

The Onomastic Treasures of the CIA

Edwin D. Lawson

State University of New York, Fredonia

1995 ANS Presidential Address

During the Cold War, the United States Central Intelligence Agency prepared documents on personal names for more than 30 languages, from the familiar (German, Russian) to the exotic (Gujerati, Telegu) and from those with relative few speakers (Estonian 1.35 million) to those with many millions (Chinese 610 million). The reports range in length from 14 pages (Slovenian) to 433 pages (Russian). The median number of pages is 46. While the documents vary in quality and coverage, they all contain much of onomastic value. Most have sections giving background on naming in the language, style of name use, pronunciation, transliteration, given names, family names (where applicable), and the use of titles. Some give the meanings of names and some list special features such as laws regarding naming, patronymics, and rules for women's names.

I first became aware of the CIA (Central Intelligence Agency) onomastic documents in 1989 when I was searching databases for a new annotated bibliography on personal names. (Published as *More Names and Naming* by Greenwood in 1995). A search of OCLC (Online Computer Library Search) returned several thousand items. Standing out from the others were the documents produced by the CIA on personal names in foreign countries. A list of the documents is shown in table 1. There are 38 reports in all, 37 on names in individual languages, and one on name style for indexing.

Most of the items were not available on Interlibrary Loan, but the quality of the few that were, and which I was able to obtain, made it seem worthwhile to request others. I had to make a formal request to the CIA for the documents, and with the help of Congressman Amo Houghton of my district, I got most of what I had requested. (The CIA reported that it could not find the document for Thai).

Recently, I rechecked CIA documents and found that eight more titles had been added to the OCLC list. I obtained one (Lithuanian) from a university library after repeated requests. I think that three documents (Nepali, Thai, and Uzbek) may be available through a new microfilm program of the Library of Congress. Four documents (French, Lao, Spanish, and Word Order) are available only through the CIA. I filed an application for them under the Freedom of Information Act.

Names 44.2 (June 1996):154-164

ISSN:0027-7738

© 1996 by The American Name Society

Table 1. CIA Documents on Personal Names.

Albanian Personal Names, 1966. viii + 29p.
**Amharic Personal Names*, 1965. v + 53p.
Arabic Personal Names, 1964. ix + 433p.
Armenian Personal Names, 1965. v + 50p.
Bulgarian Personal Names, 1964. v + 46p.
Burmese Personal Names, 1961. v + 39p.
Chinese Personal Names, 1961. vii + 65p.
Czech Personal Names, 1964. vii + 47p. Map.
Estonian Personal Names, 1965. vii + 65p.
Finnish Personal Names, 1963. v + 38p.
+*French Personal Names*, 1968, 19p.
German Personal Names, 1961. vii, 35p.
Greek Personal Names, 1962. v + 45p.
Gujerati Personal Names, 1965. v + 44p.
**Hausa Personal Names*, 1965. v + 49p.
Hindi Personal Names, 1964. vi + 75p.
Hungarian Personal Names, 1961. vii + 31p.
Korean Personal Names, 1962. vi + 46p.
**Latvian Personal Names*, 1963. iii + 70p.
+*Lao Personal Names*, 1967. 65p.
Lithuanian Personal Names, 1968. v + 36p.
Mongolian Personal Names, 1966. v + 34p.
+*Nepali Personal Names*, 1968. 37p.
Polish Personal Names, 1962. vii + 60p.
Rumanian Personal Names, 1961. vii + 20p.
Russian Personal Names, 1968. vii + 161p.
Serbo-Croatian Personal Names, 1962. vii + 32p.
Slovak Personal Names, 1964. vii + 51p. Map.
Slovenian Personal Names, 1962. v + 14p.
+*Spanish Personal Names*, 1969. v + ?
Swahili Personal Names, 1962. vii + 28p.
Swedish Personal Names, 1967. v + 46p.
Telugu Personal Names, 1964. iii + 50p.
Thai Personal Names, 1964. v. 1., 79p.
Turkish Personal Names, 1961. v + 22p.
+*Uzbek Personal Names*, 1969. 22p.
Vietnamese Personal Names, 1961. vii + 34p.
+*Word Order Standards for Indexing Personal Names*, 1970. 43p.

Records for all titles show author as: United States. Central Intelligence Agency.

*Report currently available through Interlibrary Loan.

+Reports recently identified and requested. Citation information incomplete.

