

# The Chicago Manual of Style

**SEVENTEENTH EDITION**

*The University of Chicago Press*

CHICAGO AND LONDON

# 11 · Languages Other than English

OVERVIEW 11.1

GENERAL PRINCIPLES 11.3

Words and Phrases from Other Languages 11.3

Titles of Works from Other Languages 11.6

Quotations from Other Languages 11.11

LANGUAGES USING THE LATIN ALPHABET 11.18

African Languages 11.23

French 11.25

German 11.38

Italian 11.46

Latin 11.54

Spanish 11.60

Other Languages Using the Latin Alphabet 11.70

LANGUAGES USUALLY TRANSLITERATED (OR ROMANIZED) 11.71

Arabic 11.76

Chinese and Japanese 11.82

Hebrew 11.91

Russian 11.98

South Asian Languages 11.109

CLASSICAL GREEK 11.110

Breathings and Accents 11.112

Punctuation and Numbers 11.116

Word Division 11.118

OLD ENGLISH AND MIDDLE ENGLISH 11.122

AMERICAN SIGN LANGUAGE (ASL) 11.125

## Overview

- 11.1 **Scope and organization.** This chapter provides guidelines for presenting text from languages other than English in English-language contexts. These guidelines are general: authors or editors working with languages in which they are not expert should seek additional guidance from someone who is. More than two dozen languages are covered, with those languages that commonly appear and those that present complex problems being considered most fully. The chapter begins with the treatment of words and phrases, titles of works, and quotations, the principles of which apply to most of the languages discussed (see 11.3–17). It then addresses languages using the Latin alphabet, transliterated (or romanized) languages, classical Greek, Old English and Middle English, and American Sign Language. Individual languages or groups of languages are presented in alphabetical order within their particular sections. (For the treatment of personal names, see 8.7–18.)
- 11.2 **Unicode.** Many of the letters and symbols required by the world’s languages are included in a widely used standard for character encoding called Unicode. The Unicode standard (published by the Unicode Consortium; bibliog. 2.7) is widely supported by modern operating systems and browsers and many other applications (including word processors) and is required by such standards as XML and EPUB. Unicode assigns a unique identifying hexadecimal number (or code point) and description to tens of thousands of characters. Even fonts with Unicode character mapping, however, typically support only a subset of the Unicode character set. For this reason, it is desirable to determine at the outset which characters will be needed for a publication. Table 11.1 lists special characters, with Unicode numbers and abbreviated descriptions, needed for each of the languages treated in this chapter that use the Latin alphabet. Table 11.2 lists special characters that may be needed for certain transliterated languages. For Russian (Cyrillic) and Greek characters, see tables 11.3, 11.4, and 11.5. Unicode numbers mentioned in text should be prefixed by U+ (e.g., U+00E0 for à).

## General Principles

### *Words and Phrases from Other Languages*

- 11.3 **Non-English words and phrases in an English context.** Italics are used for isolated words and phrases from another language, especially if

they are not listed in a standard English-language dictionary like *Merriam-Webster's Collegiate* (see 7.1) or are likely to be unfamiliar to readers (see also 7.54). (For proper nouns, see 11.4.) If such a word or phrase becomes familiar through repeated use throughout a work, it need be italicized only on its first occurrence. If it appears only rarely, however, italics may be retained.

The *grève du zèle* is not a true strike but a nitpicking obeying of work rules. She preferred to think of it optimistically as a *sueño reparador*—rather than, as in English, a sleep that was merely restful.

Unless the term appears in a standard English-language dictionary and is being used as such, observe the capitalization conventions of the original language. In the following examples, the German word for computer (which is the same as the English word) is capitalized because it is a noun, and the French adjective *française* is lowercase even though it would be capitalized in English (as “French”). See also 11.18.

The German word for computer is *Computer*. The French word is *ordinateur*. In Spanish, the word is either *computadora* or *ordenador*, depending on region or context.

We were prepared to learn the nuances of *la langue française*.

The plurals of non-English words should be formed as in the original language (see also 7.12).

We were sent off with some beautiful *Blumen* (not *Blumes* [italic ess] and not *Blumes* [roman ess]).

An entire sentence or a passage of two or more sentences in another language is usually set in roman and, unless it is set as a block quotation or extract (see 13.9–29), enclosed in quotation marks (see 11.11).

- 11.4 **Non-English proper nouns in an English context.** With the exception of titles of books and the like, proper nouns from other languages are generally *not* italicized, even on first mention (cf. 11.3). This usage extends to named places and structures, institutions and companies, brand names, and other categories as discussed in chapter 8. (For titles of works, see 11.6–10.) Capitalization should follow predominant usage in the original language. In some cases, this may entail observing a preference for capitalization that runs counter to the conventions for generic text. If the editor is unfamiliar with the language, an expert, or the author, should be consulted; when in doubt, opt for sentence-style capitalization (see

8.158). See also 11.18. An initial *the* may be used if the definite article would appear in the original language.

She won the Premio Nadal for her second novel, *Viento del norte*.

Mexico City's Ángel de la Independencia is known familiarly as "El Ángel."

The Real Academia Española was founded in 1713.

A history of the Comédie-Française has just appeared.

The Académie française dates to the reign of Louis XIII.

I prefer the Bibliothèque nationale by day and the Bois de Boulogne by night.

He is a member of the Société d'entraide des membres de l'ordre national de la Légion d'honneur.

Leghorn—in Italian, Livorno—is a port in Tuscany.

When he asked her to meet him along Unter den Linden, she was amused by the consecutive prepositions—one in English and one in German, just like them.

Translations of proper nouns from other languages should be capitalized headline-style (see 8.159).

He is a member of the Mutual Aid Society for Members of the National Order of the Legion of Honor.

Original (or transliterated) names of proper nouns presented as glosses should not be italicized (but see 11.5).

The number of cases adjudicated by the Supreme People's Court of the People's Republic of China (Zhonghua renmin gongheguo zuigao renmin fayuan) has increased sharply.

11.5 **Translations of terms from other languages.** A translation following a word, phrase, or title from another language is enclosed in parentheses or quotation marks. See also 6.96, 11.3, 11.4, 11.9, 14.99.

The word she wanted was *pécher* (to sin), not *pêcher* (to fish).

The Prakrit word *majjao*, "the tomcat," may be a dialect version of either of two Sanskrit words: *madjaro*, "my lover," or *marjaro*, "the cat" (from the verb *mrij*, "to wash," because the cat constantly washes itself).

A group of German expressionists known as Die Brücke (The Bridge) were influential in the decade leading up to the First World War.

Leonardo Fioravanti's *Compendio de i secreti rationali* (Compendium of rational secrets) became a best seller.

If a non-English word other than a proper noun is presented as a parenthetical gloss, it should be presented in italics as in running text (but see 11.4).

He said that to fish (*pêcher*) was to sin (*pécher*).

For quotations from other languages, see 11.11–17.

### *Titles of Works from Other Languages*

- 11.6 **Capitalization of titles from other languages.** For titles of works from other languages, whether these appear in text, notes, or bibliographies, Chicago recommends a simple rule: capitalize only the words that would be capitalized in normal prose—the first word of the title and subtitle and all proper nouns or any term that would be capitalized under the conventions of the original language. That is, use *sentence style* (see 8.158). This rule applies equally to titles using the Latin alphabet and to transliterated titles. For examples, see 14.98. For special considerations related to German capitalization, see 11.39. For variations in French, see 11.27.
- 11.7 **Punctuation of titles from other languages.** When a non-English title is included in an English-language context, the following changes are permissible: a period (or, more rarely, a semicolon) between title and subtitle may be changed to a colon (and the first word of the subtitle may be capitalized); guillemets (« ») or other non-English styles for quotation marks may be changed to regular quotation marks (“ ” or ‘ ’); and any space between a word and a mark of punctuation that follows may be eliminated. Commas should not be inserted (even in a series or before dates) or deleted, nor should any other mark of punctuation be added or deleted. See also 8.165.
- 11.8 **Italic versus roman type for titles from other languages.** Titles of works in languages that use the Latin alphabet (including transliterated titles) are set in italic or roman type according to the principles set forth in 8.156–201—for example, books and periodicals in italic; poems and other short works in roman.
- Stendhal’s *Le rouge et le noir* was required reading in my senior year.  
 We picked up a copy of the *Neue Zürcher Zeitung* to read on the train.  
 She published her article in the *Annales de démographie historique*.  
 Strains of the German carol “Es ist ein’ Ros’ entsprungen” reached our ears.  
 Miguel Hernández’s poem “Casida del sediento” has been translated as “Lament of the Thirsting Man.”
- 11.9 **Non-English titles with English translation.** When the title of a work in another language is mentioned in text, an English gloss may follow

in parentheses (see 6.96). If the translation has not been published, the English should be capitalized sentence-style (as in the first example below; see 8.158) and should appear neither in italics nor within quotation marks. A published translation, however, is capitalized headline-style (as in the second and third examples; see 8.159) and appears in italics or quotation marks depending on the type of work (see 8.156–201). Some editorial discretion may be required, especially if the translation is incorporated into running text (as in the third example). For translations of non-English titles in notes and bibliographies, see 14.99. See also 11.10.

Leonardo Fioravanti's *Compendio de i secreti rationali* (Compendium of rational secrets) became a best seller.

Proust's *À la recherche du temps perdu* (*Remembrance of Things Past*) was the subject of her dissertation.

*but*

*La ciudad y los perros*, which literally means “the city and the dogs,” was published in English under the title *The Time of the Hero*.

- 11.10 **Original-language title of work versus translation.** Readership and context will determine whether to use the original or the translated title of a non-English work mentioned in running text. In a general work, titles that are widely known in their English translation could be cited in English first, with the original following in parentheses; in some cases, the original can be omitted entirely. Some authors prefer to cite all non-English titles in an English form, whether or not they have appeared in English translation. As long as the documentation clarifies what has been published in English and what has not, translated titles standing in for the original may be capitalized headline-style and treated like other English-language titles (see 8.159, 8.163). See also 11.9.

“The West” in the title of the Chinese classic *Journey to the West* (*Xī yóu jì*) refers mainly to the Indian subcontinent.

Molière's comedy *The Miser* may have drawn on an obscure late-medieval French treatise, *The Evils of Greed*, recently discovered in an abandoned château.

### *Quotations from Other Languages*

- 11.11 **Typographic style of quotations from other languages.** Quotations from a language other than English that are incorporated into an English text are normally treated like quotations in English, set in roman type and run in or set off as block quotations according to their length. (For a complete discussion of quotations, see chapter 13.) They are punctu-

ated as in the original except that quotation marks can usually replace guillemets (or their equivalents), and punctuation relative to quotation marks and spacing relative to punctuation are adjusted to conform to the surrounding text (see 11.19). For isolated words and phrases, see 11.3. For excerpts from the original language following an English translation, see 11.12.

The narrator's "treinta o cuarenta molinos de viento" become Quixote's "treinta, o pocos más, desaforados gigantes," a numerical correspondence that lets the reader trust, at the very least, the hero's basic grasp of reality.

If em dashes rather than quotation marks are used for dialogue in the original (see 11.31, 11.47, 11.64, 11.101), they should be retained in a block quotation but may be replaced by quotation marks if only a phrase or sentence is quoted in running text.

- 11.12 **Translations relative to quotations.** A translation may follow the original in parentheses—or, as in 11.13, the original may follow a translation. Quotation marks need not be repeated for the parenthetical translation (or parenthetical original, as the case may be); any internal quotation marks, however, should be included (as in the second example). See also 6.96, 11.5. If a long sentence or more than one sentence appears in parentheses or brackets, as in the second example, closing punctuation of the original and the translation should remain distinct.

A line from Goethe, "Wer nie sein Brot mit Tränen aß" (Who never ate his bread with tears), comes to mind.

À vrai dire, Abélard n'avoue pas un tel rationalisme: "je ne veux pas être si philosophe, écrit-il, que je résiste à Paul, ni si aristotélicien que je me sépare du Christ." (As a matter of fact, Abelard admits no such rationalism. "I do not wish to be so much of a philosopher," he writes, "that I resist Paul, nor so much of an Aristotelian that I separate myself from Christ.")

Whether to provide translations of quoted passages depends on the linguistic abilities of the intended audience. For example, in a work to be read by classicists, Latin or Greek sources may be quoted freely in the original. Or in a literary study of, say, Goethe, quotations from Goethe's work may be given in the original German only. For a wider readership, translations should be furnished.

- 11.13 **Source of quotation plus translation.** When both a source and a translation are required in text, the source may be placed in parentheses, with



the original (or translation, as the case may be) following, separated by a semicolon. The following example quotes a thirteenth-century author writing in Middle Dutch. See also 13.68–69.

Hadewijch insists that the most perfect faith is “unfaith,” which endlessly stokes desire and endlessly demands love from God. “Unfaith never allows desire to rest in any faith but always distrusts her, [feeling] that she is not loved enough” (letter 8:39; *Ende ontrowe en laet gegherthen nieuwers ghedueren in gheenre trowen, sine mestrou hare altoes, datse niet ghenoech ghemint en es*).

