical words in the dictionary is phonetic (broad) rather than phonological. The intent is to show phonetic aspects that characterize Figuig variety and compare it to other Amazigh varieties. Apart from spirantization, devoicing of some geminates is an important feature that is noted at the phonetic level. The voiced geminates /bb, dd, ḍd, gg, ggʷ/ are realized as [pp, ɬɬ, tt, kk, kkʷ] both at the lexical and morpho-phonemic level. The voiceless non aspirated geminate [ɬɬ] (IPA [tt]) is indicated through the use of the IPA diacritic for devoicing, the under-ring, to distinguish it from the voiceless aspirated geminate [tt] (IPA [ttʰ]) (Yeou et al., 2011).

The phonetic approach is also motivated by the intent to record dialectal variation inside Figuig Berber itself. For example, the data below shows that
absence / presence of pharyngealization or aspiration in an important indicator of dialectal variation:

(2)  

<table>
<thead>
<tr>
<th>Upper Figuig</th>
<th>Zenaga</th>
</tr>
</thead>
<tbody>
<tr>
<td>ttrid</td>
<td>ṭṛрид</td>
</tr>
<tr>
<td>ḏḍir</td>
<td>deyyer</td>
</tr>
<tr>
<td>iv zer</td>
<td>ḳvr</td>
</tr>
<tr>
<td>tšar</td>
<td>ṭšar</td>
</tr>
<tr>
<td>ḏḍer</td>
<td>tṭer</td>
</tr>
<tr>
<td>tter</td>
<td>tṭer</td>
</tr>
</tbody>
</table>

“thin pancakes”  
“foot of mountain”  
“river, wadi”  
“fill (up), be filled (up)”  
“live”  
“ask for charity”

As far as the representation of morpheme boundaries is concerned, I adopt a ‘syntactic’ approach based on the criterion of syntactic mobility (see Stroomer, 1994). Hence spaces will preferably used rather than hyphens to mark morphological segmentation. For example, transcription type (a) will be adopted rather than type (b):

(3)  

1a. Inna yas sad isek tid ḏart nnes  “He told him that he will build his house”  
1b. Inna-yas sad isek tid ḏart-nnes  
2a. ppas d mmi ṭrahén ḳres  “His father and his son went to see him/her”  
2b. ppa-s d mmi-s ṭrahén ḳ r-es

As regards assimilation processes, I generally adopt a phonological approach:

- The subjunctive and future marker ad is always transcribed as /ad/, given that its assimilation to the following consonant is predictable: ad ṭv er [at ṭv er] “that she will study”, ad ṭv er [an ṭv er] “that we will study”
- The prefix /ṭ-/, which assimilates in voicing to the following consonant, is also transcribed phonologically: tezdev [dezdev] “she lived”, težžey [deţţey] “she recovered”

However, a phonetic approach was preferred in the case where there might be some dialectal variation, namely for the feminine morpheme suffix /-t/, the causative prefix /s(s)-/, and the intensive form prefix /tt-/:
2.4. The layout of the entries

The present dictionary uses a root-based approach, even if this approach has its practical weaknesses. This approach was partly motivated by the desire to serve the needs and interests of the academic community of linguists interested in researching the Amazigh language, or related languages.

The layout of the entries is organized as follows:

- Roots in bold capitals are arranged in the following alphabetic order: B, D, Ḍ, F, G, H, I, K, Kʷ, L, M, N, Y, R, S, Ş, Š, T, Ṭ, W, X, Y, Z, Ž, Ė. Roots starting with Ḍ, R and Ž are, however, listed with D, R and Z, respectively.

- The items in a root entry are grouped according to their semantic relation.

- For each root entry, simple verbs are listed first, and then the derived verbs with the following prefixes: /s-/, /m-/, /ttw-/, for the causative, the reciprocal and the passive, respectively. After that, verbal nouns for both simple and derived verbs are given, and finally, nouns and adjectives.

