

LESSICO NATURALISTICO SOMALO

Introduzione

Annarita Puglielli e Giampaolo Barosso

I. Origine del progetto

Il progetto lessicale sulla lingua somala – una parte del progetto Studi Somali iniziato nel 1977-78 e che aveva come obiettivo una descrizione sistematica di questa lingua a tutti i livelli – ha prodotto in prima istanza il Dizionario Somalo-Italiano, pubblicato nel 1985. I circa 35000 lemmi che ne costituiscono la base lessicale sono diventati l'input iniziale per la compilazione del Grande Dizionario Somalo, tuttavia avevamo la consapevolezza della necessità di approfondimento di specifiche aree lessicali. A tale scopo si sono messe a punto strategie diverse per raccogliere dati relativi ai diversi settori (cfr. Puglielli A., 'The Somali Lexical Project – Aims and Methods' in M.Putz, ed., *Discrimination through language in Africa*, Berlin –N.Y., Mouton de Gruyter, 1995).

II. Il lessico naturalistico somalo (LNS)

Il Lessico Naturalistico Somalo che qui presentiamo, riguarda in particolare due aree: la botanica e la zoologia, aree su cui si è cominciato a lavorare nella seconda metà degli anni '80.

La ricerca sul mondo naturale somalo prevedeva innanzitutto una ricerca bibliografica, poi un primo spoglio delle fonti bibliografiche esistenti e la creazione di un database con i termini somali trovati nella letteratura. Successivamente si sarebbe fatto un intervento sul territorio da parte di linguisti, botanici ed esperti di zoologia che avrebbero dovuto realizzare la seconda parte della ricerca mirata al controllo dei termini somali raccolti dalle fonti bibliografiche e all'individuazione delle categorie conoscitive e classificatorie del mondo naturale utilizzate dalla comunità linguistica somala.

Ciò che presentiamo, includendolo nella Collezione Inediti dell'Archivio Somalia, è il risultato della prima fase del lavoro, un risultato quindi non completo, in quanto a causa dello scoppio della guerra civile è stato impossibile realizzare la seconda fase del lavoro. I dati in esso contenuti possono rappresentare un utile strumento per chi volesse occuparsi di questi ambiti.

Tutto il lavoro relativo al Lessico Naturalistico Somalo che qui presentiamo è stato svolto da Giampaolo Barosso, uno studioso eclettico e di grande capacità e raffinatezza intellettuale che lo ha realizzato a partire dalla bibliografia fino alla costruzione del database. La sua scomparsa non ha reso possibile inserire in questa presentazione un suo scritto di commento assai interessante e il cui testo purtroppo non abbiamo in forma completa.

L'evoluzione in campo informatico, ha poi fatto sì che il lavoro originario abbia dovuto essere 'ricconvertito' in modo da renderlo fruibile con i pc e i sistemi attuali.

Il LSN si compone di:

- a) Bibliografia
- b) DB botanica
- c) DB zoologia

Bibliografia

La bibliografia – unica per i due settori – contiene 212 titoli e copre un arco di circa un secolo. I testi identificati non tutti sono stati reperibili e consultati, ma naturalmente nei DB è citata la fonte da cui i termini sono stati rilevati. I lavori prodotti prima della scrittura ufficiale della lingua somala (1972) ovviamente non seguono le norme della lingua standard, né era prevista una loro normalizzazione in questa prima fase del lavoro; la decisione sulla scrittura da utilizzare nel vocabolario si sarebbe presa dopo la verifica dei dati sul campo. La conseguenza di quanto detto è che lo stesso termine presente in fonti diverse può apparire con grafia diversa e quindi essere registrato nel DB più di una volta, tuttavia l'informazione relativa alle classificazioni scientifiche ecc. consente al lettore di poter utilizzare l'informazione. Altra osservazione da fare è che potrebbe essere problematica anche l'attendibilità del dato linguistico dal momento che non abbiamo alcuna notizia sulla tecnica utilizzata per l'elicitazione dei dati linguistici e sulle fonti dei dati stessi.