If adding a translation or the original in text creates too much clutter, it may be placed in a note, in which case it is enclosed in quotation marks but not in parentheses or brackets. If the parenthetical passage in the second example in 11.12 were to appear in text without the French, as either a run-in or a block quotation, a note could read as follows:

1. “À vrai dire, Abélard n’avoue pas un tel rationalisme: ‘je ne veux pas être si philosophe, écrit-il, que je résiste à Paul, ni si aristotélicien que je me sépare du Christ.’”

See also 13.30.

- 11.14 **Crediting the translation of a quoted passage.** When quoting a passage from a language that requires a translation, authors should use a published English translation if one is available and give credit to the source of that translation, including the title of the translation, the translator’s name, relevant bibliographic details, and page number (see 14.99). Authors providing their own translations should so state, in parentheses following the translation, in a note, or in the prefatory material—for example, “my translation” or “Unless otherwise noted, all translations are my own.” If an individual other than the author provided the translations, that person should be credited in a similar manner, but by name. See also 11.9.
- 11.15 **Adjusting translated quotations.** An author using a published translation may occasionally need to adjust a word or two; “translation modified” or some such wording must then be added in parentheses or in a note (see also 13.62). In addition, it is recommended that such modifications be indicated by square brackets (see 13.59, 13.60). These devices should be used sparingly. If a published translation is unsuitable for the author’s purpose, it should be abandoned and all quoted passages newly translated.

- 11.16 **Editing translated quotations.** Quotations from published translations can be modified only with respect to the permissible changes described in 13.7. In new translations furnished by the author, however, capitalization, punctuation, spelling, and idiom may be adjusted for consistency with the surrounding text.
- 11.17 **The sin of retranslation.** Never should a passage from a work originally published in English (or any other language, for that matter) be retranslated from a version that has been translated into another language. For example, an author quoting from a German study of Blackstone's *Commentaries* that quotes from Blackstone in German must track down the original Blackstone passages in English and reproduce them. If unable to locate the original, the author must resort to paraphrase.

## Languages Using the Latin Alphabet

- 11.18 **Capitalization—English versus other languages.** Capitalization is applied to more classes of words in English than in any other Western language (but see 8.1). Most of the other languages discussed in this chapter follow a simpler set of rules. Except where stated to the contrary, the language in question is assumed to lowercase all adjectives (except those used as proper nouns), all pronouns, months, and days of the week. In addition, capitals are used more sparingly than in English for names of offices, institutions, and so on. Translated terms, however, are subject to Chicago's recommendations for capitalization of names and terms (see chapter 8). For personal names, see 8.7–18.
- 11.19 **Punctuation—original language versus English context.** The remarks in this chapter related to punctuation point out the more obvious departures from what is familiar to readers of English. For the purposes of illustration, quotation marks in the style of the original language have been preserved in the examples; however, spacing relative to these and other punctuation marks has been adjusted to conform to the typographic style of this manual. In quotations from other languages (and in translations), regular English-style quotation marks can usually replace the guillemets or whatever is used in the original (with the placement of periods and commas adjusted as needed; see 6.9–11). Dashes used to mark dialogue, however, should be preserved in block quotations presented in the original language. See 11.11. Another exception is the punctuation at the beginning of Spanish questions and exclamations (see 11.62), which should

be preserved for quotations in Spanish (but omitted when the passage is translated).

- 11.20 **Word division for languages other than English.** Though conventions for dividing words at the ends of lines vary widely, the following general rules apply to non-English languages as well as to English: (1) Single-syllable words should never be broken. (2) No words should be broken after one letter, nor should a single letter be carried over to another line (see also 7.37). (3) Hyphenated words and solid compounds should be broken at the hyphen or between elements, if at all possible. See also 7.40; for proper nouns, see 7.42. Specific rules for some of the languages covered in this chapter appear in the relevant sections below.
- 11.21 **Special characters in the Latin alphabet.** Words, phrases, or titles from another language that occur in an English-language work must include any special characters that appear in the original language. Those languages that use the Latin alphabet may include letters with accents and other diacritical marks, ligatures, and, in some cases, alphabetical forms that do not normally occur in English. Table 11.1 lists the special characters that might be required for each language treated in this section. Most authors will have access to Unicode-compliant software (see 11.2) and will therefore be able to reproduce each of these characters without the addition of any specialized fonts. Authors should nonetheless supply a list of special characters used within a manuscript (see 2.16) to ensure the correct conversion to a particular font required for publication or, for electronic projects, to ensure compatibility across systems that may not support Unicode. If type is to be reproduced from an author's hard copy, marginal clarifications may be needed for handwritten accents or special characters. In either case, use table 11.1 to correctly identify the character by name and Unicode number (e.g., for Ð or ð, indicate "D with stroke [U+0110]" or "d with stroke [U+0111]"). For diacritical marks used in transliteration, see 11.74.
- 11.22 **International Phonetic Alphabet (IPA).** Phonetic symbols using IPA notation are based on the Latin alphabet and are defined for Unicode (see 11.2). For the latest version of the IPA alphabet, consult the website of the International Phonetic Association. For additional information on the subject of phonetics, including treatment of other systems of notation, consult Geoffrey K. Pullum and William A. Ladusaw, *Phonetic Symbol Guide* (bibliog. 5).

TABLE 11.1. Special characters (and Unicode numbers) for languages using the Latin alphabet

Character (and Unicode number)	Description	Languages that use it
„ (201E), “ (201C)	double low-9 quotation mark, left double quotation mark	German
« (00AB), » (00BB)	double angle quotation marks (guillemets)	French, German (reversed), Italian, Spanish
‘ (2018)	‘okina (represented by left single quotation mark)	Hawaiian
À (00C0), à (00E0)	A/a with grave	French, Italian, Portuguese
Á (00C1), á (00E1)	A/a with acute	Czech, Hungarian, Icelandic, Portuguese, Spanish
Â (00C2), â (00E2)	A/a with circumflex	French, Moldovan, Portuguese, Romanian, Turkish
Ã (00C3), ã (00E3)	A/a with tilde	Portuguese
Ä (00C4), ä (00E4)	A/a with diaeresis	Finnish, German, Swedish, Turkmen
Å (00C5), å (00E5)	A/a with ring above	Finnish, Danish, Norwegian, Swedish
Ā (0100), ā (0101)	A/a with macron	Hawaiian, Latin
Ă (0102), ă (0103)	A/a with breve	Latin, Moldovan, Romanian
Ą (0104), ą (0105)	A/a with ogonek	Polish
Æ (00C6), æ (00E6)	ligature Æ/æ	Danish, Icelandic, Norwegian, Old English and Middle English
Ɓ (0181), ɓ (0253)	B/b with hook	Hausa
Ç (00C7), ç (00E7)	C/c with cedilla	Albanian, Azeri, French, Portuguese, Turkish, Turkmen
Ć (0106), ć (0107)	C/c with acute	Bosnian, Croatian, Montenegrin, Polish, Serbian
Č (010C), č (010D)	C/c with caron (haček)	Bosnian, Croatian, Czech, Montenegrin, Serbian
Ð (00D0), ð (00F0)	eth	Old English and Middle English, Icelandic
Đ (010E), đ (010F)	D/d with caron (haček)	Czech
Đ (0110), đ (0111)	D/d with stroke	Bosnian, Croatian, Montenegrin, Serbian
Ɖ (018A), ɖ (0257)	D/d with hook	Hausa
È (00C8), è (00E8)	E/e with grave	French, Italian, Portuguese
É (00C9), é (00E9)	E/e with acute	Czech, French, Hungarian, Icelandic, Italian, Portuguese, Spanish
Ê (00CA), ê (00EA)	E/e with circumflex	French, Portuguese
Ë (00CB), ë (00EB)	E/e with diaeresis	Albanian, French
Ē (0112), ē (0113)	E/e with macron	Hawaiian, Latin
Ĕ (0114), ĕ (0115)	E/e with breve	Latin
Ę (0118), ę (0119)	E/e with ogonek	Polish
Ě (011A), ě (011B)	E/e with caron (haček)	Czech
Ʒ (021C), ɟ (021D)	yogh	Old English and Middle English
Ə (018F), ə (0259)	schwa	Azeri
Ğ (011E), ğ (011F)	G/g with breve	Azeri, Turkish
Ì (00CC), ì (00EC)	I/i with grave	Italian, Portuguese

(continued)

TABLE 11.1. (continued)

Character (and Unicode number)	Description	Languages that use it
Í (00CD), í (00ED)	I/i with acute	Czech, Hungarian, Icelandic, Portuguese, Spanish
Î (00CE), î (00EE)	I/i with circumflex	French, Moldovan, Romanian, Turkish
Ï (00CF), ï (00EF)	I/i with diaeresis	French, Portuguese
Ī (012A), ī (012B)	I/i with macron	Hawaiian, Latin
İ (012C), i (012D)	I/i with breve	Latin
İ̇ (0130)	I with dot above	Azeri, Turkish
ı (0131)	dotless i	Azeri, Turkish
Ƙ (0198), ƙ (0199)	K/k with hook	Hausa
Ł (0141), ł (0142)	L/l with stroke	Polish
Ñ (00D1), ñ (00F1)	N/n with tilde	Spanish
Ń (0143), ń (0144)	N/n with acute	Polish
Ň (0147), ň (0148)	N/n with caron (haček)	Czech, Turkmen
Ò (00D2), ò (00F2)	O/o with grave	Italian, Portuguese
Ó (00D3), ó (00F3)	O/o with acute	Czech, Hungarian, Icelandic, Polish, Portuguese, Spanish
Ô (00D4), ô (00F4)	O/o with circumflex	French, Portuguese
Õ (00D5), õ (00F5)	O/o with tilde	Portuguese
Ö (00D6), ö (00F6)	O/o with diaeresis	Azeri, Finnish, German, Hungarian, Icelandic, Swedish, Turkish, Turkmen
Ø (00D8), ø (00F8)	O/o with stroke	Danish, Norwegian
Ō (014C), ō (014D)	O/o with macron	Hawaiian, Latin
Ö (014E), ö (014F)	O/o with breve	Latin
Ő (0150), ő (0151)	O/o with double acute	Hungarian
Œ (0152), œ (0153)	ligature Œ/œ	French
Ř (0158), ř (0159)	R/r with caron (haček)	Czech
Š (015A), š (015B)	S/s with acute	Polish, Montenegro
Ş (015E), ş (015F)	S/s with cedilla	Azeri, Turkish, Turkmen
Ș (0218), ș (0219)	S/s with comma below	Moldovan, Romanian
Š (0160), š (0161)	S/s with caron (haček)	Bosnian, Croatian, Czech, Montenegro, Serbian
ß (00DF)	sharp S (eszett)	German
Ț (021A), ț (021B)	T/t with comma below	Moldovan, Romanian
Ț̣ (0164), ț̣ (0165)	T/t with caron (haček)	Czech
Þ (00DE), þ (00FE)	thorn	Old English and Middle English, Icelandic
Û (00D9), ù (00F9)	U/u with grave	French, Italian, Portuguese
Ú (00DA), ú (00FA)	U/u with acute	Czech, Hungarian, Icelandic, Portuguese, Spanish
Û̇ (00DB), û̇ (00FB)	U/u with circumflex	French, Turkish
Û̈ (00DC), ü̈ (00FC)	U/u with diaeresis	Azeri, French, German, Hungarian, Portuguese, Spanish, Turkish, Turkmen
Ů (016E), ů (016F)	U/u with ring above	Czech
Ū (016A), ū (016B)	U/u with macron	Hawaiian, Latin
Ů (016C), ů (016D)	U/u with breve	Latin
Ű (0170), ű (0171)	U/u with double acute	Hungarian
Ý (00DD), ý (00FD)	Y/y with acute	Czech, Icelandic, Turkmen

TABLE 11.1. (continued)

Character (and Unicode number)	Description	Languages that use it
Ƴ (01B3), Ʒ (01B4)	Y/y with hook	Hausa
Ż (0179), ż (017A)	Z/z with acute	Polish, Montenegrin
Ž (017B), ž (017C)	Z/z with dot above	Polish
Ž (017D), ž (017E)	Z/z with caron (haček)	Bosnian, Croatian, Czech, Montenegrin, Serbian, Turkmen

## African Languages

- 11.23 **African capitalization and punctuation.** Most African languages—with the exception, most notably, of Arabic (see 11.76–81)—use the Latin alphabet and follow English capitalization and punctuation. The most widespread of these is Swahili, spoken by many different ethnic groups in eastern and central Africa. Hausa, Fulfulde, Yoruba, Igbo, Wolof, and Bambara are also spoken by millions, largely in western Africa; the same is true for Kikongo (or Kongo) and Lingala in the Congo-Zaire region and of Amharic and Somali in the Horn of Africa region. Amharic and other Ethiopian Semitic languages such as Tigrinya use the Ge'ez alphabet, not covered here. Xhosa and other “click” languages spoken in southern Africa do not follow English capitalization. The names of African languages themselves vary widely from ethnic group to ethnic group and from region to region. It is now standard practice to capitalize the names of African languages in the traditional way—for example, Kiswahili rather than KiSwahili or KISwahili. Xhosa speakers refer to and spell their language “isiXhosa” but “Isixhosa” (sometimes “Isizhosa”) is also found in English-language publications.
- 11.24 **African special characters.** Swahili uses no additional letters or diacritics. Among the more than two thousand other African languages, however, many rely on diacritics and phonetic symbols to stand for sounds that cannot be represented by letters or combinations of letters. Hausa, which is spoken by millions of people across western Africa, requires the following special characters (see also table 11.1):

B̄b, D̄d, K̄k, ƳƷ

In Nigeria, both the upper- and the lowercase *y* with a “hook” are represented instead with an apostrophe (’Y’*y*). Additional diacritics, too numerous to be listed here, may be needed in other African languages.