- For each verb, the first line gives the aorist as the basic form, followed by the other forms, mainly the preterite, the negative preterite, the intensive, and the negative intensive. The slash separates variant forms if there are any.

- For each noun or adjective we list the marker of the construct state between parentheses: (u-), (w-), etc.

2.5. The strengths of the present dictionary

2.5.1. Exhaustiveness and volume

This dictionary project aims to develop a comprehensive dictionary of Figuig Berber with French translation and extensive dialectal cross-references. The present dictionary tries to avoid the following weaknesses of previous work on the lexis of Figuig: (1) word-for-word translation; (2) lack of real uttered sentences; (3)
limitation to one community dialect; (4) limitation to the literal dimension of meaning.

In this project, I will bring together not only an extensive compilation of words in Figuig with authentic sample sentences and their French translations, but also figurative and idiomatic uses of some of these words. In addition to that, some patterns of expression like proverbs, riddles and excerpts from songs and tales are included, because they reflect the culture of the Figuigui community more than every other kind of linguistic unit.

The Amazigh Figuig variety is characterized by some minor dialectal variation due to the fact that Figuig comprises seven separate ksours or communities situated on two levels: The upper level consists of Laâbidate (At nnež), Loudaghir (At ɛdджi), Oulad Slimane (At slimane), Hamam Tahtani (At waǰďay), Hamam Foukani (At Ʌamer) and El Maïz (At lemɅiz), and the lower level consists of Zenaga (Iznayen). There is complete mutual intelligibility across the communities, and the small variation that exists will be noted in the dictionary. The default dialect is upper Figuig, namely Loudaghir, but variant forms are listed in the variant field (\va).

2.5.2. Dialectal cross references

At the bottom of each lexical entry, the dictionary lists attested forms from Amazigh languages or dialects given in Table (6), and whose references are given in the bibliography. It also indicates if the meaning is different or similar to that of Figuig. If the word is borrowed, etymological information about the source language and the original form is given.

(7)

<table>
<thead>
<tr>
<th>Mc</th>
<th>Tamazight of central Morocco</th>
<th>Zen</th>
<th>Zenaga of Mauritania</th>
</tr>
</thead>
<tbody>
<tr>
<td>McS</td>
<td>Tamazight of south central</td>
<td>Aha</td>
<td>Tahaggart</td>
</tr>
<tr>
<td></td>
<td>Morocco</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chl</td>
<td>Tachelhit</td>
<td>Nig</td>
<td>Tamajeq (Tawellemmet, Tayert)</td>
</tr>
<tr>
<td>Rif</td>
<td>Tarifit</td>
<td>Mal</td>
<td>Tamasheq of Mali</td>
</tr>
<tr>
<td>Izn</td>
<td>Beni Iznassen</td>
<td>Nef</td>
<td>Nafusi</td>
</tr>
<tr>
<td>Sen</td>
<td>Senhaja de Srair</td>
<td>Ght</td>
<td>Ghat</td>
</tr>
<tr>
<td>Ntf</td>
<td>Ntifa</td>
<td>Ghad</td>
<td>Ghadamès</td>
</tr>
</tbody>
</table>
### 3. Electronic Version of the dictionary

To generate the electronic version of the database, Lexique Pro was used. Lexique Pro is a free program developed by the Summer Institute of Linguistics (SIL). It transforms data from a Toolbox and formats it in a dynamic viewer.

