The Typology of Number Borrowing in Berber

Lameen Souag

SOAS

The question of which numbers are most easily borrowed, and in which contexts, has implications for an understanding both of historical change and language contact and of the extent to which the linguistic behaviour of numbers can be related to independent cognitive factors. In the Berber languages of North Africa, numbers are commonly clear-cut loanwords from Arabic; some languages retain as few as one or two non-Arabic number words, while others preserve a complete inventory. Closer examination reveals differences in intensity of borrowing even within single languages, depending on the numbers' functional usage. The languages in question are closely related to one another and are all influenced by mutually comprehensible varieties of Maghrebi Arabic, allowing what amounts to a controlled experiment, with extremely similar contact situations in different areas yielding a wide spectrum of possible outcomes. Careful examination of this spectrum allows us to set up a typology of numeral borrowing in Arabic-Berber contact, showing how linguistic, social, and cognitive factors all affect the process of number borrowing and how synonymy may emerge as an unstable transitional stage in the adoption of a new system.

1 INTRODUCTION

The key question in studying language contact is: what factors affect the borrowing of a given feature? The greater the differences between the situations compared, the harder it is to isolate the relevant factors. Berber offers a particularly promising opportunity to disentangle this issue; it is spoken in numerous relatively isolated enclaves across North Africa, mostly in fairly similar sociolinguistic situations and in contact with the same languages. While numbers tend to receive relatively little attention from linguists, this particular semantic domain has some notable advantages in investigating contact. Numbers are readily borrowed in Berber, and are much better documented than most features of interest in the language; moreover, the Arabic and Berber number systems are sufficiently different to ensure no difficulty in identifying the loanwords.

1.1 Current situation

Berber is a family of closely related languages indigenous to North Africa, touching the Mediterranean and Atlantic to the north and west, and spoken as far east as Siwa in Egypt and as far south as northern Burkina Faso (with emigrant groups even further afield.). The term “Tamazight”, the traditional autonym of a number of Berber languages, is increasingly used as an alternative. The largest Berber languages by population are, from west to east: Tashelhiyt/Shilha (south Morocco); Middle Atlas Tamazight (central Morocco); Tarifit/Rifi (north Morocco); Taqbaylit/Kabyle (northeast Algeria); Tashawit/Chaouia (northeast Algeria). While these are spoken across reasonably large, densely populated, continuous areas, many varieties are restricted to a handful of villages (eg Ghomara in north Morocco) or a single oasis (eg Siwa in Egypt), often with little contact with other Berber speakers. The mutual intelligibility of Berber varieties varies substantially, making a division into languages difficult in practice; the whole family could be seen as consisting of two more or less broken up dialect continua, one in the North and one in the South, with a few more divergent outliers around the edges.
Almost every Berber language is surrounded by colloquial Maghreb Arabic speakers on all sides, and is spoken in a state whose official language is Modern Standard Arabic, and in which the ex-colonial language (usually French) remains significant in official domains. The languages of the Tuareg (a sparse, partly nomadic population spread across a vast expanse of the Sahara) have come under much less Arabic influence than others; along with Zenaga (the nearly extinct Berber language of Mauritania), they have also been influenced by sub-Saharan African languages.

1.2 History

Berber is a subgroup of Afro-Asiatic; as such, it is distantly related to Arabic (and other Semitic languages of the Middle East), Egyptian, Somali (and other Cushitic or Omotic languages of East Africa), and Hausa (and other Chadic languages of West Africa.) It was already spoken in North Africa before the Roman conquest, as the Tifinagh inscriptions of the Numidian kings attest.

In the 7th century, the Arab Umayyad Empire conquered most of North Africa. While the area resumed independence within a century or so, the results of this conquest were lasting; most northern Berbers converted to Islam, and Arabic became an important part of city life, widely used in government and trade. In the 11th century, large Arab tribes (in particular the Banu Hilal) immigrated en masse from Arabia via Egypt, leading to the collapse of state authority in much of North Africa and to the presence of large rural Arabic-speaking groups. This seems to have been a turning point in the Arabization of North Africa; with Arabic both useful on a local scale and prestigious on a broader scale, many Berber groups gradually shifted to Arabic. Some have done so within living memory; for example, the village of Sened in Tunisia was still largely Berber-speaking at the beginning of the 20th century, but is now entirely Arabic-speaking. The long-standing influence of Arabic is reflected in most Berber languages' tendency to use Arabic numbers.