Languages such as French, Portuguese, and Arabic that are used in Africa are addressed in separate sections in this chapter.

### *French*

- 11.25 **French—additional resources.** As is the case with many languages, there is considerable variation in French publications with respect to capitalization and punctuation. For excellent advice, with frequent reference to the Académie française and numerous examples from literature, consult the latest edition of *Le bon usage*, known to many by the name of its original editor, Maurice Grevisse (bibliog. 5). Further guidance may be had at the website of the Académie française.
- 11.26 **French capitalization.** Generic words denoting roadways, squares, and the like are lowercased, whether used alone or with a specific name as part of an address. Only the proper name is capitalized.

le boulevard Saint-Germain  
la place de l'Opéra  
13, rue des Beaux-Arts

In most geographical names, the generic word is lowercased and the modifying word capitalized.

la mer Rouge      le pic du Midi

Names of buildings are usually capitalized.

l'Hôtel des Invalides      le Palais du Louvre

In names of organizations and institutions, only the first substantive and any preceding modifier are capitalized, but not the preceding article (except at the beginning of a sentence).

l'Académie française      la Légion d'honneur      le Grand Théâtre de Québec

In hyphenated names, both elements are capitalized.

la Comédie-Française      la Haute-Loire

Names of religious groups are usually lowercased.

un chrétien      des juifs

In names of saints, the word *saint* is lowercased. But when a saint's name is used as part of a place-name or the name of a church or other institution, *saint* is capitalized and hyphenated to the following element.

le supplice de saint Pierre      *but*      l'église de Saint-Pierre

Adjectives formed from proper nouns are usually lowercased.

une imagination baudelairienne

See also 11.18.

- 11.27 **Titles of French works.** French publications vary in the way they capitalize titles of works. In general, Chicago recommends sentence-style capitalization (see 8.158), the rule followed by Grevisse, *Le bon usage* (see 11.6, 11.25). Note that a superscript ordinal letter should remain in the superior position, as in the last example (cf. 14.88). An exception may be made for the French newspaper *Le Monde*, which always appears thus.

<i>L'Apollon de Bellac: Pièce en un acte</i>	<i>Le père Goriot</i>
<i>L'assommoir</i>	<i>Paris au XX<sup>e</sup> siècle</i>
<i>L'exil et le royaume</i>	<i>but</i>
<i>Les Rougon-Macquart</i>	<i>Le Monde</i>

According to an alternative practice advocated by the Académie française and others (and exemplified by the title *Le Monde*), for titles beginning with a definite article (*Le, La, L', Les*), the article and the first substantive (noun or noun form) and any intervening modifier are capitalized (e.g., *La Grande Illusion*). Titles that begin with a modifier are treated in the same way, with the modifier and first substantive capitalized (e.g., *Mauvais Sang*); any other titles, including those beginning with an indefinite article (*Un, Une*) are capitalized sentence-style (e.g., "Un cœur simple"). This style, if adopted for French titles, should be used consistently. For punctuation in titles, see 11.7.

- 11.28 **Spacing with French punctuation.** In French typeset material, fixed thin spaces generally occur before colons, semicolons, question marks, and exclamation marks; between guillemets (« ») and the text they enclose (see 11.29); and after an em dash used to introduce dialogue (see 11.31). In electronic documents, fixed (i.e., nonbreaking) spaces can be used to



avoid stranding a mark at the beginning of a line or, in the case of an opening guillemet, at the end (see 6.121). In an English context, the typographic conventions of the publication as a whole can be observed, and such spacing need not be duplicated. (If for any reason French spacing is required, however, it must be followed consistently and according to French practice for all marks.) See also 11.19.

- 11.29 **French use of guillemets.** For quotation marks, the French use guillemets (« »), often with a fixed thin space (or, especially in electronic documents, a regular nonbreaking space; see 6.121) to separate the guillemets from the quoted matter. If such guillemets are retained in an English context, as for a quotation in French (but see 11.19), they can usually be spaced like regular quotation marks (see also 11.28). Such tags as *écrit-il* or *dit-elle* are often inserted within the quoted matter without additional guillemets. Only punctuation belonging to the quoted matter is placed within the closing guillemets; other punctuation follows them.

«Mission accomplie?» a-t-il demandé.

En ce sens, «avec» signifie «au moyen de».

À vrai dire, Abélard n'avoue pas un tel rationalisme: «je ne veux pas être si philosophe, écrit-il, que je résiste à Paul, ni si aristotélicien que je me sépare du Christ».

As in English (see 13.32), when a quotation (other than a block quotation) continues for more than one paragraph, opening guillemets appear at the beginning of each additional paragraph; closing guillemets appear only at the end of the last paragraph. See also 11.30.

- 11.30 **Quotation marks in French.** For quotations within quotations, double (or sometimes single) quotation marks are used. Formerly, additional guillemets were used, with opening guillemets repeated on each runover line. (Note that when guillemets are used, if the two quotations end simultaneously, only one set of closing guillemets appears.) See also 11.29.

«Comment peux-tu dire, “Montre-nous le père”?»

Regular quotation marks are sometimes seen in French contexts in lieu of guillemets—especially in email correspondence and other electronic settings. This usage is considered informal.

- 11.31 **French dialogue.** In dialogue, guillemets are often replaced by em dashes. In French publications, the dash is usually followed by a thin space; in English publications, the space is not necessary (see 11.28).

Such dashes are used before each successive speech but are not repeated at the end of a speech. To set off a quotation within a speech, guillemets may be used. See also 11.29.

—Vous viendrez aussitôt que possible? a-t-il demandé.

—Tout de suite.

—Bien. Bonne chance!

—Tu connais sans doute la parole «De l'abondance du cœur la bouche parle».

—Non, je ne la connais pas.

- 11.32 **French ellipses.** The French often use an ellipsis to indicate an interruption or break in thought. An ellipsis is also sometimes used in lieu of *and so forth*. In French practice, an ellipsis consists of three unspaced dots closed up to the word they follow (*like this... rather than this . . .*); in English contexts, they may be spaced in the manner recommended elsewhere in this manual (see 13.50-58) and shown in the examples below. See also 11.19.

«Ce n'est pas que je n'aime plus l'Algérie . . . mon Dieu! un ciel! des arbres! . . . et le reste! . . . Toutefois, sept ans de discipline . . .»

To indicate omissions, the French use unspaced ellipses enclosed in brackets, with thin spaces between the brackets and the dots. In English contexts, spaced periods may be used (but with no space between the brackets and the periods they enclose; see 13.58).

«Oh, dit-elle avec un mépris écrasant, des changements intellectuels! [. . .]»  
Les deux amis se réunissaient souvent chez Luc [. . .].

- 11.33 **French word division—vowels.** In French, a word is divided after a vowel wherever possible. One-letter syllables at the ends or beginnings of lines should be avoided (see 11.20).

ache-ter (*not* a-cheter)      in-di-vi-si-bi-li-té      tri-age

Two or more vowels forming a single sound, or diphthong, are never broken.

écri-vain      fouet-ter      Gau-guin      éloi-gner      vieux

- 11.34 **French word division—consonants.** A division is normally made between two adjacent consonants, whether the same or different.

der-riè-re	Mal-raux	<i>but</i>
feuil-le-ter	ob-jet	qua-tre
ba-lan-cer	par-ler	ta-bleau

Groups of three adjacent consonants are normally divided after the first.

es-prit      res-plen-dir

- 11.35 **French words containing apostrophes.** Division should never be made immediately after an apostrophe.

jus-qu'au      au-jour-d'hui

- 11.36 **French words best left undivided.** Since there are as many syllables in French as there are vowels or diphthongs (even if some are unsounded except in poetry), the French break words that appear to English speakers to be of only one syllable (e.g., *fui-te*, *guer-re*, *sor-tent*). French practice also permits division after one letter (e.g., *é-tait*). In English-language publications, however, such breaks should be avoided, since they may confuse readers not fluent in French. Words of four or fewer letters should in any case be left undivided. See also 7.37.

- 11.37 **French accents and ligatures.** French employs the following special characters (see also table 11.1):

À à, Â â, Ç ç, É é, Ê ê, Ê ê, Ë ë, Î î, Ï ï, Ô ô, Œ œ, Ù ù, Û û, Ü ü

Although French publishers have often omitted accents on capital letters (especially *A*) and may set the ligature *Œ* as two separate letters (*OE*), all the special characters needed for French—including capitalized forms—are widely available, and they should be retained wherever needed in English-language contexts. This practice, advocated by the Académie française, is especially helpful to readers who may not be familiar with French typographic usage.

## *German*

- 11.38 **The new German orthography.** The new rules for German orthography (including spelling and capitalization) adopted in 1998 and made mandatory for schools and public documents in 2005 (subject to certain revisions) have been controversial. Some publications have continued to follow traditional rules, or a combination of house style and traditional

rules, whereas others have adopted the new rules. Some book publishers honor the preference of their authors and, by a similar token, do not update spelling when reprinting older works. Material quoted from German should therefore reflect the spelling in the source. For principles and details of the new orthography, consult the latest edition of *Duden: Die deutsche Rechtschreibung* (bibliog. 5). The recommendations and examples in this section reflect the new orthography.

- 11.39 **German capitalization.** In German, all nouns and words used as nouns are capitalized, whether in ordinary sentences or in titles of works (see 11.6).

ein Haus	Deutsch (the German language)
die Weltanschauung	eine Deutsche (a German woman)
das Sein	etwas Schönes

Adjectives derived from proper names are generally lowercased. Exceptions include invariable adjectives ending in *er* (often referring to a city or region) and adjectives that themselves are part of a proper name. For further exceptions, consult *Duden* (see 11.38).

die deutsche Literatur  
 nordamerikanische Sprachen  
 die platonischen Dialoge  
*but*  
 eine berühmte Berliner Straße  
 der Nahe Osten  
 der Deutsch-Französische Krieg

The pronouns *Sie*, *Ihr*, and *Ihnen*, as polite second-person forms, are capitalized. As third-person pronouns they are lowercased. The familiar second-person forms *du*, *dich*, *dein*, *ihr*, *euch*, and so on—once routinely capitalized—are now lowercased.

- 11.40 **German apostrophes.** An apostrophe is used to denote the colloquial omission of *e*.

wie geht's      was gibt's      hab' ich

Although an apostrophe rarely appears before a genitive *s*, an apostrophe is used to denote the omission of the *s* after proper names ending in an *s* sound (*ce*, *s*, *ss*, *ß*, *tz*, *x*, or *z*) or in a silent *s*, *x*, or *z*.

Alice' Geburtstag	Cixous' Theaterstücke
Jaspers' Philosophie	Leibniz' Meinung

- 11.41 **German quotation marks.** In German, quotations usually take reversed guillemets (» «); split-level inverted quotation marks („ “); or, in Switzerland, regular guillemets (see 11.29). Other punctuation is placed outside the closing quotation marks unless it belongs to the quoted matter.

*Eros* bedeutet für sie primär »zusammen-sein mit« und nicht »anschauen«. Denn: „An die Pferde“, hieß es: „Aufgesessen!“

- 11.42 **German word division—vowels.** In German, division is made after a vowel wherever possible. See also 11.20.

Fa-brik	hü-ten	Bu-ße
---------	--------	-------

Two vowels forming a single sound, or diphthong, are never broken.

Lau-ne	blei-ben
--------	----------

Further, a break should never be made after a single vowel at the beginning or end of a word (*aber, Ofen, Treue*).

- 11.43 **German word division—consonants.** Two or more adjacent consonants, whether the same or different, are divided before the last one unless they belong to different parts of a compound (see also 11.20).

klir-ren	Meis-ter
Was-ser	<i>but</i>
Verwand-te	Morgen-stern

The consonant combinations *ch, ck, ph, sch,* and *th* are not divided unless they belong to separate syllables. (Until the 1998 spelling change, *st* was subject to this rule. The combination *ck*, on the other hand, used to be changed at the end of a line to *kk* and divided between the *k*'s.)

Mäd-chen	<i>but</i>
Zu-cker	Klapp-hut
Philo-so-phia	Häus-chen
rau-schen	

- 11.44 **German word division—compounds.** Compound words should be divided between their component elements whenever possible (see also 11.20).

Meeres-ufer	Rasier-apparat
mit-einander	Tür-angel

- 11.45 **German special characters.** For setting German in roman type (the old Gothic or Fraktur type having long been out of use), the eszett, or sharp s ( $\beta$ ), and three unlauded vowels are needed (see also table 11.1).

Ää, Öö, ß, Üü

Although unlauded vowels are occasionally represented by omitting the accent and adding an *e* (*ae*, *Oe*, etc.), the availability of unlauded characters in text-editing software makes such a practice unnecessary. The eszett ( $\beta$ ), also widely available, must not be confused with, or replaced by, the Greek beta ( $\beta$ ). In the new spelling it is replaced by *ss* in certain words. Consult a German dictionary published after 1998. In German-speaking areas of Switzerland, the eszett is rarely used.