![Figure 2: snapshot of the electronic version of the dictionary](image)

<table>
<thead>
<tr>
<th>Kab</th>
<th>Kabyle</th>
<th>Snd</th>
<th>Sened</th>
</tr>
</thead>
<tbody>
<tr>
<td>Che</td>
<td>Chenoua</td>
<td>Djr</td>
<td>Djerba</td>
</tr>
<tr>
<td>Cha</td>
<td>Tachawit or Chaoui</td>
<td>Chn</td>
<td>Chenini</td>
</tr>
<tr>
<td>Sns</td>
<td>Beni Snous</td>
<td>Dw</td>
<td>Douiret</td>
</tr>
<tr>
<td>Ace</td>
<td>Central Algeria</td>
<td>Ght</td>
<td>Ghat</td>
</tr>
<tr>
<td>Mzb</td>
<td>Tumzabt of Mzab</td>
<td>Ghad</td>
<td>Ghadamès</td>
</tr>
<tr>
<td>Oua</td>
<td>Tagargrent of Ouargla</td>
<td>Siw</td>
<td>Siwa</td>
</tr>
<tr>
<td>Tim</td>
<td>Gourara, Touat, Tidikelt</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The generated interactive dictionary has the advantage of displaying the database in a user-friendly format. The user can navigate by clicking on hyperlinks to related entries such as synonyms, variants, cross-references, semantic categories. Figure 2 shows a snapshot of a lexical entry from the electronic version of the dictionary. On the left we can either access the entries, the dictionary roots, by clicking on “Figuig”, or access the reversed French finder list by clicking on “French”. Both the roots and the French words are organized in alphabetical order. Moreover, we can also view entries by semantic domain if we click on the Category button. Work on thematic categorization is ongoing and only the following semantic fields have been entered so far:

- Kinship terms
- Animals
- Body Parts
- Food & Drink
- Clothing
- Agriculture & Vegetation
- Religion & Beliefs

The advantage of the electronic version generated by Lexique Pro is that it can be distributed as a packaged setup program and can also be exported to a set of web pages, in text, html, or xml formats.

4. Conclusion

The dictionary has been on the making for several years. We hope that it might be completed in approximately six months. It is expected that the dictionary will have roughly eight thousand words when completed.

5. Bibliography


Amaniss, A. (non publié). *Dictionnaire tamazight-français*.


Appendix: example of dictionary page and finder list

DN

**adan** (w-) n. ♦ intestins (en général), le petit intestin. ◊ **adanen n tmurt** litt. intestins de terre: lombric, ver de terre. cf. 'adan n tmuat' (Rif): m. sn. Cf.: *tameswadant, √MSWDN; imedren, √DRN; tasuft, √SF. Pl.: adanen. dial & étym.: adan, adan (Chl), adan (Mc, Izn, Rif, Sen, Cha, Che, Ace, Sns, Mzb, Oua, Ghd, Ght, Nef), adan, aden (Snd), ădân (Aha): boyau, intestins comme sn. commun; dér. du v. eden (Aha): graisser, être graissé.

tadunt (dd-/td-) n. ♦ graisse (d'origine animale). ◊ **mi dd un (ul) tšēh ddunt ukk ῦušš iqqar (iqqar) as yexx / mikk un (ul) tɛrxis ddunt ukk ῦušš iqqar (iqqar) ammu tufh** litt. quand le chat n'arrive pas à avoir la graisse il lui dit "pouah!" / qu'est ce qu'elle pue!: se dit pour qqn. qui minimise l'importance d'une chose souhaitable, mais qu'il est incapable de réaliser (dicton). syn.: *tilebdin, LBḌ. dial & étym.:* tadunt (Chl, Ntf, Sen, Mc, Izn, Rif, Cha, Sns, Mzb, Oua, Nef, Tim, Snd), tadunt, tadwent (Ace), tâdent (Aha), tadent, tedent (Nig): m. sn.

adun n. ♦ augment. du précédent.