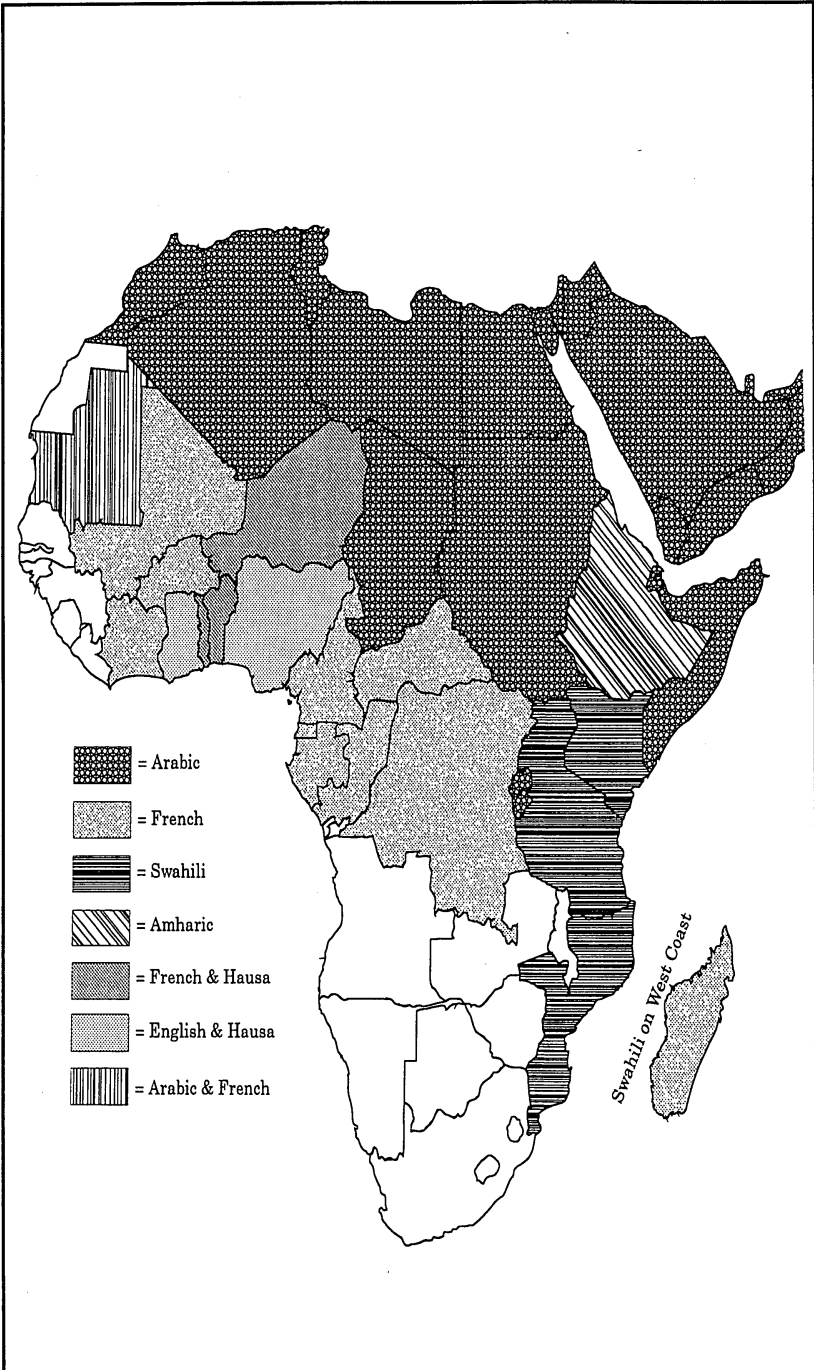
Map 1. Areas of the World Included in the CIA Documents