Il DB Botanica

I dati relativi alla botanica sono registrati in records con la seguente struttura:

Nome Somalo; Nome Scientifico; Famiglia, Sottofamiglia, Ordine, Riferimento (RIF), Note.

Il campo RIF contiene i codici numerici corrispondenti alle fonti bibliografiche utilizzate. In moltissimi casi un singolo nome scientifico è registrato in più records, tanti quanti sono i lemmi somali a cui esso corrisponde. La decisione di registrare tanti records per ciascun nome somalo per una stessa specie è stata presa in considerazione della impossibilità - qualora qualcuno volesse farlo - di procedere a elaborazioni, ordinamenti, confronti, nel caso i lemmi somali fossero stati presenti in sequenza in un unico campo. Della varietà di grafie con cui i nomi somali sono registrati abbiamo già detto, sono quelle trovate nelle fonti, né si riteneva di poter prediligere una grafia sulle altre senza un opportuno ricontrollo dei dati sul campo.

Il DB zoologia

I dati che si riferiscono alla zoologia sono ripartiti come segue:

Mammiferi	(ZMAM)
Uccelli	(ZUCC)
Animali domestici	(ZDOM)
Rettili e anfibi	(ZR&A)
Invertebrati	(ZINV)
Pesci	(ZPES).

La struttura dei records è la seguente:

Nome Somalo, Italiano, Altre lingue, Identific., Nome Scientifico, Sottofamiglia, Famiglia, Ordine, Classe, Sottotipo, Tipo, Riferimento, Note.

Il campo Identificatore è reso necessario per l'impossibilità di servirsi uniformemente di uno stesso termine di riferimento. Alcune fonti, infatti, presentano solo il riferimento per la famiglia o l'ordine; per animali come i domestici, inoltre, è assente un riferimento sistematico. Il campo prevede quindi la segnalazione di un identificatore che per priorità è l'identificatore scientifico seguito da altre possibilità come genere, famiglia scientifica, o, nell'eventualità, nome italiano.

Per i gruppi Classe, Sottotipo e Tipo sono state utilizzate le seguenti abbreviazioni:

Classe	Sottotipo	Tipo
Ma = Mammiferi	V = Vertebrati	C = Cordati
U = Uccelli		
R = Rettili		
Anf = Anfibi		
Pes = Pesci		

LNS nell'Archivio Somalia : indicazioni per la consultazione degli allegati

- I. L'Introduzione e le indicazioni per la consultazione.
- II. La bibliografia che include sia botanica che zoologia.
- III. I dati della Botanica e della Zoologia in archivi separati in formato Excel: si tratta di liste sempre ordinate alfabeticamente per nome somalo.
- IV. Un DB Access interrogabile sempre per tutte e due le aree.
- V. Cartella ZIP per la consultazione offline dell'intero DB tramite un'interfaccia HTML semplificata.

La presenza dei diversi allegati è finalizzata a dare la possibilità di accedere al LNS al più ampio numero di utenti, questo significa che gli stessi dati e gli stessi contenuti sono forniti in diversi formati di lettura e consultazione.

Di seguito alcune indicazioni pratiche per la consultazione del materiale relativo al database sopra elencato:

- La bibliografia è allegata in un duplice formato: PDF, consultabile scaricando gratuitamente il programma Adobe Acrobat Reader, e Microsoft Excel, consultabile soltanto se si è in possesso dell'applicativo Microsoft Office. Sia all'interno del file PDF che del file Excel è possibile effettuare delle ricerche.
- I dati della Botanica e della Zoologia sono anch'essi messi a disposizione nei due formati PDF e Microsoft Excel, come accade per la bibliografia.
- Il DB Access contiene gli stessi dati contenuti nei file citati al punto 3, si differenzia invece per il formato, in questo caso un database Microsoft Access, consultabile soltanto se si è in possesso dell'applicativo.
- Per utilizzare il DB offline in formato HTML, è necessario scaricare sul proprio pc la cartella ZIP ed estrarne i contenuti. Successivamente, occorre aprire la cartella relativa e fare doppio clic sul file index.html, si accederà così alla Homepage tramite cui consultare il database. In questo formato, la sezione di Zoologia è suddivisa nelle seguenti sottocategorie: Mammiferi, Invertebrati, Pesci, Rettili e anfibi, Uccelli; mentre la sezione di Botanica è unica ordinata alfabeticamente. Una volta effettuata la scelta all'interno della sezione, si accede agli elenchi e ai dati presenti. La colonna della Bibliografia riporta dei simboli di libri sui quali è possibile cliccare per aprire il riferimento bibliografico in una piccola finestra pop-up.