## *Italian*

- 11.46 **Italian capitalization.** In Italian, a title preceding a proper name is normally lowercased.

il commendatore Ugo Emiliano      la signora Rossi

In commercial correspondence, the formal second-person pronouns are capitalized in both their nominative forms, *Lei* (singular) and *Voi* (plural), and their objective forms, *La* (accusative singular), *Le* (dative singular), and *Vi* (accusative and dative plural). The older singular and plural forms *Ella* (*Le*, *La*) and *Loro* (*Loro*, *Loro*) are handled the same way. These pronouns are capitalized even in combined forms.

Posso pregarLa di farmi una cortesia?

Vorrei darLe una spiegazione.

See also 11.6, 11.18. For a fuller treatment of this and other matters of style, consult Roberto Lesina, *Il nuovo manuale di stile* (bibliog. 5).

- 11.47 **Italian quotations and dialogue.** Italian uses guillemets (« ») to denote quoted matter, but usually without the space between guillemets and quoted text that appears in many French publications. Regular quotation marks (double or single) are also frequently used in Italian—sometimes as scare quotes (see 7.57) in the same text in which guillemets are used for quotations. Note that periods and commas are correctly placed *after* the closing guillemet or quotation mark.

«Cosa pensi del fatto che io possa diventare “un qualcosa di imperial regio”?  
Questo non è proprio possibile».

In dialogue, em dashes are sometimes used, as in French. The dash is used before each successive speech. Unlike in French, however, another dash is used at the end of the speech if other matter follows in the same paragraph. The spaces that typically surround the dashes in Italian texts need not be used in English contexts (see 11.19).

—Avremo la neve,—annunziò la vecchia.

—E domani?—chiese Alfredo, voltandosi di scatto dalla finestra.

- 11.48 **Italian apostrophes.** An apostrophe is used to indicate the omission of one or more letters. A space should appear after an apostrophe that follows a vowel; after an apostrophe that follows a consonant, however, *no* space should appear.

po' duro      de' malevoli      l'onda      all'aura

- 11.49 **Italian ellipses.** Italian, like French (see 11.32), uses ellipses to indicate interruptions or breaks in thought. To indicate omitted material, the dots are enclosed in brackets. Though Italian typography usually calls for unspaced dots, in English publications Chicago recommends spaced periods wherever ellipses occur (see 13.50–58). See also 11.19.

Voglio . . . quattro milioni.      Davvero? [. . .] Non ci avevo pensato.

- 11.50 **Italian word division—vowels.** In Italian, division is made after a vowel wherever possible. One-letter syllables at the ends or beginnings of lines should be avoided (see 11.20).

acro-po-li (*not* a-cropoli)      mi-se-ra-bi-le      ta-vo-li-no

Consecutive vowels are rarely divided, and two vowels forming a single sound, or diphthong, are never divided.

miei pia-ga Gio-van-ni Giu-sep-pe pau-sa gio-iel-lo

- 11.51 **Italian word division—consonants.** Certain consonant groups must never be broken: *ch, gh, gli, gn, qu, sc,* and *r* or *l* preceded by any consonant other than itself.

ac-qua-rio la-ghi pa-dre ri-flet-te-re  
fi-glio na-sce rau-che so-gna-re

Three groups of consonants, however, may be divided: double consonants; the group *cqu*; and any group beginning with *l, m, n,* or *r*.

bab-bo ac-qua cam-po den-tro  
af-fre-schi cal-do com-pra par-te

- 11.52 **Italian word division—words containing apostrophes.** Division should never be made immediately after an apostrophe (but see 11.48).

dal-l'accusa del-l'or-ga-no quel-l'uomo un'ar-te l'i-dea

- 11.53 **Italian special characters.** In Italian, the following special characters are required (see also table 11.1):

Àà, Èè, Éé, Ìì, Òò, Ùù

Although the grave accent on capitalized vowels is sometimes dropped, in stressed final syllables it must be retained to avoid confusion.

CANTÒ (he sang) CANTO (I sing) PAPÀ (daddy) PAPA (pope)

Especially in older works, an apostrophe is sometimes seen with a capital letter in place of the accent on a stressed final (or single) vowel. In direct quotations, such usage should be retained.

E' (it is) E (and) PAPA' (daddy)

## Latin

- 11.54 **Latin capitalization—titles of works.** Titles of ancient and medieval Latin works should usually be capitalized in sentence style—that is, only the first word in the title and subtitle, proper nouns, and proper adjectives are capitalized (see 8.158).



*De bello Gallico*      *De viris illustribus*      *Cur Deus homo?*

Renaissance and modern works or works in English with Latin titles, on the other hand, can usually be capitalized headline-style (see 8.159). (If there is any doubt about the era to which the title belongs, opt for sentence style.)

*Novum Organum*      *Religio Medici*

See also 11.6.

- 11.55 **Latin word division—syllables.** A Latin word has as many syllables as it has vowels or diphthongs (*ae, au, ei, eu, oe, ui*, and, in archaic Latin, *ai, oi, ou*) and should be divided between syllables (see also 11.20).

na-tu-ra      cae-li-co-la      in-no-cu-us

- 11.56 **Latin word division—single consonants.** When a single consonant occurs between two vowels, the word is divided before the consonant unless it is an *x*. Note that *i* and *u* sometimes act as consonants (and, when they do, are sometimes written as *j* and *v*).

Cae-sar      me-ri-di-es      in-iu-ri-or (or in-ju-ri-or)      *but*      lex-is

- 11.57 **Latin word division—multiple consonants.** When two or more consonants come together, the word is divided before the last consonant, except for the combinations noted below.

om-nis      cunc-tus

The combinations *ch, gu, ph, qu,* and *th* are treated as single consonants and thus never separated.

co-phi-nus      lin-gua      ae-qua-lis

The following consonant groups are never broken: *bl, br, chl, chr, cl, cr, dl, dr, gl, gr, phl, phr, pl, pr, thl, thr, tl,* and *tr*.

pan-chres-tus      li-bris      ex-em-pla      pa-tris

- 11.58 **Latin word division—compounds.** Compound words are divided between parts; within each part the rules detailed elsewhere in this section

apply. The commonest type of compound word begins with a preposition or a prefix (e.g., *ab-*, *ad-*, *in-*, *re[d]-*).

ab-rum-po      ad-est      red-eo      trans-igo

- 11.59 **Latin special characters.** Latin requires no special characters for setting ordinary copy. Elementary texts, however, usually mark the long vowels with a macron and, occasionally, the short vowels with a breve, as follows. (See also table 11.1.)

Āā, Ăă, Ēē, Ĕĕ, Īī, Ĭĭ, Ōō, Ŏö, Ūū, Ŭŭ

## Spanish

- 11.60 **Spanish—additional resources.** There is considerable variation in Spanish-language publications throughout the world with respect to capitalization, punctuation, and other matters. For further guidance, consult the extensive resources available from the Real Academia Española, including such essential guides as the *Diccionario panhispánico de dudas* and the *Ortografía de la lengua española* (bibliog. 5).
- 11.61 **Spanish capitalization.** In Spanish, a title preceding a proper name is normally lowercased. When abbreviated, however, titles are capitalized.

el señor Jaime López	<i>but</i>
la señora Lucía Moyado de Barba	el Sr. López
doña Perfecta	

Nouns as well as adjectives denoting membership in nations are lowercased, but names of countries are capitalized.

los mexicanos      la lengua española      Inglaterra

Names of organizations and institutions, historical events, buildings, streets, and the like are usually capitalized (see also 8.159).

Real Academia Español  
 Universidad Nacional Autónoma de México  
 Plaza del Dos de Mayo

See also 11.4, 11.6, 11.18.

- 11.62 **Spanish question marks and exclamation points.** A question or an exclamation in Spanish is preceded by an inverted question mark or exclamation point and followed by a regular mark.

¿Qué pasa, amigo?      ¡Olvidalo en ese caso!

If a vocative or dependent construction precedes a question or exclamation, it is written as follows:

Amigo, ¿qué pasa?      En ese caso, ¡olvidalo!

Because the opening marks are integral to Spanish punctuation, they should be retained even when Spanish is being quoted in an English context (see 11.19).

- 11.63 **Spanish guillemets and quotation marks.** Spanish traditionally uses guillemets (« ») as quotation marks. Only punctuation belonging to the quoted matter is placed within the closing guillemets; other punctuation follows them. Within a quotation, em dashes may be used to set off words identifying the speaker. In Spanish publications, the opening dash is usually *preceded* by a space; the closing dash is then *followed* by a space unless immediately followed by punctuation. In English contexts, such spaces need not be used (see also 11.19). (For quotations within quotations, regular quotation marks are used, as in French; see 11.30.)

«Vino el negocio a tanto—comenta Suárez—, que ya andaban muchos tomados por el diablo».

In lexical studies, it is typical to see single quotation marks used for glosses, with no punctuation preceding the gloss (cf. 11.5).

Muchos adverbios se forman añadiendo *-ly* al adjetivo: *courteous* ‘cortés’, *courteously* ‘cortésmente’, *bold* ‘atrevido’, *boldly* ‘atrevidamente’.

Increasingly, Spanish-language publications use regular quotation marks rather than guillemets for all quotations. Where this is the case, the rules for punctuation marks relative to the quotation marks are the same as they are for guillemets (but see 11.11).

- 11.64 **Spanish dialogue.** In dialogue, an em dash (or, less frequently, a guillemet) introduces each successive speech. Any other matter that follows the quoted speech in the same paragraph is generally preceded by a dash or a comma. See also 11.63.

—Esto es el arca de Noé, afirmó el estanciero.  
 —¿Por qué estás aquí todavía?—preguntó Juana alarmada.

- 11.65 **Spanish ellipses.** In Spanish, as in French (see 11.32), ellipses are used to indicate interruptions or breaks in thought. In Spanish publications, these dots are generally unspaced; in English contexts, they may be spaced as recommended elsewhere in this manual (see 13.50–58). To indicate omitted material, the dots are enclosed in brackets. See also 11.19.

Hemos comenzado la vida juntos . . . quizá la terminaremos juntos también . . .  
 La personalidad más importante del siglo XIX es Domingo Faustino Sarmiento [. . .], llamado el hombre representante del intelecto sudamericano.  
 [. . .] El gaucho [. . .] servía de tema para poemas, novelas, cuentos y dramas.

- 11.66 **Spanish word division—vowels.** In Spanish, division is made after a vowel whenever possible. See also 11.20.

ca-ra-co-les      mu-jer      re-cla-mo      se-ño-ri-ta

Two or more vowels that form a single syllable (a diphthong or a triphthong) may not be divided.

cam-bias      fue-go      miau      tie-ne      viu-da

If adjacent vowels belong to separate syllables, however, they are divided between syllables.

ba-úl      cre-er      pa-ís      te-a-tro

- 11.67 **Spanish word division—consonants.** If two adjacent consonants form a combination that would generally not occur at the beginning of a Spanish word, the break is made between them.

ac-cio-nis-ta      ad-ver-ten-cia      al-cal-de      an-cho      efec-to      is-leño

The consonant groups *bl*, *br*, *cl*, *cr*, *dr*, *fl*, *fr*, *gl*, *gr*, *pl*, *pr*, and *tr*—all pairs that can occur at the beginning of Spanish words—are inseparable (unless each belongs to a different element of a compound, as in *sub-lu-nar*; see 11.68, 11.20).

ci-fra      li-bro      no-ble      re-gla  
 co-pla      ma-dre      pa-tria      se-cre-to  
 im-po-si-ble      ne-gro      re-fle-jo      te-cla  
 le-pra

Groups of three consonants not ending with one of the inseparable pairs listed above always have an *s* in the middle. They are divided after the *s*.

cons-pi-rar      cons-ta      ins-tan-te      obs-cu-ro      obs-tan-te

Spanish *ch* and *ll* were long considered single characters, alphabetized as such, and never divided. The Spanish Royal Academy has now declared that these combinations are to be alphabetized as two-letter groups, and new publications have adopted this convention. Along with *rr*, however, they still cannot be divided, since they represent single sounds. For details, consult Real Academia Española, *Ortografía de la lengua española* (bibliog. 5).

ci-ga-rri-llo      mu-cha-cho

- 11.68 **Dividing Spanish compounds.** Compound words are often but not always divided between their component parts.

des-igual      mal-es-tar      semi-es-fe-ra      sub-lu-nar  
in-útil      trans-al-pi-no      bien-aven-tu-ra-do      sub-ra-yar  
*but*

no-so-tros (no longer considered a compound by Spanish speakers)

- 11.69 **Spanish special characters.** Spanish employs the following special characters (see also table 11.1):

Áá, Éé, Íí, Ññ, Óó, Úú, Üü

### *Other Languages Using the Latin Alphabet*

- 11.70 **Special considerations for other languages using the Latin alphabet.** In addition to the languages covered elsewhere in this section, there are dozens of other languages that use the Latin alphabet. Special considerations for a number of them are listed below. For the special characters required for each of these languages, see table 11.1. See also 11.6, 11.18.