2 Sample

To establish a database of patterns in Berber number borrowing, I examined a total of 53 source-dialect pairs, with dates ranging from 1883 to 2007. Some were by native speakers, some by linguists, some by colonial administrators; but one advantage of examining this topic is that even amateur non-speakers can usually be trusted to get cardinal numbers right.

2.1 Sample size

Giving an exact count of languages is less than meaningful when many of them seem to fall into dialect continua, but if the classification of the Ethnologue (2005) is followed, my sample includes 19 out of 26 languages. Of the missing ones, Guanche is excluded deliberately: it has been extinct since the 1600s, the scanty materials surviving are hard to interpret, its classification is not entirely certain, and its speakers were not in regular contact with Arabic speakers. No adequate data on the number system appears to have been published for Ghomara or Awjila, and none was found for Sokna, Temacine, or Judeo-Berber. In short, this sample is nearly complete. However, the Ethnologue count is problematic in that some varieties - eg those of the ksours of southwestern Algeria, or southern Tunisia - are classed by the Ethnologue as part of a neighbouring language despite showing rather different characteristics, while at least one, Judeo-Berber, is split despite being mutually comprehensible with Tamazight (Galand-Pernet et al. 1970:14.)
A less ambiguous measure of sample diversity is the number of geographically separated enclaves examined - 25. Even this, however, runs into problems. The enclave with the largest population - central/south Morocco - displays several distinct patterns of number borrowing, suggesting that an appropriate unit in this region would be smaller.

2.2 Results

The results may be summarised as follows. (See appendix for fuller details.) Words for “million” and above are excluded, since, where attested, these tend to be French or other European loanwords (often via Arabic). “Langs” counts the number of Ethnologue languages in the type, and “Sources” the number of source-dialect pairs.

<table>
<thead>
<tr>
<th>Type</th>
<th>Berber</th>
<th>Arabic</th>
<th>Sources</th>
<th>Langs</th>
<th>Enclaves</th>
</tr>
</thead>
<tbody>
<tr>
<td>R*</td>
<td>all attested</td>
<td>none</td>
<td>10</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Transitional</td>
<td>1 to ≥10,000</td>
<td>≥ 1000</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Transitional</td>
<td>1 to 29</td>
<td>≥ 3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Transitional</td>
<td>1 to ≥50</td>
<td>≥ 21</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Transitional</td>
<td>1 to 99</td>
<td>≥ 20</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Transitional</td>
<td>1 to ≥30</td>
<td>≥ 11</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Transitional</td>
<td>1 to 10</td>
<td>≥ 2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>R2</td>
<td>1 to 2</td>
<td>≥ 3</td>
<td>21</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>R3</td>
<td>1 to 3</td>
<td>≥ 4</td>
<td>8</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>R1</td>
<td>1 only</td>
<td>≥ 2</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Mixed</td>
<td>1-6, 8-10</td>
<td>7 (&gt;10?)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mixed trans.</td>
<td>1 to 19, partly 21-99</td>
<td>≥ 10*</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mixed trans.</td>
<td>1 to 19, partly 21-99</td>
<td>≥ 20*</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1 - Typology of Number Borrowing in Berber

The meaning of the type labels will be clear; R* refers to languages (such as Tamashek) which have retained all Berber numbers, Rn to those which consistently use Berber below n and Arabic numbers above it. In all but a few “mixed” cases, all cardinal numbers above some point are Arabic, and all below some point are Berber. One of the exceptions - Jerbi - is based on inadequate data from a single informant, while the other two (dialects of Tashelhiyt) appear to be rare cases where units, rather than numbers, have been borrowed; in these, the number 20 has been borrowed from Arabic and is used in forming higher numbers vigesimally, using Berber words for the coefficients. However, in a number of rather diverse “transitional” cases (including the two Tashelhiyt varieties), some numbers can be expressed in either Arabic or Berber.