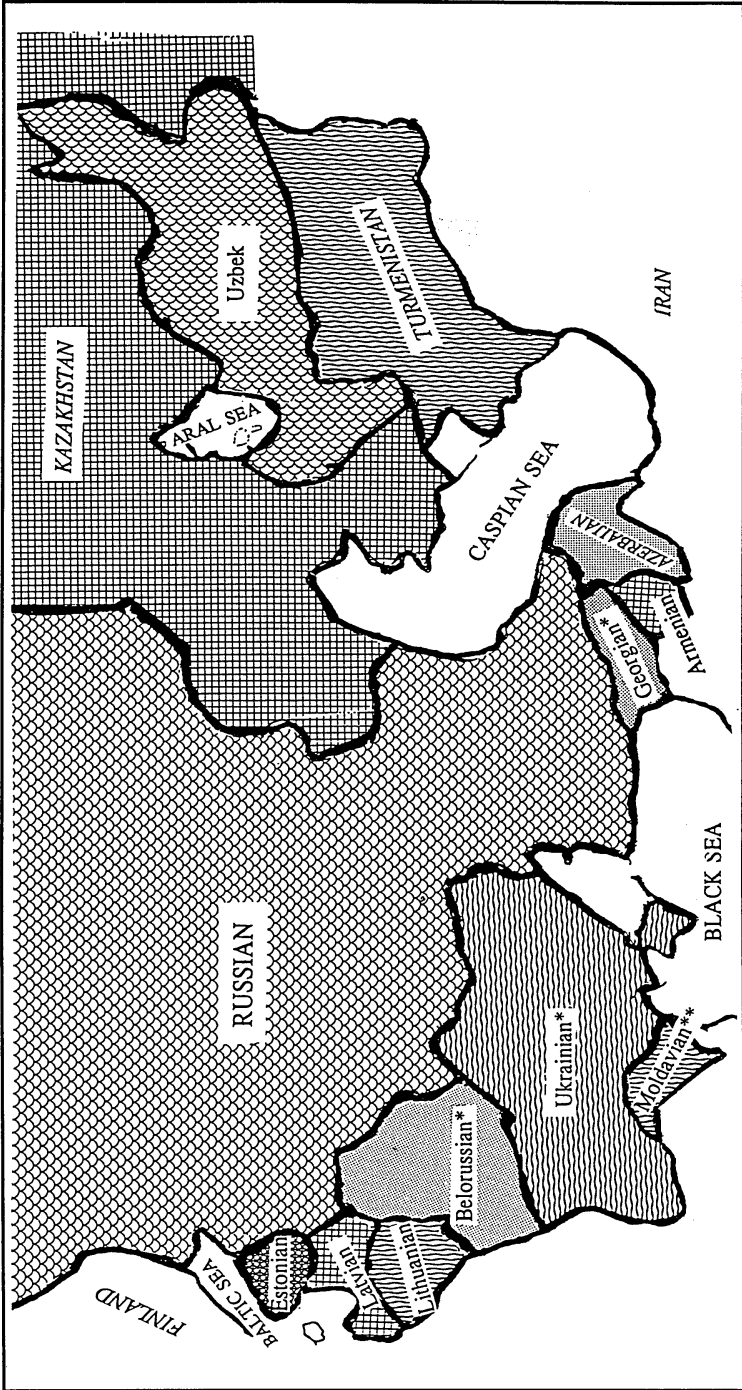
Map 2. European Language Areas Included in the CIA Documents