**Albanian.** Since 1972, Albanian has had a single, unified orthography, based on a standard originally adopted in 1909. Writers and editors working with older texts may need to take historical context into account and determine whether a spelling is conditioned by the specific time when it was used or whether it is preferable to follow the current norm.

**Croatian and Bosnian.** The former Serbo-Croatian language used both Latin and Cyrillic alphabets. The modern Bosnian and Croatian standard languages use only the Latin version of that same alphabet. Although the substitution of *dj* for *đ* is sometimes seen (e.g., in informal correspondence), standard orthographic practice in all the successor languages of Serbo-Croatian distinguishes these two consistently. See also Serbian and Montenegrin.

**Czech,** a Slavic language written in the Latin alphabet, uses many diacritical marks to indicate sounds not represented by this alphabet, as shown in table 11.1. Note that the lowercase *d* and *t* with caron (the single glyphs *ď* and *ť*, respectively) are often seen with an apostrophe instead.

**Danish.** The polite second-person pronouns *De*, *Dem*, and *Deres* (increasingly rare, and not to be confused with the third-person pronouns *de*, *dem*, and *deres*) and the familiar *I* are capitalized in Danish. Until the middle of the twentieth century, common nouns were capitalized, as in German.

**Dutch.** For the capitalization of particles with personal names, see 8.10. Proper adjectives (as well as nouns) are capitalized as in English. When a word beginning with the diphthong *ij* is capitalized, both letters are capitals: *IJsland*. When a single letter begins a sentence, it is lowercased, but the next word is capitalized: *'k Heb niet . . .*

**Finnish.** Because Swedish is the second official language in Finland, the Finnish alphabet taught in schools and the standard keyboard used in Finland include the Swedish *a* with ring above (see table 11.1).

**Hawaiian.** The Hawaiian alphabet was developed in the nineteenth century from the Latin alphabet. In addition to the five vowels with macrons listed in table 11.1, Hawaiian uses the *'okina*, a glottal stop represented by a left single quotation mark (U+2018)—for example, in the place-name *Hawai'i*. See also 6.115.<sup>1</sup>

**Hungarian** uses a wide variety of accented vowels, as shown in table 11.1.

**Icelandic** includes the consonants *Ðð* (eth) and *Þþ* (thorn), which were also used in Old and Middle English (see 11.122–24). (The eth, which never begins a word, is capitalized only in contexts where all capitals are used.) In addition to featuring an acute-accented version of each regular vowel (including *Ýý*), Icelandic includes the vowels *Ææ* and *Öö*.

---

1. A modifier letter turned comma (U+02BB) may be used instead of a left single quotation mark to represent the *'okina* and is preferred by some authors and publishers as a means of differentiating the glottal stop from the common mark of punctuation. In many typefaces, however, the two glyphs have an identical appearance.

**Norwegian.** The polite second-person pronouns *De*, *Dem*, and *Deres* (increasingly rare, and not to be confused with the third-person pronouns *de*, *dem*, and *deres*) are capitalized in Norwegian. Until the middle of the twentieth century, common nouns were capitalized, as in German.

**Polish.** In formal address the second-person plural pronoun *Państwo* (you) is capitalized, as are related forms: *Czekam na Twój przyjazd* (I await your arrival); *Pozdrawiam Cię!* (Greetings to you!). Division of Polish words is similar to that of transliterated Russian (see 11.98–108). Division normally follows syllabic structure (e.g., *kom-pi-ka-cja*; *sta-ro-pol-ski*). Note that the conjunction *i* (and) should never appear at the end of a line but must be carried over to the beginning of the next.

**Portuguese.** Titles and nouns or adjectives denoting nationality are capitalized as in Spanish (see 11.61). Accented capitals, sometimes dropped in Portuguese running text, should always be used when Portuguese is presented in an English context.

**Romanian and Moldovan** are now both written using the same Latin orthography. Note that *Șș* and *Țț*—Latin *Ss* and *Tt* with comma below—often appear instead with a cedilla, though the comma is correct. *Â â* and *Î î* represent identical sounds but have different etymological origins. The use of *Â â* has been restricted, eliminated, and reinstated in whole or in part during various orthographic reforms. Writers and editors, therefore, should take care to determine whether a spelling is conditioned by the specific time when it was used or whether it is preferable to follow the current norm.

**Serbian and Montenegrin.** The former Serbo-Croatian language used both Latin and Cyrillic alphabets. In the modern Montenegrin standard language, both versions of that alphabet are official. In the modern Serbian standard language, the Cyrillic version of that same alphabet is official, though the Latin alphabet is also used, as regulated by law. Note that although the substitution of *dj* for *đ* is sometimes seen (e.g., in informal correspondence), standard orthographic practice in all the successor languages of Serbo-Croatian distinguishes these two consistently. In addition to the letters needed for Serbian, two extra letters are required for Montenegrin: *Š š* and *Ž ž* (see table 11.1). See also Croatian and Bosnian.

**Swedish.** In Swedish, the second-person pronouns *Ni* and *Er*, traditionally capitalized in correspondence, are now lowercased in all contexts.

**Turkish and Azeri.** Modern Turkish has undergone a number of orthographic reforms since the original change to the Latin alphabet in 1928. Differences in the spellings of a name or word can therefore depend on the time period. Writers and editors should take care to

determine whether a spelling is conditioned by the specific time when it was used or whether it is preferable to follow the current norm. In Turkish, as in English, the names of months and days of the week are capitalized. The Azeri (Azerbaijani) standard alphabet in use since 1992 is identical to the Turkish alphabet except for the presence of *ə*, *Qq*, and *Xx* (lacking in Turkish), and the absence of vowels with circumflex. Conventions for capitalization and spelling are similar to those for Turkish. Note that in both languages, the letter *i* retains its dot when capitalized.

**Turkmen and Uzbek.** Turkmenistan has successfully transitioned from a Cyrillic to a Latin alphabet. In Uzbekistan the transition is still ongoing. Uzbek requires no special characters aside from the left single quotation mark in the letters *O'o'* and *G'g'* (not shown in table 11.1).

## Languages Usually Transliterated (or Romanized)

- 11.71 **Transliteration.** In nonspecialized works it is customary to transliterate—that is, convert to the Latin alphabet, or romanize—words or phrases from languages that do not use the Latin alphabet. For discussion and illustration of scores of alphabets, see Peter T. Daniels and William Bright, eds., *The World's Writing Systems* (bibliog. 5). For alphabetic conversion, the most comprehensive resource is the Library of Congress publication *ALA-LC Romanization Tables* (bibliog. 5), available online. Do not attempt to transliterate from a language unfamiliar to you. Note that the recommendations elsewhere in this chapter related to capitalization (11.18), punctuation (11.19), and word division (11.20) for languages that use the Latin alphabet apply equally to transliterated text.
- 11.72 **Character sets for non-Latin alphabets.** Modern word-processing software readily allows users to enter words in a number of non-Latin alphabets. For a given alphabet, there may be a variety of non-Unicode character sets available as specialized fonts, but authors who want to include such copy should generally opt for a font that includes the correct Unicode characters if at all possible (see 11.2), after consulting their publisher. See also 2.16.
- 11.73 **Proofreading copy in non-Latin alphabets—a warning.** Anyone unfamiliar with a language that uses a non-Latin alphabet should exercise extreme caution in proofreading even single words set in that alphabet. Grave errors can occur when similar characters are mistaken for each



other. If in doubt, editors should query the author; it may be advisable to consult the Unicode number and description (see 11.2) when referring to a given character or diacritical mark.

- 11.74 **Diacritics—specialized versus general contexts.** Nearly all systems of transliteration require diacritics—including, in the languages discussed below, macrons, underdots, and overdots, to name just a few. Except in linguistic studies or other highly specialized works, a system using as few diacritics as are needed to aid pronunciation is easier on readers, publisher, and author. Most readers of a nonspecialized work on Hindu mythology, for example, will be more comfortable with Shiva than Śiva or with Vishnu than Viṣṇu, though many specialists would want to differentiate the *Sh* in Shiva from the *sh* in Vishnu as distinct Sanskrit letters. For nonspecialized works, the transliterated forms without diacritics that are listed in the latest editions of the Merriam-Webster dictionaries (bibliog. 3.1) are usually preferred by readers and authors alike.
- 11.75 **Italics versus roman for transliterated terms.** Transliterated terms (other than proper names) that have not become part of the English language are italicized. If used throughout a work, a transliterated term may be italicized on first appearance and then set in roman. Words listed in the dictionary are usually set in roman. See also 11.3–5.

The preacher pointed out the distinction between agape and eros.

*but*

Once the Greek words *erōs* and *agapē* had been absorbed into the English language, it became unnecessary to italicize them or to use the macrons.

## *Arabic*

- 11.76 **Arabic transliteration.** There is no universally accepted form for transliterating Arabic. One very detailed system may be found in the *ALA-LC Romanization Tables* (bibliog. 5). Another system is followed by the *International Journal of Middle East Studies* (bibliog. 5). Having selected a system, an author should stick to it with as few exceptions as possible. In the following examples, only the hamza (ʾ) and the ʿayn (ع) are used (see 11.77). Letters with underdots and some of the other special characters used in transliteration from Arabic are included in table 11.2. (The Arabic alphabet may be found in the alphabet table in *Merriam-Webster's Collegiate Dictionary* [bibliog. 3.1], among other sources.)

TABLE 11.2. Special characters (and Unicode numbers) for transliterated Arabic, Hebrew, Japanese, and South Asian languages

Character (and Unicode number)	Description	Languages that use it
´ (02B9)	modifier letter prime (see 11.92)	Arabic, Hebrew
˘ (02BF)	ʿayn <i>or</i> ʿayin (modifier letter left half ring)	Arabic, Hebrew
˙ (02BE)	alif (hamza) <i>or</i> ʾalef (modifier letter right half ring)	Arabic, Hebrew
Ā (0100), ā (0101)	A/a with macron	Arabic, Hebrew, Japanese, South Asian languages
Ǻ (01CD), ǻ (01CE)	A/a with caron (haček)	Hebrew
Á (00C1), á (00E1)	A/a with acute	Arabic
Ḑ (1E0C), ḑ (1E0D)	D/d with dot below	Arabic, South Asian languages
Ē (0112), ē (0113)	E/e with macron	Hebrew, Japanese, South Asian languages
Ě (011A), ě (011B)	E/e with caron (haček)	Hebrew
ə (0259)	small schwa	Hebrew
Ḥ (1E24), ḥ (1E25)	H/h with dot below	Arabic, Hebrew, South Asian languages
Ī (012A), ī (012B)	I/i with macron	Arabic, Hebrew, Japanese, South Asian languages
Ḷ (1E32), ḷ (1E33)	K/k with dot below	Arabic, Hebrew
Ḹ (1E36), ḹ (1E37)	L/l with dot below <sup>1</sup>	South Asian languages
Ḻ (1E38), ḻ (1E39)	L/l with dot below and macron	South Asian languages
Ṁ (1E40), ṁ (1E41)	M/m with dot above	South Asian languages
Ṃ (1E42), ṃ (1E43)	M/m with dot below	South Asian languages
Ñ (00D1) ñ (00F1)	N/n with tilde	South Asian languages
Ñ̄ (004E+0304), ñ̄ (006E+0304)	N/n with macron (combining character)	South Asian languages
Ṇ (1E44), ṇ (1E45)	N/n with dot above	South Asian languages
Ṋ (1E46), ṋ (1E47)	N/n with dot below	South Asian languages
Ō (014C), ō (014D)	O/o with macron	Hebrew, Japanese, South Asian languages
Ǫ (01D1), ǫ (01D2)	O/o with caron (haček)	Hebrew
Ṛ (1E5A), ṛ (1E5B)	R/r with dot below <sup>2</sup>	South Asian languages
Ṝ (1E5C), ṝ (1E5D)	R/r with dot below and macron	South Asian languages
Ś (015A), ś (015B)	S/s with acute	Hebrew, South Asian languages
Š (0160), š (0161)	S/s with caron (haček)	Hebrew
Ş (1E62), ş (1E63)	S/s with dot below	Arabic, South Asian languages
Ṭ (1E6C), ṭ (1E6D)	T/t with dot below	Arabic, Hebrew, South Asian languages
Ū (016A), ū (016B)	U/u with macron	Arabic, Hebrew, Japanese, South Asian languages
Ṽ (1E7E), ṽ (1E7F)	V/v with dot below	Hebrew
Ỳ (1E8E), ỳ (1E8F)	Y/y with dot above	South Asian languages
Ẑ (1E92), ẑ (1E93)	Z/z with dot below	Arabic

<sup>1</sup>Variations of L/l with a combining ring below (U+0325) rather than a dot below may also be required.

<sup>2</sup>As with L/l, R/r variations may require a combining ring below (U+0325) rather than a dot.

- 11.77 **The hamza and the ‘ayn.** The hamza (ʾ) and the ‘ayn (ع) frequently appear in transliterated Arabic words and names. Writers using hamzas or ‘ayns must on every occurrence make it clear, by coding or by careful instructions to the editor or typesetter, which of the two marks is intended. It should be noted that the Arabic characters are not the same as the ones used for transliteration; see table 11.2 for the preferred Unicode characters for hamza and ‘ayn in transliteration. The hamza is sometimes represented—especially in nonspecialized works—by an apostrophe, as in Qur’an, and the ‘ayn by a single opening quotation mark (‘ayn). (Since an ‘ayn often occurs at the beginning of a word, a quotation mark must be used with caution.) Most transliteration systems drop the hamza when it occurs at the beginning of a word (anzala *not* ’anzala). See also 6.115, 6.117.
- 11.78 **Arabic spelling.** Isolated references in text to well-known persons or places should employ the forms familiar to English-speaking readers.