2.3 Geographical distribution

R* is concentrated in the south (mainly Tuareg and Zenaga), and R2 is spread across most of the north. R3 is restricted to central Morocco and the Gourara oases of SW Algeria (formerly ruled by central Moroccan tribes), while R1 is restricted to Tarifit in Northern Morocco and nearby enclaves of what amounts to the same language. Transitional/mixed types are found between them, in southern Morocco, in the Mzab and Ouargla in central Algeria, and on Jerba in southern Tunisia. In short: in the South, where Arabic influence is least, the original Berber numbers are retained. In the North, where Arabic influence is greatest, the Arabic system almost entirely replaces Berber.
3 IMPLICATIONS

The sample is consistent with three typological claims, true of Berber and likely to be true elsewhere:

a) 1 is always retained. (Greenberg 1978:290 notes that he knows of no exceptions to this; in the few cases of which I am aware, such as Japanese and Korean, in which a word for “one” has been borrowed, it coexists with a native word for “one”.)

b) 2, and then 3, are far more often retained than any higher numbers.

c) If $n$ is retained, so are all numbers $< n$ (cp. Greenberg 1978:289's near-universal 54: “If an atomic numeral expression is borrowed from one language into another, all higher atomic expressions are borrowed.”)

3.1 Cognitive explanations

A number of motivations may be found for these empirical rules. The universal retention of ‘one’ may be related to the fact that it serves not only as a number but as a determiner across most Berber languages (as in Maghreb Arabic), eg:

(1) *haeret iyen*
    thing one
    ‘a (certain) thing’ (Tamashek - Heath 2005:250)

This second function gives it a role outside the numeric system proper, and renders it considerably more frequent than would otherwise have been the case, discouraging replacement.

The second generalisation - that 2 and 3 are much more widely retained than higher numbers - is paralleled by several other typological facts. Dual numbers are well-attested cross-linguistically, and trial numbers, if rare, are found in a number of languages; yet higher numbers, where found, are always vague, referring not to exactly $n$ items but to an inexact large or small number of items (Corbett 2000.) The highest number found in non-productive number systems - ones with no productive means of forming higher numbers - is never greater than 5, and is commonly 2 or 3 (Greenberg 1978:256.) A cognitive motivation may plausibly be sought for these typological facts. A variety of independent facts suggest that small numerals, especially 2 and 3, are perceived rather differently than other numerals. 2 and 3 can be subitised - perceived without counting - by the typical person (Hurford 1987.) Some aspects of the concepts of “2” and “3” are acquired much earlier than other numbers - one-week-old newborns distinguish 2 dots from 3 dots, but not 4 from 6 (Antell & Keating 1983.) The mean short-term memory limit is 3-5 items (Cowan 2001), implying that 4 or more items will be outside the short term memory of many people some of the time; only 3 or fewer items can be relied upon to fit in short-term memory.

3.2 Sociolinguistic explanations

One of the most interesting aspects of this typology is that - while in the South we see the original system and in the North we see the results of a far-reaching shift - in much of the middle, in a zone stretching from southern Morocco to northern Libya, we can witness the shift still underway. This is shown by direct evidence for Tumzabt (central Algerian Sahara), Ghadames and Nafusi (NW Libya), where (respectively) sources from 1898/1904/1898 show more Berber numbers than ones from 1969/1973/1942. Indirect evidence is available for Tashelhiyt (south Morocco) and Teggargarent (central Algerian Sahara), and shows some very interesting points about how the change takes place. In both the latter, the choice of
Arabic or Berber numbers varies partly with the speaker's gender. For the Tashelhiyt dialect of Tazerwalt, Stumme tells us:

Also, for the numbers 3-29 one frequently chooses the Arabic terms (ie. those in §171.) The women and small children of the Tazērwalt-Shlûh by preference count (as far as possible) with the Berber numbers, the men by preference (from 10 up) with the Arabic ones. Therefore the Shlûh call the Berber numbers lahsâb ntîmgârin, and the Arabic ones lahsâb niirgâzên - ie women's counting vs. men's counting. (Stumme 1899:102)

For Teggargarent, the language of Ouargla in Algeria, early reports likewise indicate that ‘[chared, Berber for '3'] is scarcely used except by women, men usually using tlata [Arabic for '3']’ (Biarnay 1925:188.) This matches very well with what would be expected from consideration of North African social norms. Men travel more freely than women, normally do their household's shopping, and are much more likely to engage in trade. They thus have more occasions to speak to Arabic monolinguals, and more incentive to use Arabic numbers; it is to be expected that, as observed in these two cases, they will switch sooner than women.

In Teggargarent religious affiliation is also a factor. Some speakers are Maliki Muslims, along with most Algerians - including their Arabic-speaking neighbours. Others are Ibadî Muslims, belonging to a small, closely-knit religious minority with a traditional tendency to avoid close relations with non-members. Most Algerian Ibâdis speak Tumzâbî, the Berber language of the nearby Mzab region; and Tumzâbî has retained Berber numbers up to at least 20, so one might expect the Ibâdi Ouarglis to use more Berber numbers than their Maliki neighbours. Sure enough, forms above 3 are used mainly by Ibâdis, according to Delheure (1987). Curiously, Biarnay (1925) reported that Teggargarent used only Arabic numbers above three. Either his informants on this matter were Malikis and the situation was much the same then as in 1987, or easier contact with the Mzab in the twentieth century has led to a resurgence of previously disused Berber numbers among the Ibâdis.