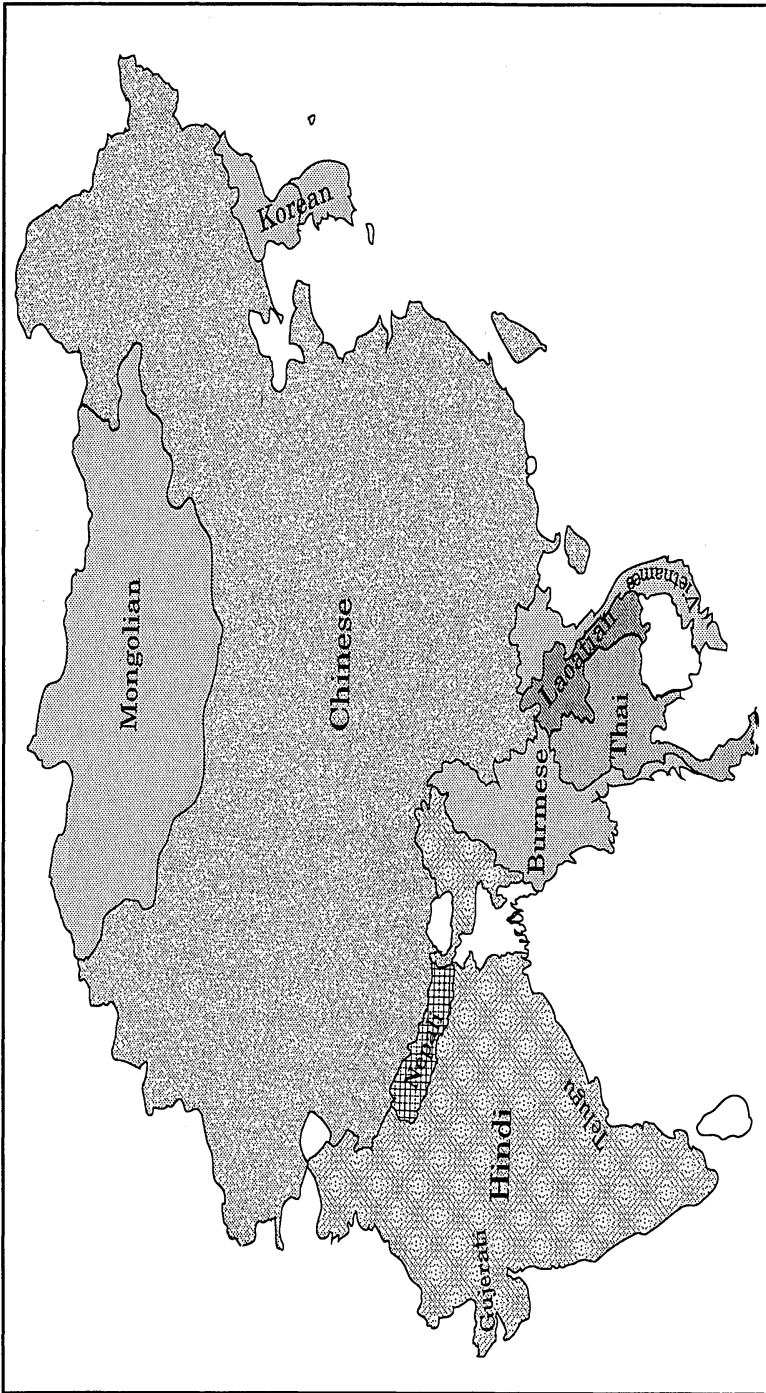
Map 3. African Language Areas Included in the CIA Documents



Map 4. Western Parts of the Former Soviet Union and Language Areas Included in the CIA Documents



Map 5. Asian Language Areas Included in the CIA Documents



Maps 1-5 show the areas covered by the documents. The shaded portions of map 1 show the world regions covered. Map 2 shows that Europe is completely covered except for Portugal, Holland, Denmark, Norway, and Luxembourg. Africa, map 3, is fairly well covered and it should be noted that names are written in Cyrillic for those languages lying on a line from the east where Swahili is spoken, through the center of the continent where French and Swahili are spoken, to the western Hausa-speaking areas.

Contents of the Reports

Since the reports vary considerably, it is difficult to summarize them together. Nevertheless, they all contain some basic information. This is presented in table 2.

The following categories are used in the documents:

- Location:** Where the language is spoken. Usually, this involves more than one country but generally there is one country where the language is centered. Albania is on the SE coast of the Adriatic, bounded by Greece on the S, Yugoslavia on N & E.
- Speakers:** Status and number of speakers. In Albania, where the official language is Albanian, there are 1.8 million speakers plus another million in Yugoslavia and 350,000 in Italy. Arabic stands in strong contrast. The locations where Arabic is spoken range from Morocco in northwest Africa as far as Oman on the southeast coast of the Arabian Peninsula. In addition, speakers of Kurdish, Urdu, Pashto, and Iranian use Arabic script.
- Carding:** Carding, a term new to me, is the way librarians index names. Many countries follow the style used in the US, but certainly not all. Albania, whose carding for libraries may be similar to ours, indexes its telephone directories by given name. There are also other countries that index directories by given name and by title as well.
- Style:** How individuals use their names in everyday life. Muslim countries vary in style. In some there were no surnames until recently: 1926 in Iran; 1935 in Turkey; 1959 in Tunis. Patronyms are important in some cultures, less so in others.
- Language Group:** Language Family Affiliation. Armenian is an independent branch of Indo-European; Burmese belongs to the Indo-Chinese group.
- Alphabets:** All documents have notes on the alphabet used. Several show the actual alphabet. Many languages have their own variations on an existing alphabet, such as Roman or Cyrillic. These are indicated by a + sign in table 2. Hausa and Turkish, which now use a modified Roman script, previously used Arabic script. An additional note is

that Croatian and Serbian differ primarily in that Croatian uses the Roman alphabet and Serbian uses Cyrillic. The same difference occurs between Rumanian (Roman) and Moldavian (Cyrillic).