Abstract

The materials we present here are part of the Lexical Research Project aimed at the production of Somali dictionaries started in 1977-78 as part of the Somali Studies Project - a cooperation project financed by the Ministry of Foreign Affairs between The Somali National University and the Università di Roma 'La Sapienza'. For a description of its aims and methods see Puglielli A., *The Somali Lexical Project. Aims and methods*, in M. Putz (ed.), *Discrimination through Language in Africa*, Berlin-N.Y., Mouton De Gruyter, 1995.

In particular this work was meant to collect terms from the fields of botany and zoology. The work started with a bibliographical research and a scrutiny of the Somali terms referring to plants and animals found in the literature. In a successive phase the lexical items identified would have to be checked on the field. This second phase was never reached because of the civil war.

The decision to put the collected materials in the Archive was taken because they could represent useful information for people interested in these fields.

The SNL is constituted by:

- a) Bibliography
- b) Botany Database
- c) Zoology Database.

All the work was carried out by Giampaolo Barosso, an eclectic researcher, with fine intellectual capacity, who realized it from the bibliographical research to the construction of the Databases.

Bibliography

It consists of 212 items that cover 100 years. The texts from which the terms are taken are cited as sources. The terms are cited as found in the sources and have not been normalized according to the official writing system for Somali introduced in 1972. The standardization was planned after the field work and the control of all the terms to be included. As a consequence, the same term found in different sources may appear with a different spelling, and therefore they are recorded as different entries in the DBs.

Botany DB

The records have the following structure:

Somali name

Scientific name

Family

Subfamily

Order

Reference

Notes

Zoology DB

The data are divided into the following categories: Mammals (Ma), Birds (U), Domestic Animals (DOM), Reptiles (R), Amphibious (Anf), Fishes (Pes).

The records have the following structure:

Somali Name

Italian name

Other Languages name

Identifier

Scientific Name

Family

Sub-family

Order

Class Type

Sub-type

Reference

Notes

The SNL is included in the Archive – within the ‘Unpublished documents’ Collection - with a single bibliographic record to which the following material is attached:

- 1.** Introduction and practical advices for consultation.
- 2.** The bibliography which includes both botany and zoology.
- 3.** The data of Botany and Zoology databases in separate archives in Excel format: these lists are sorted alphabetically by Somali name.
- 4.** A Microsoft Access DB including both areas.
- 5.** A ZIP folder to consult the entire DB offline through a simplified HTML interface.

The presence of the various attachments is aimed at ensuring access to the SNL to the largest number of users, this means that the same data and the same contents are provided in different formats.

Here are some **practical indications** to consult the above-mentioned material related to the database:

- The bibliography is attached in a double format: PDF, which can be consulted by downloading the Adobe Acrobat Reader software for free, and Microsoft Excel, which can only be consulted if you have the Microsoft Office application. The user can perform searches both in PDF and Excel files.
- Botany and Zoology databases data are also made available in both PDF and Microsoft Excel format, as is the case for the bibliography.
- The Access DB contains the same data of the files mentioned in point 3, but it differs in the format, in this case a Microsoft Access database, which can only be consulted if you have the application.
- To use the DB offline in HTML format, you need to download the ZIP folder on your PC and extract its contents. Subsequently, you need to open the folder and double-click on the index.html file, you will then access the Homepage through which to consult the database. In this format, the Zoology section is divided into the following sub-categories: Mammals, Invertebrates, Fish, Reptiles and amphibians, Birds; while the Botany section is in alphabetical order. Once you have selected the section, you access the related lists and data. The Bibliography column shows the icon of one or more books, by clicking on the icons the bibliographic reference will open in a small pop-up window.

