Appendix C: Diacritics and Special Characters

BACKGROUND: **Diacritics** are marks placed over, under, or through letters or combinations of letters that indicate differences in pronunciation and meaning between the marked letter and its unmarked counterpart.   Common diacritics are the umlaut (for example, ü), the acute (for example, ó), the grave (for example, à), and the macron (for example, ō).   **Double diacritics** are two diacritics modifying a single letter, for example, Nguyễn.   **Special characters** occupy a place of their own rather than modifying a letter.   Common special characters include the Polish l (ł), mi͡agkiĭ znak (ʹ), the ayn (ʻ), and the alif (ʼ).   **Superscript characters** are letters, numbers, or symbols written or printed higher on a line of text; **subscript characters** are letters, numbers, or symbols written or printed lower on a line of text.

PLACEMENT OF DIACRITICS IN LCSH: In *LCSH* diacritics and special characters are generally printed in accordance with either the LC romanization conventions for a particular language or with a conventional form found in reference sources.

1. Policy for headings established in the name authority file and used as subject headings*.* Transcribe from name authority records valid personal, corporate, and jurisdictional names and uniform titles, with all indicated diacritics and special characters placed in accordance with LC's transcription or romanization conventions for a particular language.

> ***Accent marks over uppercased letters in French, Portuguese, and Spanish****.* In February 2006 the previous prohibition on use of accent marks over uppercased letters in French, Portuguese, and Spanish was rescinded.   Name headings established or revised after that date reflect the changed policy.

2. Policy for headings established in the subject authority file*.*   Transcribe existing headings and subdivisions exactly as they have been established in the subject authority file, using diacritics and special characters as indicated.

3. Establishing subject headings with or without diacritics.   When, as a result of doing authority research, a choice must be made between a term without diacritics and the otherwise identical term with diacritics, generally prefer the term without diacritics if it is *clearly* the conventional English form.   If there is no conventional English form, prefer the form of the term as found in the *Encyclopaedia Britannica* or in reference sources of the particular discipline or language, using diacritics to conform to LC's transcription or romanization conventions.

Appendix C: Diacritics and Special Characters

4. Diacritic use and LC practice. If it is determined that a heading or reference is to have diacritics, use diacritics that conform to LC transcription or romanization conventions, even if reference sources, including BGN, indicate different diacritics.   Use judgment in determining whether to make UF references from forms with variant diacritics that are spelled differently from the proposed heading.   (For forms that have variant diacritics but are spelled the same as the proposed heading, see sec. 5, below.)   *Examples:*

BGN approved form: Jabal S̲ā'id, mountain

Heading as established:

     151 ## $a Sā'id Mountain (Saudi Arabia)
> *[BGN left hook under* S *changed to dot under* S *to conform to LC romanization practice; reference not made from* S̲a'id *since spelling is identical.]*

BGN approved form: Nísos Andíparos

Heading as established:

     151 ## $a Antiparos Island (Greece)
     451 ## $a Andiparos Island (Greece)
     451 ## $a Nísos Andíparos (Greece)
> *[Accent marks supplied by BGN in second 451 field]*

Encyclopaedia Britannica: Ksatriya

Heading as established:

     150 ## $a Kshatriyas
     450 ## $a Ksatriyas
> *[Since spelling is not identical, UF reference made from the form based on Sanskrit with dot under* s *to conventional English plural form.]*

Appendix C: Diacritics and Special Characters

4. Diacritic use and *LC* practice.   *(Continued)*

> ***Accent marks over uppercased letters in French, Portuguese, and Spanish.***   Add diacritics according to the usage of the language.   *Example:*

```
151 ## $a Basques, Île aux (Québec)
```

5. Diacritics in subject headings and references.   Because some online systems cannot distinguish between a term without diacritics and the otherwise identically spelled term with diacritics, do not add 4XX fields for variants of this type, even if the terms do not represent the same word in a particular language.   The following are examples of 4XX fields that should not be made:

```
100 3# $a Cheng family
400 3# $a Chʻeng family

100 3# $a Ota family
400 3# $a Ōta family
```

> *Note: For instructions on headings for family names of different meanings or different ethnic backgrounds that are distinguished only by the use of diacritical marks, see H 1631, sec. 3.c.*

```
150 ## $a Nirvana
450 ## $a Nirvāna

151 ## $a Drava River
451 ## $a Dráva River
```

6. Superscript/Subscript characters.   Since superscript or subscript characters are positioned on the same line as the rest of the heading both in the online authority and bibliographic records, as well as in the printed edition of *LCSH*, treat them as normal characters rather than as superscript or subscript characters. *Examples:*

Term:	C^*-algebras
Transcribe as:	`150 ## $a C*-algebras`
Term:	L_1 algebras
Transcribe as:	`150 ## $a L1 algebras`
Term:	L_p spaces
Transcribe as:	`150 ## $a Lp spaces`

Appendix C: Diacritics and Special Characters

7. Greek or Hebrew letters.   Romanize Greek or Hebrew letters when establishing subject headings.   Do not use Greek or Hebrew letters in headings or references, even though they appear as such in reference sources.   *Examples:*

```
      150 ## $a Chi Rho symbol
[not 150 ## $a XP symbol]

      150 ## $a Chi-square test
[not 150 ## $a χ-square test]

      150 ## $a Lambda calculus
[not 150 ## $a λ calculus]

      150 ## $a Pi
[not 150 ## $a π ]
```

8. Ampersands.   Do not use ampersands in form or topical subdivisions, or to connect two words in the qualifier of a geographic heading.   In subject headings, spell out ampersands using the word **and**.