Avicenna (*not* Ibn Sina)  
 Damascus (*not* Dimashq)  
 Mecca (*not* Makka *or* Makkah)

- 11.79 **The Arabic definite article.** Though there is considerable variation across publications, Chicago recommends joining the Arabic definite article, *al*, to a noun with a hyphen.

al-Islam      al-Nafud      Bahr al-Safi      al-Qaeda (*or* al-Qaida)

In speech the sound of the *l* in *al* is assimilated into the sounds *d*, *n*, *r*, *s*, *sh*, *t*, and *z*. Where rendering the *sound* of the Arabic is important (for example, when transliterating poetry), the assimilations are often shown, as in the examples below. In most other situations, the article-noun combination is written without indication of the elision, as above.

an-Nafud      Bahr as-Safi

Some authors drop the *a* in *al* and replace it with an apostrophe when it occurs after a long syllable (Abū ’l-Muhallab). Some also drop the *a* when it occurs connected with a particle (wa ’l-layl). Others do not replace the dropped *a* with anything (Abū l-Muhallab; wa l-layl).

- 11.80 **Arabic capitalization.** Since the Arabic alphabet does not distinguish between capital and lowercase letter forms, practice in capitalizing transliterated Arabic varies widely. Chicago recommends the practice outlined

in 11.6: capitalize only the first word and any proper nouns. This practice applies to titles of works as well as to names of journals and organizations. Note that *al*, like *the*, is capitalized only at the beginning of a sentence or a title. See also 11.9.

‘Abd al-Rahman al-Jabarti, *‘Aja’ib al-athar fi al-tarajim wa al-akhbar* (The marvelous remains in biography and history)

For citing and alphabetizing Arabic personal names, see 8.14, 16.75.

- 11.81 **Arabic word division.** Breaking transliterated Arabic words or names at the ends of lines should be avoided wherever possible. If necessary, a break may be made after *al* or *Ibn*. A break may be made after two letters if the second has an underdot (e.g., *it-baq*). Breaks must never be made between the digraphs *dh*, *gh*, *kh*, *sh*, or *th* unless both letters have underdots. Nor should breaks be made before or after a hamza. Aside from these niceties, the rules governing English word division may be followed (see 7.36–47). It should be noted, however, that untransliterated (or unromanized) Arabic is read from right to left; if a line break occurs within an untransliterated Arabic phrase, the words must still be read right to left on each line. For an example of this in Hebrew, see 11.96.

## Chinese and Japanese

- 11.82 **Chinese romanization.** The Hanyu Pinyin romanization system, introduced in the 1950s, has largely supplanted both the Wade-Giles system and the place-name spellings of the *Postal Atlas of China* (last updated in the 1930s), making Pinyin the standard system for romanizing Chinese. Representing sounds of Chinese more explicitly, Pinyin has been widely accepted as the system for teaching Chinese as a second language. As of 2000, the Library of Congress issued new romanization guidelines reflecting the conversion of its entire online catalog records for the Chinese collection to comply with Pinyin. Although a few authors, long familiar with Wade-Giles or other older systems (or Tongyong Pinyin, a more recent system still used by some in Taiwan), have not switched to Pinyin in their writings, Chicago joins librarians in urging that Pinyin now be used in all writing about China or the Chinese language. (In some contexts it may be helpful to the reader to add the Wade-Giles spelling of a name or term in parentheses following the first use of the Pinyin spelling.) The *ALA-LC Romanization Tables* (bibliog. 5) available online from the Library of Congress should be used with caution by anyone unfamiliar with Chinese.

- 11.83 **Exceptions to Pinyin.** Even where Pinyin is adopted, certain place-names, personal names, and other proper nouns long familiar in their older forms may be presented that way in English texts. Or, for greater consistency, the old spelling may be added in parentheses after the Pinyin version. If in doubt, consult the latest edition of *Merriam-Webster's Collegiate Dictionary* (bibliog. 3.1); names not listed there in older forms should be presented in Pinyin. Editors who wish to alter spellings should do so in consultation with the author.
- 11.84 **Apostrophes, hyphens, and tone marks in Chinese romanization.** Pinyin spellings often differ markedly from Wade-Giles and other older spellings. Personal names are usually spelled without apostrophes or hyphens, but an apostrophe is sometimes used when syllables are run together (as in Xi'an to distinguish it from Xian), even in contexts where tone marks are used (e.g., Xi'an). The Pinyin romanization system of the Library of Congress does not include tone marks, nor are they included in many English-language publications. However, tone marks may be appropriate in certain contexts (e.g., textbooks for learning Chinese).
- 11.85 **Some common Chinese names.** Some names frequently encountered are listed below.

DYNASTIES		PERSONAL NAMES	
<i>Wade-Giles</i>	<i>Pinyin</i>	<i>Wade-Giles</i>	<i>Pinyin</i>
Chou	Zhou	Fang Li-chih	Fang Lizhi
Ch'in	Qin	Hua Kuo-feng	Hua Guofeng
Ch'ing	Qing	Lin Piao	Lin Biao
Sung	Song	Lu Hsün	Lu Xun
T'ang	Tang	Mao Tse-tung	Mao Zedong
Yüan	Yuan	Teng Hsiao-p'ing	Deng Xiaoping

The names Sun Yat-sen and Chiang Kai-shek, among a few others, usually retain the old spellings.

#### GEOGRAPHICAL NAMES

<i>Wade-Giles</i>	<i>Postal atlas</i>	<i>Pinyin</i>
Kuang-tung	Kwangtung	Guangdong
Pei-ching (Pei-p'ing)	Peking (Peiping)	Beijing
Shang-hai	Shanghai	Shanghai
Su-chou	Soochow	Suzhou
Ta-lien	Dairen	Dalian

- 11.86 **Japanese romanization.** The Japanese language in its usual written form is a mixture of Chinese characters (called *kanji* in Japanese) and two *kana* syllabaries. (A syllabary is a series of written characters, each used to represent a syllable.) Since romanized Japanese, *rōmaji*, was introduced into Japan in the sixteenth century, a number of systems of romanization have been developed. The one in most common use since the early part of the Meiji period (1868–1912) is the modified Hepburn (or *hyōjun*) system. This system is used in *Kenkyūsha's New Japanese–English Dictionary* (bibliog. 3.2) and most other Japanese–English dictionaries (and is the basis of the Japanese romanization tables available online from the Library of Congress); outside Japan, it is also used almost exclusively, notably in Asian collections in libraries throughout the world.
- 11.87 **Modified Hepburn system.** In the modified Hepburn system, an apostrophe is placed after a syllabic *n* that is followed by a vowel or *y*: *Gen'e*, *San'yo*. A macron is used over a long vowel (usually an *o* or a *u*, though some systems allow for macrons over *a*, *i*, and *e*) in all Japanese words except well-known place-names (e.g., Tokyo, Hokkaido, Kobe) and words such as *shogun* and *daimyo* that have entered the English language and are thus not italicized. (When the pronunciation of such names or words is important to readers, however, macrons may be used: Tōkyō, Hokkaidō, Kōbe, shōgun, daimyō.) Hyphens should be used sparingly: *Meiji jidai-shi* (or *jidaiishi*) *no shinkenkyū*. *Shinjuku-ku* (or *Shinjukuku*) *no meisho*.
- 11.88 **Chinese and Japanese—capitalization and italics.** Although capital letters do not exist in Japanese or Chinese, they are introduced in romanized versions of these languages where they would normally be used in English (see chapter 8). Personal names and place-names are capitalized. In hyphenated names, only the first element is capitalized in romanized Chinese, though both elements may be capitalized in Japanese. Common nouns and other words used in an English sentence are lowercased and italicized (see 11.3, 11.5). Names of institutions, schools of thought, religions, and so forth are capitalized if set in roman, lowercased if set in italics.

Donglin Academy; the Donglin movement

Buddhism, Taoism, feng shui [see 7.54], and other forms . . .

Under the Ming dynasty the postal service was administered by the Board of War (*bingbu*) through a central office in Beijing (*huitong guan*).

The heirs of the Seiyūkai and Minseitō are the Liberal and Progressive Parties of Japan.

It was Genrō Saionji (the *genrō* were the elder statesmen of Japan) who said . . .  
(note that *genrō* is both singular and plural)

- 11.89 **Titles of Japanese and Chinese works.** As in English, titles of books and periodicals are italicized, and titles of articles are set in roman and enclosed in quotation marks (see 8.156–201). The first word of a romanized title is always capitalized, as are many proper nouns (especially in Japanese).

Chen Shiqi, *Mingdai guan shougongye de yanjiu* [Studies on government-operated handicrafts during the Ming dynasty], . . .

Hua Linfu, “Qingdai yilai Sanxia diqu shuihan zaihai de chubu yanjiu” [A preliminary study of floods and droughts in the Three Gorges region since the Qing dynasty], *Zhongguo shehui kexue* 1 (1999): 168–79.

Okamoto Yoshitomo, *Jūrokuseiki Nichi-Ō kōtsūshi no kenkyū* [Study of the intercourse between Japan and Europe during the sixteenth century], . . .

Akiyama Kenzō, “Goresu wa Ryūkyūjin de aru” [The Gores are Ryūkyūans], *Shigaku-Zasshi* (or *Shigaku Zasshi*) . . .

- 11.90 **Inclusion of Chinese and Japanese characters.** Chinese and Japanese characters, immediately following the romanized version of the item they represent, are sometimes necessary to help readers identify references cited or terms used. They are largely confined to bibliographies and glossaries. Where needed in running text, they may be enclosed in parentheses. The advent of Unicode has made it easier for authors to include words in non-Latin alphabets in their manuscripts, but publishers need to be alerted of the need for special characters in case particular fonts are needed for publication (see 11.2).

Harootunian, Harry, and Sakai Naoki. “Nihon kenkyū to bunka kenkyū” 日本研究と文化研究. *Shisō* 思想 7 (July 1997): 4–53.

Hua Linfu 華林甫. “Qingdai yilai Sanxia diqu shuihan zaihai de chubu yanjiu” 清代以來三峽地區水旱災害的初步研究 [A preliminary study of floods and droughts in the Three Gorges region since the Qing dynasty]. *Zhongguo shehui kexue* 中國社會科學 1 (1999): 168–79.

That year the first assembly of the national Diet was held and the Imperial Rescript on Education (*kyōiku chokugo* 教育勅語) was issued.

## Hebrew

- 11.91 **Hebrew transliteration systems.** There are several acceptable romanization systems for Hebrew, including the one in the *ALA-LC Romaniza-*

*tion Tables* (see bibliog. 5). Any such system may be used, but it is the author's responsibility to use it consistently in a given work. (The Hebrew alphabet may be found in the alphabet table in *Merriam-Webster's Collegiate Dictionary* [bibliog. 3.1], among other sources.)

- 11.92 **Diacritics in transliterated Hebrew.** In transliterated Hebrew, the following accents and characters are sometimes needed (though usually only in specialist materials): underdots (Ḥḥ, Ḷḷ, Ṭṭ, Ṽṽ); macrons (Āā, Ēē, Īī, Ōō, Ūū); acute accents (Śś); haček, or caron (Ǻǻ, Ǽǽ, Ǫǫ, Šš); and superscript schwa (◌◌̆). The 'alef and the 'ayin may be represented in the same way as the Arabic hamza and 'ayn (see 11.77 and table 11.2). In some systems, a prime may also be needed (to separate two distinct consonant sounds that might be mistaken for a digraph).
- 11.93 **Hebrew prefixes.** In Hebrew, several prepositions, conjunctions, and articles appear as prefixes. Some authors use apostrophes or hyphens after these prefixes in romanized text, and some do not. (In Hebrew no such marker is used.) Either approach is acceptable if used consistently.
- 11.94 **Hebrew capitalization and italics.** The Hebrew alphabet has no capital letters, and there is no universally used system for capitalizing romanized Hebrew. Writers may follow normal English usage—capitalizing proper names, book titles, and so forth (see 11.6, 11.18). Some writers eschew capitalization altogether. As always, the author must ensure internal consistency. For italics in romanized Hebrew, the normal English usage may also be followed (see 11.8).
- 11.95 **Hebrew word division.** For romanized Hebrew, or Hebrew words incorporated into English, the principles set forth in 7.36–47 may be followed. When a double consonant occurs at the point of division, one consonant goes with each division.

Rosh Ha-shana      Yom Kip-pur

- 11.96 **Unromanized Hebrew phrases.** Hebrew is read from right to left. In English sentences that contain an unromanized Hebrew phrase, the Hebrew order is maintained within the sentence. (Modern operating systems can often handle a mix of left-right and right-left input in the same context.)

The first phrase in Lamentations is איכה ישבה בדד (How she sits in solitude!).

If a line break occurs within a Hebrew phrase, the words must still be read right to left on each line. Thus, if the Hebrew phrase in the example



above had to be broken, the Hebrew words would appear to be in a different order.