APPENDICE

Includiamo in appendice, allo scopo di dare alle persone che consulteranno questo database un'idea del lavoro ad esso sotteso, alcune note e commenti di Giampaolo Barosso sul Lessico Naturalistico Somalo di botanica e zoologia.

Note e commenti di G. Barosso

Botanica

1. La struttura dei records di BTN è questa:

Campo	Nome campo	Tipo	Dim
1	NOMESOMALO	Carattere	40
2	NOMESCIENT	Carattere	60
3	FAMIGLIA	Carattere	34
4	RIF	Carattere	40
5	NOTE	Memo	10
Totale:			165

2. Il numero dei records è attualmente (26/04/89) di 6.113 - pari al numero di occorrenze (NON singole) di nomi scientifici [genere & specie (varietà/sottospecie)].
3. In moltissimi casi un singolo nome scientifico è registrato in più records, tanti quanti sono i "nomi somali" che ad esso corrispondono. Il numero dei SINGOLI NOMI SCIENTIFICI è di 3.148 - pari quindi al numero di singole specie [varietà/sottospecie] registrate (meno i pochi casi in cui il "nome somalo" corrisponde a una specie non identificata o a più specie di un certo genere, e dove il nome scientifico ha perciò la forma "<genere> sp." o "<genere> spp.").
4. Il numero di SINGOLE FAMIGLIE è di 137. Per circa 120 generi (186 specie) l'indicazione relativa alla famiglia manca (sono i casi in cui non ho potuto sopperire alla mancanza d'indicazione sul testo-fonte).
5. Il numero complessivo (occorrenze NON singole) di "nomi somali" è di 4.524. Il numero di nomi scientifici ai quali non corrisponde alcun "nome somalo" è quindi di 1.589.
6. Il numero di "NOMI SOMALI" SINGOLI è di 3.733.
7. Definire "nomi somali" le grafie registrate sotto questa voce mi sembra del tutto improprio (perciò ho messo il termine tra virgolette); direi che al momento debbano essere considerate niente di più che mere grafie sulla carta - alle quali, così a occhio, credo sarà difficile attribuire una maggiore significatività; ma per stabilirlo occorrono naturalmente un esame e considerazioni meno "a occhio".
8. DIMENSIONI DELLA STRUTTURA. Rispetto all'impostazione iniziale - all'incirca uguale alla versione iniziale della struttura relativa alle voci zoologiche - il numero dei campi che

compongono l'attuale struttura dei record è molto ridotto (v. osservazioni sulle dimensioni della struttura nelle 'Note' relative alla zoologia).

9. L'eliminazione del campo 'Varianti' - adottando invece il criterio di compilare un record per ogni singolo "nome somalo" - è dovuta a ragioni sia (a) di sostanza sia (b) di metodo, e pratiche.
10. La suddivisione dei dati in "Nome somalo" e "Varianti", suggerita dai primi testi esaminati (zoologici), e all'inizio apparentemente ragionevole, mi è poi sembrata del tutto arbitraria e sviante. Man mano che procedeva lo spoglio dei testi botanici (finora solo 4, del resto), il numero di grafie "somale" riferite a un singolo nome scientifico diventava sempre più alto, particolarmente in certi casi (v. per es. *Acacia tortilis*, *Cenchrus ciliaris*, *Cordeauxia edulis*; a proposito di quest'ultima specie, Kazmi osserva di aver rilevato nella letteratura 15 grafie diverse (io ne ho rilevate 25), "nessuna delle quali riproduce fedelmente il suono comprensibile a un Somalo"); d'altro lato diventava sempre più evidente che nessun criterio di giudizio consentiva di differenziare (tantomeno in questa fase preliminare di raccolta dati) lo status linguistico di una grafia rispetto a quello delle altre (v. nota 7. - V. anche l'osservazione di Madany su come "una lettera possa produrre una notevole differenza; per es., 'Dhuur' è il nome di alberi del genere *Tamarix*, mentre 'Duur' è l'erba *Andropogon kelleri*"; io, finora, per *Andropogon kelleri* ho raccolto un 'Duhr', un 'Dur', più altro di completamente diverso, ma nessun 'Duur'; e per *Tamarix* un 'Dur', un 'Dhur', un 'Dor', un 'Dorr', un 'Dot', ecc. e nessun 'Dhuur').
11. Tenere raggruppate tutte le grafie (meno una) in un singolo campo avrebbe impedito procedure di elaborazione, ordinamento, confronto, ecc. dei dati (procedure forse utili, tra l'altro, anche rispetto ai problemi di cui sopra), che la distribuzione in singoli record invece consente. Infine, il numero di "varianti" diventava troppo elevato per la capienza massima di un campo dBASE.
12. I campi 'Italiano' e 'Altre lingue' li ho tolti in quanto inutili per mancanza di dati. I campi relativi a gruppi sistematici superiori alla 'Famiglia' non mi sembrano al momento di alcuna utilità, né credo lo saranno in futuro (comunque resta sempre semplicissimo trasferire i dati esistenti in una eventuale nuova struttura più ampia).
13. Non ho registrato Licheni, Alghe, Muschi e Funghi, per il grande numero di specie che ne fanno parte, e per le quali è del resto improbabile esistano nomi corrispettivi in somalo.