> *Exception:*   When assigning as a subject heading a name heading established with an ampersand, use the ampersand.   *Examples:*
>
> ```
>       610 20 $a C.S. Wertsner & Son.
>       610 20 $a Dow Jones & Co.
> ```

9. Common diacritics and special characters:

a. Acute   (´)

> Avalokiteśvara (Buddhist deity)
> Cévennes Mountains (France)
> Daśnāmīs
> Mossoró (Race horse)

> Do not confuse this diacritic with the **mi͡agkiĭ znak** (′) and the **apostrophe** ('), which are special characters.

Appendix C: Diacritics and Special Characters

9. Common diacritics and special characters: *(Continued)*

b. Alif (')

S̲ā'id Mountain (Saudi Arabia)

Do not confuse with the **acute** (´), the **mi͡agkiĭ znak** (′), the **apostrophe** ('), or the **ayn** (').

c. Angstrom *or* Circle above letter (°)

Skovsgård (Denmark)
Småland (Sweden)
Lundsgård (Denmark)

Do not confuse with the **dot above letter** (·).

d. Apostrophe (')

Coups d'état
Martin's Hundred Site (Va.)

Do not confuse with the **acute** (´), the **mi͡agki̇̆ znak** (′), the **alif** ('), the **ayn** ('), or the **grave** (`).

e. Ayn (')

'Īd al-Ghadīr
Shī'ah

Do not confuse with the **acute** (´), the **mi͡agki̇̆ znak** (′), the **alif** ('), or the **apostrophe** (')

Appendix C: Diacritics and Special Characters

9. Common diacritics and special characters: *(Continued)*

f. Breve   (˘)

>>> Bolʹshoĭ prospekt (Saint Petersburg, Russia)
>>> Hanmun sosŏl (Korean fiction)
>>> Umileniĭe TSaritsy Nebesoĭ (Icon)

>> Do not confuse with the **hacek** (ˇ)

g.   Cedilla   (¸)

>>> Mogi-Guaçu River (Brazil)
>>> Redução Jesuítica de Jesus Maria Site (Brazil)

>> Do not confuse with the **left hook** (˛).

h. Circle below letter   (̥)

>>> Ksattr̥ (The Sanskrit word)

>> Do not confuse with the **dot below letter** (̣).

i. Circumflex   (^)

>>> Château de Ripaille (Thonon-les-Bains, France)
>>> Rodovia Transamazônica (Brazil)
>>> Nkâmi (African people), USE Nkomi (African people)

j. Dot above letter *or* Superior dot   (˙)

>>> Pasvik Ėlʹf, USE Pats River
>>> Kórkyra Mélaina (Croatia), USE Korcula Island (Croatia)

>> Do not confuse with the **angstrom** or **circle above letter** (°).

Appendix C: Diacritics and Special Characters

9. *Common diacritics and special characters*:   *(Continued)*

k. *Dot below letter*   (̣)

> Ksattr̥ (The Sanskrit word)
> Kundalinī
> Hzh (The Hebrew root)

Do not confuse with the **circle below letter** (̥).

l. *Double acute*   (″)

> Hungary.   Földművelésügyi Minisztérium

This diacritic appears only in Hungarian.   Do not confuse with the **tverdyi znak** (″), a special character that follows the letter it modifies, the **Umlaut** (¨), or the standard quote mark (").

m. *Grave*   (`)

> Corbières Mountains (France)
> Monge-Ampère equations
> Pietà

Do not confuse with the **apostrophe**   (').

n. *Hacek*   (ˇ)

> Český kras (Czech Republic)
> Pelješac Peninsula (Croatia)
> Šumadija (Serbia : Region)

Do not confuse with the **breve** (˘).

Appendix C: Diacritics and Special Characters

9. Common diacritics and special characters: *(Continued)*

 o. Left hook (˛)

    Mun̦tii Parângu (Romania), use Parîng Mountains (Romania)
    Oaş Mountains (Romania)

  Do not confuse with the **cedilla** (¸).

 p. Ligature (͡)

    I͡Amal Peninsula (Russia)
    Iverskai͡a Svi͡atai͡a i Chudotvornai͡a Bogomater′ (Icon)
    Ili͡ushin airplanes

 q. Macron (¯)

    Ādityas (Hindu deities)
    Osakajō (Osaka, Japan)
    Sonārāya (Hindu deity)

 r. Mi͡agkiĭ znak (′)

    Bol′shoĭ prospekt (Saint Petersburg, Russia)
    Ob′ River (Russia)

  Do not confuse with the **acute** (´), the **alif** (ʾ), the **apostrophe** ('), or the **ayn** (ʿ).

 s. Polish l (ł)

    Pałac Rady Ministrów (Warsaw, Poland)
    Łososina River (Poland)

  Note that Polish also has a regular l.

9. Common diacritics and special characters: *(Continued)*

t. Slash o (ø)

> Andøya (Norway)
> Eigerøy (Norway)
> Køge Bay (Denmark)

u. Tilde (~)

> Akwe̶-Shavante language, USE Xavante language
> El Niño Current
> Peñarroya, Nuestra Señora de
> São Francisco River (Brazil)

v. Turkish i (ı)

> Note that ı is *not* dotted.   If uppercased, render as I.
> Turkish also has a regular i.

w. Tverdyi znak (″)

> Ob″edinenie Stavropol′neftegaz

> Do not confuse this special character with the **double acute** (″) of Hungarian, the **umlaut** (¨), or the standard quote mark (").

x. Umlaut (¨)

> Moësa River (Switzerland)
> Östermalm (Stockholm, Sweden)
> Ukraïna (The Ukrainian word)

> Do not confuse this diacritic with the **double acute** (″) of Hungarian, the **tverdyi znak** (″), or the standard quote mark (").