The first phrase in Lamentations is איכה ישבה

בדד (How she sits in solitude!).

or

The first phrase in Lamentations is איכה

ישבה בדד (How she sits in solitude!).

As a safeguard, the author should highlight all the words in Hebrew phrases and furnish detailed instructions on how to implement line breaks.

- 11.97 **A note on Hebrew vowels.** Most Hebrew vowels are not letters; they are marks attached to the letters, most of which are consonants. In Hebrew texts the vowel marks (as well as dots that modify the pronunciation of consonants) rarely appear. Among texts in which the marks do appear are prayer books, printed Bibles, and poetry.

### *Russian*

- 11.98 **Russian transliteration.** Of the many systems for transliterating Russian, the most important are summarized in table 11.3. Journals of Slavic studies generally prefer a “linguistic” system that makes free use of diacritics and ligatures. In works intended for a general audience, however, diacritics and ligatures should be avoided. For general use, Chicago recommends the system of the United States Board on Geographic Names. Regardless of the system followed, the spellings for names listed in the Merriam-Webster dictionaries (bibliog. 3.1) should prevail.

Catherine the Great	Moscow
Chekhov	Nizhniy (or Nizhni) Novgorod
Dnieper River	Tchaikovsky

- 11.99 **Russian capitalization.** Capitalization conventions in Cyrillic are much like those of French and should be preserved in transliteration. Pronouns, days of the week, months, and most proper adjectives are lowercased. Geographic designations are capitalized when they apply to formal institutions or political units but otherwise lowercased.

Tverskaya guberniya	Moskovskiy universitet
tverskoye zemstvo	russkiy kompozitor

TABLE 11.3. Russian alphabet (and Unicode numbers) and romanization

Basic Russian (Cyrillic) alphabet (and Unicode numbers)				US Board on Geographic Names	Library of Congress	Linguistic system <sup>2</sup>
Upright		Cursive <sup>1</sup>				
А (0410), а (0430)		А, а		a		
Б (0411), б (0431)		Б, б		b		
В (0412), в (0432)		В, в		v		
Г (0413), г (0433)		Г, г		g		
Д (0414), д (0434)		Д, д		d		
Е (0415), е (0435)		Е, е		ye, <sup>3</sup> e	e	e
Ё (0401), ё <sup>4</sup> (0451)		Ё, ё		yë, <sup>3</sup> ë (OOEB)	ë	e, ë
Ж (0416), ж (0436)		Ж, ж		zh		ž
З (0417), з (0437)		З, з		z		
И (0418), и (0438)		И, и		i		
Й (0419), й (0439)		Й, й		y	ÿ (012D)	j
К (041A), к (043A)		К, к		k		
Л (041B), л (043B)		Л, л		l		
М (041C), м (043C)		М, м		m		
Н (041D), н (043D)		Н, н		n		
О (041E), о (043E)		О, о		o		
П (041F), п (043F)		П, п		p		
Р (0420), р (0440)		Р, р		r		
С (0421), с (0441)		С, с		s		
Т (0422), т (0442)		Т, т		t		
У (0423), у (0443)		У, у		u		
Ф (0424), ф (0444)		Ф, ф		f		
Х (0425), х (0445)		Х, х		kh		x
Ц (0426), ц (0446)		Ц, ц		ts	ṭs <sup>5</sup>	c
Ч (0427), ч (0447)		Ч, ч		ch		č
Ш (0428), ш (0448)		Ш, ш		sh		š
Щ (0429), щ (0449)		Щ, щ		shch		šč
Ъ (042A), ъ <sup>6</sup> (044A)		Ъ, ъ		” (201D) <sup>7</sup>	” (02BA) <sup>8</sup>	” (02BA) <sup>8</sup>
Ы (042B), ы <sup>6</sup> (044B)		Ы, ы		y		
Ь (042C), ь <sup>6</sup> (044C)		Ь, ь		’ (2019) <sup>9</sup>	’ (02B9) <sup>10</sup>	’ (02B9) <sup>10</sup>
Э (042D), э (044D)		Э, э		e	è (0117)	è (00E8)
Ю (042E), ю (044E)		Ю, ю		yu	ÿ <sup>5</sup>	ju
Я (042F), я (044F)		Я, я		ya	ÿa <sup>5</sup>	ja

NOTE: The Library of Congress and linguistic systems employ the same characters as the US Board system except where noted.

<sup>1</sup>The Unicode numbers are the same for the upright and cursive characters; the differences in appearance depend on the italic version of a given typeface.

<sup>2</sup>The term *linguistic* describes a system generally preferred by journals of Slavic studies (see 11.98).

<sup>3</sup>Initially and after a vowel or ъ or ь.

<sup>4</sup>Not considered a separate letter; usually represented in Russian by *е*.

<sup>5</sup>Character tie, sometimes omitted, may be produced by using the combining double inverted breve (U+0361).

<sup>6</sup>Does not occur initially. <sup>7</sup>Right double quotation mark.

<sup>8</sup>Modifier letter double prime (hard sign). <sup>9</sup>Right single quotation mark.

<sup>10</sup>Modifier letter prime (soft sign).

- 11.100 **Titles of Russian works.** Only the first word and any proper nouns are capitalized in titles.

N. A. Kurakin, *Lenin i Trotskiy*

O. I. Skorokhodova, *Kak ya vosprinimayu i predstavlyayu okruzhayushchiy mir*  
[How I perceive and imagine the external world]

Note that in the original Cyrillic, titles are set in ordinary type; the Cyrillic *kursiv* is used more sparingly than our italic and never for book titles. In transliterations, however, italic should be used.

- 11.101 **Russian quotations and dialogue.** Russian generally resembles French in its use of guillemets (« ») for dialogue and quoted material and of dashes for dialogue (see 11.29, 11.31).

«Bozhe, bozhe, bozhe!» govorit Boris.

—S kem ya rabotayu?

—S tovarishchem.

—Kak my rabotayem?

—S interesom.

To set off a quotation within a speech, guillemets may be used, as in French. For an example, see 11.31.

- 11.102 **Russian ellipses.** Ellipses are used as in French (see 11.32) to indicate interruptions or breaks in thought.

Ya . . . vy . . . my tol'ko chto priyekhali.

In Russian, an exclamation point or a question mark often takes the place of one of the dots; this convention may be regularized to three dots in English publications.

Mitya! . . . Gde vy byli? . . .

- 11.103 **Russian uses of the dash.** A dash is sometimes inserted, with a space on either side, between subject and complement when the equivalent of *is* or *are* is omitted.

Moskva — stolitsa Rossii.

Similarly, a dash, preceded and followed by a space, is used in place of a verb omitted because it would be identical to the preceding verb.

Ivan i Sonya poyedut v Moskvu poyezdom, Lev i Lyuba — avtobusom.

11.104 **Russian word division—general.** Transliterated Russian should be divided according to the rules governing word division in the Cyrillic original. The guidelines in this section are adapted from the transliteration system of the United States Board on Geographic Names.

11.105 **Combinations not to be divided in Cyrillic transliteration.** Combinations representing single Cyrillic letters—*ch, kh, sh, shch, ts, ya, ye, yë, yu, zh*—should never be divided, nor should combinations of a vowel plus short *i* (or yod, transliterated *y*): *ay, ey, yey*, and so on.

11.106 **Division between Russian consonants.** Words may be divided between single consonants or between a consonant and a consonant combination.

ubor-ku      chudes-nym      mol-cha      sred-stvo      mor-skoy

The following consonant combinations are not normally divided: *bl, br, dr, dv, fl, fr, gl, gr, kl, kr, ml, pl, pr, sk, skr, skv, st, str, stv, tr, tv, vl, vr, zhd*. They may, however, be divided if they fall across the boundary of a prefix and a root or other such units (e.g., ob-lech', ras-kol).

11.107 **Division of Russian words after prefixes or between parts.** Words may be divided after a prefix, but generally the prefix itself should not be divided.

bes-poryadok    pere-stroyka    za-dat'    pred-lozhit'    pro-vesti    obo-gnat'

Compound words should be divided between parts.

radio-priyëmnik      gor-sovet      kino-teatr

11.108 **Division of Russian words after vowel or diphthong.** Words may be divided after a vowel or a diphthong before a single (Cyrillic) consonant.

Si-bir'      voy-na      Gorba-chev      da-zhe

Division after a vowel may also be made before a consonant combination.

puteshe-stvennik      khi-trit'      pro-stak      ru-brika

## South Asian Languages

- 11.109 **South Asian special characters.** Transliteration of the principal South Asian languages may require the following special characters (see also table 11.2):

Āā, Ḍḍ, Ēē, Ḣḥ, Īī, Ĺĺ, Ļļ, M̄m̄, Ṁṁ, N̄n̄, Ṅṅ, Ññ, Ñ̄ñ̄, Ōō, R̄r̄, Ṙṙ, Śś, Ś́ś́, Ṭṭ, Ūū, Ýý

Many writers using South Asian languages, however, employ a simplified style that does not use diacritics at all—for example, substituting *sh* for various *s*'s, ignoring subscript dots for dental consonants, and omitting macrons altogether.

## Classical Greek

- 11.110 **Transliterating Greek.** Isolated Greek words and phrases in works not focusing on ancient Greece are usually transliterated. Table 11.4 shows the Greek alphabet (with Unicode numbers) and corresponding letters of the Latin alphabet. In transliteration, all Greek accents are omitted. The macron is used to distinguish the long vowels eta (*ē*) and omega (*ō*) from the short vowels epsilon (*e*) and omicron (*o*). The iota subscript is transliterated by an *i* on the line, following the vowel it is associated with (*ἀνθρώπων*, *anthrōpōi*). The rough breathing is transliterated by *h*, which precedes a vowel or diphthong and follows the letter *r* (as in the English word *rhythm*). The smooth breathing is ignored, since it represents merely the absence of the *h* sound. If a diaeresis appears in the Greek, it also appears in transliteration. Transliterated Greek words or phrases are usually italicized unless the same words occur frequently, in which case they may be italicized at first mention and then set in roman.
- 11.111 **Typesetting Greek.** Authors who need to present Greek should use a Unicode-enabled font if at all possible (see 11.2). Publishers need to make sure that a Greek font is available for publication; Greek may need to be set in a slightly different size to make it visually match the surrounding type. Greek is normally not set in italics. Extra white space must occasionally be added where more than one diacritic appears over a vowel.

TABLE 11.4. Greek alphabet (and Unicode numbers) and romanization

Name of letter	Greek alphabet (and Unicode numbers)	Transliteration
alpha	Α (0391), α (03B1)	a
beta	Β (0392), β (03B2)	b
gamma	Γ (0393), γ <sup>1</sup> (03B3)	g
delta	Δ (0394), δ <sup>2</sup> (03B4)	d
epsilon	Ε (0395), ε (03B5)	e
zeta	Ζ (0396), ζ (03B6)	z
eta	Η (0397), η (03B7)	ē (0113)
theta	Θ (0398), θ <sup>3</sup> (03B8)	th
iota	Ι (0399), ι (03B9)	i
kappa	Κ (039A), κ (03BA)	k
lambda	Λ (039B), λ (03BB)	l
mu	Μ (039C), μ (03BC)	m
nu	Ν (039D), ν (03BD)	n
xi	Ξ (039E), ξ (03BE)	x
omicron	Ο (039F), ο (03BF)	o
pi	Π (03A0), π (03C0)	p
rho	Ρ (03A1), ρ (03C1)	r; <i>initially, rh; double, rrh</i>
sigma	Σ (03A3), σ (03C3), ς <sup>4</sup> (03C2)	s
tau	Τ (03A4), τ (03C4)	t
upsilon	Υ (03A5), υ (03C5)	u; <i>often y, exc. after a, e, ē, i</i>
phi	Φ (03A6), φ <sup>5</sup> (03C6)	ph
chi	Χ (03A7), χ (03C7)	kh, ch
psi	Ψ (03A8), ψ (03C8)	ps
omega	Ω (03A9), ω (03C9)	ō (014D)

<sup>1</sup>Note that γγ becomes ng, and γκ becomes nk.

<sup>2</sup>Sometimes incorrectly appears as ∂ (U+2202, partial differential).

<sup>3</sup>Also ϑ (U+03D1). <sup>4</sup>Final letter. <sup>5</sup>Also ϕ (U+03D5).

## Breathings and Accents

- 11.112 **Greek breathing marks.** When Greek is set in the Greek alphabet, every initial vowel or diphthong or rho must be marked with a breathing, either rough (´, *dasia*) or smooth (˘, *psili*). The breathing mark is placed over the initial lowercase vowel (or the second vowel of a diphthong). It is positioned to the left of capital letters. Note that a single quotation mark cannot function as a breathing because it is the wrong size and does not sit close enough to the letter.

αὔτε ἕτεροι Ἕλληνας ἥβη Ἴρις ὑπέχω ὠκύς ῥάδιος

- 11.113 **Greek accent marks.** There are three Greek accent marks: acute, or *oxia* (´); circumflex, or *perispomeni*, either tilde-shaped or rounded (˜ or ˘),

depending on the typeface; and grave, or varia ( ` ). Accents in Greek occur only over vowels. The circumflex occurs only on the two final syllables of a word. The grave accent occurs only on the last syllable. Like breathings, accents are placed over lowercase vowels, over the second vowel of a diphthong, and to the left of capital vowels. A diaeresis is used to indicate that two successive vowels do not form a diphthong but are voiced separately (as in French *naïf*).