Zoologia

1. I dati sono ripartiti come segue:

MAMMIFERI	<ZMAM>
UCCELLI	<ZUCC>
ANIMALI DOMESTICI	<ZDOM>
RETTILI E ANFIBI	<ZR&A>
INVERTREBRATI	<ZINV>
PESCI	<ZPES>

2. L'attuale struttura è questa:

Campo	Nome campo	Tipo	Dim
1	NOMESOMALO	Carattere	40
2	ITALIANO	Carattere	60
3	ALTRLINGUE	Carattere	68
4	IDENTIFIC	Carattere	60
5	NOMESCIENT	Carattere	60
6	SOTTOFAMGL	Carattere	30
7	FAMIGLIA	Carattere	30
8	ORDINE	Carattere	30
9	CLASSE	Carattere	3
10	SOTTOTIPO	Carattere	1
11	TIPO	Carattere	2
12	RIF	Carattere	24
13	NOTE	Memo	10
Totale:			419

3. Per l'eliminazione del campo 'Varianti' valgono all'incirca le osservazioni in proposito fatte nelle 'Note' relative alla Botanica (così pure per la significatività dei "nomi somali").
4. Il campo 'Identific' si è reso necessario per l'impossibilità di servirsi uniformemente di uno stesso termine di riferimento. Inizialmente avevo assunto, come termine di riferimento, il 'Nome scient.' (Genere-specie), e questo ha funzionato per molti dati. Alcuni però davano solo il riferimento 'Famiglia', altri solo l'Ordine, ecc. Per gli animali domestici in particolare non c'è riferimento ad alcun gruppo sistematico (com'è ovvio), bensì a una descrizione "non scientifica". Il campo 'Identific', dove sono registrati i diversi riferimenti, risponde all'esigenza di avere un campo di riferimento unico e uniforme.
5. Ho lasciato i gruppi sistematici in cui nell'uno o nell'altro file, anche se non in tutti, figurano dei dati. A proposito dei gruppi sistematici - a parte il problema più ampio della relazione tra nomenclatura vernacolare e classificazione scientifica (ad es., come notavo prima, per gli animali domestici la classificazione scientifica è assai poco significativa) - c'è in subordine il problema di quanti e quali livelli tassonomici prendere in considerazione. La soluzione adottata qui è probabilmente troppo ricca e comunque tutta da rivedere.

6. Per i gruppi 'Classe', 'Sottotipo' e 'Tipo' mi sono servito di abbreviazioni:

CLASSE	SOTTOTIPO	TIPO
Ma = Mammiferi	V = Vertebrati	C = Cordati
U = Uccelli		
R = Rettili		
Anf = Anfibi		
Pes = Pesci		

In ZDOM (Animali domestici) e ZINV (Invertebrati), nei campi rel. a questi tre gruppi non ci sono dati.