Pronunciation: Notes are included in most documents. Several give pronunciation guides.

Transcription tables: The documents include transliteration tables as necessary. Several languages are translated into a number of scripts. Transcriptions into Cyrillic include Albanian, Armenian, Bulgarian, Czech, Estonian, Finnish, Greek, Hausa, Hungarian, Latvian, Mongolian, Polish, Rumanian, Slovak, Swahili, Swedish, Turkish, and Vietnamese. Swahili also has Chinese transcriptions.

Given names: When dealing with so many different cultures *given name* is more appropriate than *first name*. Part of the reason is that what we know as the first name may not be the first name in another culture. Table 2 shows the number of given names in the CIA reports. There is tremendous range: from 11 in Rumanian to 15,000 in Arabic. Further, five reports (Amharic, Finnish, Hungarian, Polish, and Serbo-Croatian) give the meanings of the names, which should be of great interest to many scholars.

Family names: This term seems more suitable than *surname*. Family names are not used in some cultures. In Burmese, a name can be either a surname or a family name. Examples are given for most of the languages that use family names; these range from 34 for Vietnamese to 1800 for Latvian.

Special features: Unique features include telecodes for Chinese and Korean. Telecodes were originally used for telegraphy to abbreviate names. Since Chinese and Korean names may require as many as 50 strokes, using a four-digit code offers many advantages. Hindi uses group names to identify individuals as a member of a social or cultural group but do not use surnames as we know them. Estonian includes some farm names.

Titles: The documents list the types of titles that appear in the language. German, for example, has titles for general social address, nobility, professional and civil service, military, and academics.

References: About a dozen reports list references. Most seem to refer to the language itself rather than to names.

These documents contain a great deal of information on names and naming and are of considerable importance to onomastic scholars. Some (perhaps much) of the information appears nowhere else.

Table 2. Basic Information on the Languages

Document Title	Alphabet	Given Names	Meanings	Family Names	Meanings	References	Additional
Albanian	Ro+	375		600			Laws on names
Amharic	Amh	1500	Yes	N/U			
Arabic	Arab.	15,000				11	New laws.
Armenian	Arm.	850		1000			
Bulgarian	Cy+	200		(400)*			*Can also be patronyms.
Burmese	Bur	(126)*		(126)*			*A name can be a given name, surname, or patronym.
Chinese	Ch	20		200		8	Telecodes.
Czech	Ro+	275		1700			
Estonian	Ro+	100		1275			Some farm names. Russ. forms.
Finnish	Ro+	100	Yes	85		2	
German	Ro+	23					
Greek	Gr	75/35		70		10	Notes on naming customs.
Gujerati	Hind+	900		200			
Hausa	Ro+	375		N/U		3	Previously used Arabic script.
Hindi	Hind	900		N/U		2	Group names are used.
Hungarian	Ro+	95/71	Yes	64			
Korean	Ko+	230		276			Telecodes.
Latvian	Ro+	800		1800			German versions shown.
Mongolian	Cy+	300		N/U			

Document Title	Alphabet	Given Names	Meanings	Family Names	Meanings	References	Additional
Polish	Ro+	650	Yes	1600		2	Moldavian similar but with Cyrillic script.
Romanian	Ro+	11/11		Endings		4	Ukrainian & Belarussian close.
Russian	Cy+	4000		Endings		2	
Serbo-Croatian	Ro+/Cy	175	Some	50-76			
Slovak	Ro+	500		1200			
Slovenian	Ro+	70		34	Some		
Swahili	Ro	80		Possible		1	Previously used Arabic script.
Swedish	Ro+	440		1300			
Telugu	Tel	220		—		11	
Turkish	Ro+	100/19		50		23	Telecodes for family names.
Vietnamese	Ro+	70		40			

Abbreviations for Alphabets:

Ro+ = Modified Roman	Cy+ = Modified Cyrillic	Hind = Hindi
Amh = Amharic	Bur = Burmese	Hind+ = Modified Hindi
Arab = Arabic	Ch = Chinese	Tel = Telegu
Arm = Armenian	Gr = Greek	Ko = Korean

Note: Given name frequencies with a slash, as Greek 75/35 indicates that the first frequency is for male names, the second for female. N/U = Not used.