- 11.114 **Unaccented Greek words.** With two exceptions, all Greek words are marked with accents—usually one, occasionally two (see below). The first exception is a group of monosyllabic words called proclitics, which are closely connected with the words following them. The proclitics are the forms of the definite article  $\acute{o}$ ,  $\acute{\eta}$ ,  $\acute{o}\iota$ ,  $\acute{\alpha}\iota$ ; the prepositions  $\epsilon\iota\varsigma$ ,  $\acute{\epsilon}\nu$ ,  $\acute{\epsilon}\kappa$  ( $\acute{\epsilon}\xi$ ); the conjunctions  $\epsilon\iota$ ,  $\acute{\omega}\varsigma$ ; and the adverb  $\acute{o}\upsilon$  ( $\acute{o}\upsilon\kappa$ ,  $\acute{o}\upsilon\chi$ ). The second exception is a group called enclitics, short words pronounced as if part of the word preceding them. Enclitics usually lose their accents ( $\text{Ἀρταξεροξής τε}$ ), and in certain circumstances the word preceding them gains a second accent ( $\text{φοβείται τις}$ ).
- 11.115 **Greek vowels.** Vowels complete with breathing marks and accents, in all combinations, are an integral part of every Greek font used in publishing. Each font, for example, should be able to provide, for lowercase eta,  $\eta$ ,  $\acute{\eta}$ ,  $\tilde{\eta}$ ,  $\grave{\eta}$ ,  $\dot{\eta}$ ,  $\acute{\eta}$ ,  $\tilde{\eta}$ ,  $\grave{\eta}$ ,  $\dot{\eta}$ ,  $\acute{\eta}$ ,  $\tilde{\eta}$ ,  $\grave{\eta}$ ,  $\dot{\eta}$ , and, for uppercase eta,  $\text{H, } \acute{\text{H}}, \tilde{\text{H}}, \grave{\text{H}}, \dot{\text{H}}, \acute{\text{H}}, \tilde{\text{H}}, \grave{\text{H}}, \dot{\text{H}}$ . Additional symbols are needed for scholarly works treating ancient manuscripts or papyri. Consult the latest Unicode character charts for Greek alphabets.

### *Punctuation and Numbers*

- 11.116 **Greek punctuation.** In Greek the period and comma are the same as in English; the colon and semicolon are both represented by a midlevel dot ( $\cdot$ ); the question mark is represented by a semicolon. The apostrophe (which looks almost like a smooth breathing mark) is used as an elision mark when the final vowel of one word is elided before a second word beginning with a vowel. In English texts, quoted words or passages in the Greek alphabet, of whatever length, should not be enclosed in quotation marks.
- 11.117 **Greek numbers.** Numbers, when not written out, are represented in ordinary Greek text by the letters of the alphabet, supplemented by three additional, obsolete Greek letters—stigma, koppa, and sampi:  $\zeta' = 6$ ,  $\rho' = 90$ ,  $\lambda' = 900$ . The diacritical mark resembling a prime (and defined for

TABLE 11.5. Greek numerals

1	α'	13	ιγ'	30	λ'	600	χ'
2	β'	14	ιδ'	40	μ'	700	ψ'
3	γ'	15	ιε'	50	ν'	800	ω'
4	δ'	16	ιζ'	60	ξ'	900	ϗ <sup>3</sup>
5	ε'	17	ιζ'	70	ο'	1,000	,α
6	ζ <sup>1</sup>	18	ιη'	80	π'	2,000	,β
7	ζ'	19	ιθ'	90	ρ <sup>2</sup>	3,000	,γ
8	η'	20	κ'	100	ρ'	4,000	,δ
9	θ'	21	κα'	200	σ'	10,000	,ι
10	ι'	22	κβ'	300	τ'	100,000	,ρ
11	ια'	23	κγ'	400	υ'		
12	ιβ'	24	κδ'	500	φ'		

<sup>1</sup>Stigma (U+03DB); also represented with digamma (U+03DD): Ϝ'.

<sup>2</sup>Archaic koppa (U+03D9); also represented with koppa (U+03DF): ϝ'.

<sup>3</sup>Sampi (U+03E1); formerly disigma (double sigma).

Unicode as the Greek numeral sign, U+0374) distinguishes the letters as numerals and is added to such a letter standing alone or to the last sign in a series. For example, ϱια' means 111. For thousands, the foregoing letters are used with a different diacritical mark (the Greek lower numeral sign, U+0375): ,α = 1,000, ,αρια' = 1,111, ,βσκαβ' = 2,222. See table 11.5.

### Word Division

- 11.118 **Greek word division—consecutive vowels.** Diphthongs (αι, αυ, ει, ευ, ηυ, οι, ου, υι, ωυ) are never divided. But two consecutive vowels that do not form a diphthong are divided.

θε-ά-ο-μαι      υί-ός      παύ-ε-τε      νε-ώς

- 11.119 **Greek word division—single consonants.** When a single consonant occurs between two vowels, the word is divided before the consonant.

φω-νή      κε-φα-λίσ      μέ-γα      δέ-δω-κεν      μή-τηρ

- 11.120 **Greek word division—two or more consonants.** If a consonant is doubled, or if a mute is followed by its corresponding aspirate (πφ, βφ, κχ, γχ, τθ, δθ), the word is divided after the first consonant.



θά-λασ-σα      συγ-χαί-ρω

If the combination of two or more consonants begins with a liquid (λ, ρ) or a nasal (μ, ν), division is made after the liquid or nasal.

ξιμ-προ-σθεν      (*but before* μν: μέ-μνημαι)

All other combinations of two or more consonants *follow* the division.

πρᾶ-γμα	τέ-χνη	βα-θμός	αι-σχρός
βί-βλί-ον	δά-κτυ-λος	σκῆ-πτρον	βά-κτρον

- 11.121 **Greek word division—compounds.** Compound words are divided between parts; within each part the rules detailed elsewhere in this section apply. The commonest type of compound word begins with a preposition or a prefix.

ἀμφ-	ἀφ-	ὕπ-	ἐξ-έβαλον
ἀν-	ἐφ-	ὕφ-	καθ-ίστημι
ἀπ-	κατ-		δύσ-μορφος

## Old English and Middle English

- 11.122 **Special characters in Old and Middle English.** Several Old English or Middle English letters not used in modern English occur in both lower-case and capital forms (see also table 11.1).

Ð ð    edh or eth  
 Þ þ    thorn

Both edh and thorn represent voiced or unvoiced *th*, as in *them* or *three*.

ȝ ȝ    Yogh; occurs in Old English representing *g* as in *good*, *y* as in *year*, or *gh* as in *light* and *thought*. Yogh sometimes occurs in Middle English representing *y* as in *year* and *gh* as in *light* and *thought*, but normally not *g* as in *good*.

Æ æ    Ligature; should *not* be printed as two letters in Old English names and text (Ælfric).

Authors should use the correct Unicode characters for the ligature and for edh, thorn, and yogh, and should provide their publisher with a list of these and any other special characters (see 11.2). For the long *s* (ſ), see 13.7.

- 11.123 **Ampersand and wynn.** In Old English and Middle English texts a sort of stylized seven (the Tironian *et*) may be found for *and*, but the modern ampersand may be substituted for this. In Old English texts  $\wp$  or  $\text{p}$  (wynn) is found for *w*; the modern *w* is often substituted for this.
- 11.124 **Old English vowels.** Modern editors of Old English sometimes distinguish between long and short vowels and diphthongs by means of a macron over the long versions (e.g.,  $\bar{a}$ ,  $\bar{x}$ ,  $\bar{e}$ ,  $\bar{e}a$ ,  $\bar{e}o$ ,  $\bar{i}$ ,  $\bar{i}o$ ,  $\bar{o}$ ,  $\bar{u}$ ,  $\bar{y}$ ). Note that, with the exception of the  $\bar{x}$  ligature, diphthongs are usually marked with a macron over only the first vowel.

## American Sign Language (ASL)

- 11.125 **Signed languages.** The visual-gestural languages used by deaf people in different parts of the world are called signed languages. Signed languages are quite different from spoken languages (although there may be regional effects of language contact), and a particular signed language may or may not share the same national or geographic boundaries as spoken languages in the same locations. The individual elements of these languages are known as signs.
- 11.126 **Components of signs.** Signs have five major articulatory components—handshape, location, orientation, movement, and (in some cases) distinctive nonmanual signals.
- 11.127 **Writing ASL.** Many formal systems for writing signed languages exist; however, none has been adopted for widespread use by deaf signers. This section offers an overview of some of the most frequently employed conventions for written transcription of signing. For additional resources, see Charlotte Baker-Shenk and Dennis Cokely, *American Sign Language: A Teacher's Resource Text on Grammar and Culture*; and Clayton Valli, Ceil Lucas, Kristin J. Mulrooney, and Miako Villanueva, *Linguistics of American Sign Language: An Introduction* (bibliog. 5).
- 11.128 **Glosses in ASL.** The written-language transcription of a sign is called a *gloss*. Glosses are words from the spoken language written in small capital letters: WOMAN, SCHOOL, CAT. (Alternatively, regular capital letters may be used.) When two or more written words are used to gloss a single sign, the glosses are separated by hyphens. The translation is enclosed in double quotation marks.

The sign for “a car drove by” is written as VEHICLE-DRIVE-BY.

One obvious limitation of the use of glosses from the spoken/written language to represent signs is that there is no one-to-one correspondence between the words or signs in any two languages.

- 11.129 **Compound signs.** Some combinations of signs have taken on a meaning separate from the meaning of the individual signs. Various typographical conventions are used to indicate these compounds, including a “close-up” mark or a plus sign. Depending on the transcription system, the sign for “parents” might be glossed as follows:

MOTHER<sup>^</sup>FATHER *or* MOTHER+FATHER

- 11.130 **Fingerspelling.** For proper nouns and other words borrowed from the spoken language, the signer may fingerspell the word, using the handshapes from a manual alphabet. (There are numerous fingerspelling alphabets used by different signed languages, among them the American Manual Alphabet.) Fingerspelled words may be transcribed in any of the following ways:

fs-JOHN *or* J-O-H-N *or* j-o-h-n

- 11.131 **Lexicalized signs.** Over time, some fingerspelled words have taken on the quality of distinct signs, either by omission of some of the individual letter signs or by a change in the orientation or movement of the letter signs. These lexicalized signs are represented by the “pound” symbol (#): #WHAT, #BACK, #DO.
- 11.132 **Handshapes.** Most of the handshapes of American Sign Language are described by the corresponding alphabetic or numerical handshape or a variation thereof. For example, APPLE is made with an X handshape; CREATE is made with a 4 handshape; ANY is made with an Open A handshape; YELL is made with a Bent 5 handshape. Handshapes without a clear relative in the fingerspelling or number system are labeled idiosyncratically according to the transcription system in use. For example, SARCASTIC is made with the HORNS handshape; AIRPLANE is made with the ILY handshape. Handshapes for signed languages that do not use the American Manual Alphabet are often described in relation to the ASL handshapes.
- 11.133 **Transcriptions of signed sentences.** Signed sentences are written as a sequence of glosses, often with the spoken/written-language translation underneath in italics or quotation marks or both. (For examples, see 11.134, 11.135.) Punctuation is generally omitted from sentence tran-

scriptions (though not from the translations). Some writers, however, add question marks and exclamation points, and a comma may be used to indicate a short pause in the sentence.

- 11.134 **Pronouns, possessives, and reference.** Pronouns are commonly transcribed either as IX (since these are frequently produced with the “index” finger) or as PRO. Either of these is followed by indication of person and sometimes number. A similar convention is used with the possessive marker, sometimes glossed as POSS. There are varying conventions about how to indicate person and number. Thus, a third-person singular pronoun in ASL (equivalent to English “he,” “she,” or “it”) might be glossed as IX<sub>3p</sub>, IX-3p, or PRO.3. A second-person plural pronoun could be glossed as IX<sub>2p-pl</sub>. Subscript indices are often used to show signs articulated in the same location or to indicate coreferential noun phrases. The following example indicates that *he* and *his* refer back to the same person:

IX<sub>3pi</sub> LOSE POSS<sub>3pi</sub> HOUSE  
*He lost his house.*

- 11.135 **Nonmanual signals.** Nonmanual gestures may be labeled based on anatomical behavior or grammatical interpretive function. These gestures, indicated by various abbreviations and terms, are typeset in a smaller font followed by a half-point rule above the ASL sentence. For example, the label *whq* is commonly used to refer to the facial expression that marks questions involving “who,” “what,” “when,” “where,” “how,” or “why.” This expression consists of a cluster of features that include furrowed brows and slightly squinted eyes. In the example below, *whq* occurs over the entire question (i.e., the expression is articulated simultaneously with all of the manual signs over which the line extends). In the same example, the label *t* indicates a topic marker that occurs simultaneously with the sign YESTERDAY. Correct alignment is critical to an accurate transcription.

*t* \_\_\_\_\_ *whq*  
 YESTERDAY, fs-JOHN SEE WHO  
*Whom did John see yesterday?*