4 INTRA-LINGUISTIC FUNCTIONAL VARIATION

A closer look at the languages of the north reveals some interesting uses of Arabic numbers in places where Berber ones would be expected. In many R2/R3 languages, Arabic numbers have replaced even the lowest numbers in certain specific contexts. The idea of time-telling using hours and minutes is a comparatively recent introduction to most North Africans, and has typically reached Berber-speaking regions through the largely Arabic-speaking cities; perhaps as a result, most Berber languages use Arabic numerals throughout in telling the time, irrespective of how the corresponding numbers would otherwise be expressed. Compare Tarîfît (Ouallah, p.c.):

(2)   rwâhda   vs.  iḍœn
      “one o'clock”   “one”

to Maghreb Arabic:

(3)   əl-wâhâd-a    vs.   wahâd
      DEF-one-F   one
      “one o'clock”   “one”

To a lesser degree, the same is often true of higher units of time. In Maghrebi Arabic itself, such units are normally counted using a slightly different system than the general
number system: the dual suffix -ayn is used rather than the number zudʒ “two”, and shortened forms of the numbers 3-10 are used (thus əl əθ, ṣəbʕə-, xəms-... instead of the full forms əl-aθa, ṣəbʕa, xəmsa...) For example, “two days” is yumayn, and “three days” əl-iyyam, whereas “two pens” is zudʒ stiluwat (never *stiluwin), and “three pens” ələθa stiluwat. These forms cannot be broken up into morphemes capable of standing alone (there is no word -ayn or əl əθ -); this has no doubt contributed to their commonly being borrowed as whole units into Berber, rather than having their component morphemes translated. For example, in Siwi (Laoust 1932) we see:

(4) iyumain vs. sən
   “two days” “two”
   təlt iyyam vs tlata
   “three days” “three”

Another domain in which Arabic numbers are, less commonly, reported to take over is sequential counting. In Ayt Ayache Tamazight ‘The Arabic numbers 1-3 are used only for counting in order without naming things... The Berber numerals are used elsewhere’. (Abdel-Massih 1971:22.)

Several easterly Berber languages have adopted the Arabic ordinal system for two and above, even while retaining a Berber cardinal number for “two”; the inherited Berber pattern is to form ordinals (other than suppletive “first”) with a relative pronoun plus the number, which burdens the lexicon considerably less but leads to longer words. Compare the Nafusi system (Beguinot 1942:127-129, notation preserved):

(5) úğun, sen, tláta, árb`a...
    one, two, three, four...
    amezwár, ettâni, ettâlet, errâba`...
    first, second, third, fourth
to the Maghreb Arabic one (Dellys dialect):

(6) wahd, zudʒ, ələθa, rəb`a...
    one, two, three, four...
    əl-luwwal, əθ-əθani, əθ-əθaləθ, ər-rab`a`...
    the first, the second, the third, the fourth

Unlike cardinal numbers, these specialised uses of numbers are not covered in sufficient numbers of sources to permit a full typological survey of borrowing patterns in them. It is to be hoped that with new fieldwork a more complete picture of the variation across Berber in these domains will emerge. The available information is enough to indicate that synonymous Arabic and Berber numbers may acquire different uses, rather than being mutually substitutable.

5 CONCLUSIONS

Number borrowing is affected by social factors, such as the degree and domain of contact; by cognitive factors, such as the possibility of subitisation and the processes involved; and by linguistic factors, such as the other functions of the number and the transparency of words
containing it. Synonymy is possible in number systems, and appears to be a diachronically unstable transitional stage in the adoption of a new system.

REFERENCES


**APPENDIX: SAMPLE DETAILS**


R1: Senhaji (Ibañez 1959), Tarifit (Sarrionandia 1905:114ff, Ouallala p.c.)


**Mixed**: "Nafusi" (Jerba: Basset 1883:308)

Lameen Souag

Linguistics Department, School of Oriental and African Studies
Thornhaugh St
Russell Square
London
WC1H 0XG
United Kingdom

lameen@gmail.com

http://lameen.googlepages.com/