DN

**aden** yuden, un (ul) yudin, ittaden, un (ul) ittiden v. ♦ couvrir, recouvrir, couvrir de couverture (dormeur); être couvert ◊ **adn it ammen ad ıtţeš** couvre-le d'une couverture pour qu'il dorme. ◊ **adfel qa ıllı yuden adıyra** la neige a complètement couvert la montagne. ◊ **tella tuden imma nnes an (al) ixef nnes** elle s'est couverte jusqu'à la tête. ♦ fig. couvrir, protéger, chercher à innocenter (coupable, accusé). ◊ **itekk ẓzbayel ppas ittaden xifes** il fait des fautes graves, mais son père ne le dénonce pas pour le protéger. *dial & étym.*: aden (Mzb, Oua, Izn, Ace, Sns, Tim, Snd, Ghd, Nef): couvrir et/ou pass.

**ttwaden** ittwaden, un (ul) ittwaden, ittwadan, un (ul) ittwidin v. ♦ être couvert, recouvert ◊ **tettwaden tmurt s wedfel (udfel)** la terre a été couverte de neige. ♦ fig. être protégé, innocenté, couvert

**idan** (y-) n. vb. ♦ fait de couvrir, de recouvrir, de couvrir de couverture: ♦ fait de chercher à innocenter.

madun (u-) n. ♦ plaque, dalle de pierre; ♦ dalle de tombeau. ◊ **iemand wala imudan** litt. il vaut mieux [s'appuyer sur] une canne qu'une dalle de tombeau: plutôt souffrir que mourir (dicton) syn.: *tadelха, √PḶ. Pl.: imadunen*. dial & étym.: madun, tmadun (Mzb), teddenen (Zen): mm. sn.

**tmadunt** n. ♦ couvercle, tout ce qui couvre; ♦ bouchon. ◊ **tmadunt n uqilil** couvercle de cruche, d'une théière. ◊ **tmadunt n uv ellay** couvercle de bouilloire. ◊ **tmadunt n qer ԑ et** bouchon de bouteille. ◊ **tmadunt n  trebuie** couvercle de marmite. ◊ **tmadunt n qer ԑ et** bouchon de bouteille. ◊ **tmadunt n  trebuie** couvercle de marmite. ◊ **tad elh a** litt. il vaut mieux [s'appuyer sur] une canne qu'une dalle de tombeau: plutôt souffrir que mourir (dicton) syn.: *tad elh a* litt. il vaut mieux [s'appuyer sur] une canne qu'une dalle de tombeau: plutôt souffrir que mourir (dicton) syn.: *tadelха, √PḶ. Pl.: imadunen*. dial & étym.: madun, tmadun (Mzb), teddenen (Zen): mm. sn.

~ 150 ~
dent $n.$ tīy mest (te-), $ÝMS.$
dent (de clé) $n.$ tiswet (te-), $SW.$
dent (de fourche) $n.$ qaššaw (u-), $QŠ(W).$
depuis prép. $sì$, $S.$
dernier $n.$ adj. anekkaru (u-), $KR.$
dernier-né $n.$ adj. aμażuẓ (u-), $MZ.$
derrière prép. adv. $n.$ deffer, $DFR.$
descendre v. hwa, $HW$;
v. zder, $ZDR.$
descente $n.$ asehnennay (u-), $HNY$;
$v.$ vb. hekku (u-), $HW.$
dessous $n.$ aḏḏay (w-), $DY.$
dessèchement $n.$ vb. asiqqur (u-), $ŶR.$
dessécher v. ssiqqur, $ŶR.$
dessécher (se) v. aḏen, $DN$;
v. hšišer, $HŠR.$
desséché $n.$ adj. maḏun (u-), $DN$;
$adj.$ $n.$ aμeqqur (u-), $ŶR.$
destinée $n.$ mimun (u-), $MN.$
dette $n.$ ameŗwāṣ (u-), $RŴS.$
deuil $n.$ anebdi (u-), $NBD.$
deux (fém.) $n.$ adj. sent, $SN.$
deux (masc.) $n.$ adj. senn, $SN.$
devancer $n.$ asul, $SL$;
v. zzar, $ZR.$
devant prép. adv. $n.$ zzat, $ZT.$
devenir v. dwel, $DWL$;
v. ḏha