# locale  
LANG=en\_US.UTF-8  
LC\_TIME="es\_ES.UTF-8"  
LC\_MESSAGES="en\_CA.UTF-8"  
LC\_ALL="de\_DE.UTF-8"

If you were to run the **date** command, which locale would be used for the formatting?

**a.** American (US) English

**b.** Spanish (ES)

**c.** Canadian (CA) English

**d.** German (DE)

**D.** D is correct because LC\_ALL takes priority over the other variables. A is not correct because the LANG variable has the lowest priority. B is not correct because LC\_TIME is overridden by LC\_ALL. C is not correct because LC\_MESSAGES is not used for date formatting.

**How Linux Uses the Locale**

Internationalization in Linux is handled with the GNU **gettext** library. If programmers write their applications with that library and annotate their messages correctly, the user can change the behavior with environment variables.

As multiple things that can be localized, such as numbers and messages, **gettext** has a series of environment variables that it checks to see which locale is appropriate.

**LANGUAGE (special case)**

**LC\_ALL (highest priority)**

**LC\_XXX (medium priority)** LC\_CTYPE, LC\_NUMERIC, LC\_TIME, LC\_COLLATE, LC\_MONETARY, LC\_MESSAGES, ...

**LANG (lowest priority)**

The **LANGUAGE** variable is only consulted when printing messages. It is ignored for formatting. Also, the colon (:) gives the system a list of locales to try in order when trying to display a system message. LANGUAGE variable is only set in extremely special case.

Not all programs have translations for all languages. By default, an English message is shown in place of a nonexistent translation. If you understand other languages, you can set up a priority list of languages. This is done through a different environment variable, called LANGUAGE. GNU gettext gives preference to LANGUAGE over LC\_ALL and LANG for the purpose of message handling, but you still need to have LANG (or LC\_ALL) set to the primary language; this is required by other parts of the system libraries. For example, some Swedish users who would rather read translations in German than English for when Swedish is not available, set LANGUAGE to ‘sv:de’ while leaving LANG to ‘sv\_SE’.

Special advice for Norwegian users: The language code for Norwegian bokmål changed from ‘no’ to ‘nb’ recently (in 2003). During the transition period, while some message catalogs for this language are installed under ‘nb’ and some older ones under ‘no’, it is recommended for Norwegian users to set LANGUAGE to ‘nb:no’ so that both newer and older translations are used.

In the LANGUAGE environment variable, but not in the other environment variables, ‘ll\_CC’ combinations can be abbreviated as ‘ll’ to denote the language’s main dialect. For example, ‘de’ is equivalent to ‘de\_DE’ (German as spoken in Germany), and ‘pt’ to ‘pt\_PT’ (Portuguese as spoken in Portugal) in this context.

Note: The variable LANGUAGE is ignored if the locale is set to ‘C’. In other words, you have to first enable localization, by setting LANG (or LC\_ALL) to a value other than ‘C’, before you can use a language priority list through the LANGUAGE variable.

**LC\_ALL** is an environment variable that overrides all of these settings in LC\_XXX and LANG. It is typically used in scripts that run particular programs. For example, configure scripts generated by GNU autoconf use LC\_ALL to make sure that the configuration tests don’t operate in locale dependent ways.

**LC\_XXX** (LC\_XXX overrides LANG) gives the administrator the power to override a locale for a particular element. For example, if **LANG** were set to en\_US.UTF-8 the user could override currency display by setting **LC\_MONETARY**. The **locale** command displays the current settings, as shown in [Example 11-10](ch11.html#ch11ex10).

**Example 11-10** Using **locale**

[**Click here to view code image**](ch11_images.html#p11ex10a)

**# locale**  
LANG=en\_CA.UTF-8  
LANGUAGE=en\_CA:en  
LC\_CTYPE="en\_CA.UTF-8"  
LC\_NUMERIC="en\_CA.UTF-8"

LC\_TIME="en\_CA.UTF-8"  
LC\_COLLATE="en\_CA.UTF-8"  
LC\_MONETARY="en\_CA.UTF-8"  
LC\_MESSAGES="en\_CA.UTF-8"  
LC\_PAPER="en\_CA.UTF-8"  
LC\_NAME="en\_CA.UTF-8"  
LC\_ADDRESS="en\_CA.UTF-8"  
LC\_TELEPHONE="en\_CA.UTF-8"  
LC\_MEASUREMENT="en\_CA.UTF-8"  
LC\_IDENTIFICATION="en\_CA.UTF-8"  
LC\_ALL=  
This example is from a typical English system. You can override just  
parts of it:  
**# LC\_TIME=fr\_FR.UTF8 date**  
samedi 7 mars 2015, 23:11:23 (UTC-0600)  
**# LC\_MESSAGES=fr\_FR.UTF8 man**  
What manual page do you want?  
**# LANGUAGE='' LC\_MESSAGES=fr\_FR.UTF8 man**  
Quelle page de manuel voulez-vous ?

In the preceding example, the time setting is switched to the French locale and the date is displayed in French. The second command sets the messages setting to French, but the English variant is used because the higher priority **LANGUAGE** is set. A French error message is used once **LANGUAGE** is set to nothing.

**LANG** is the normal environment variable for specifying a locale. As a user, you normally set this variable (unless some of the other variables have already been set by the system, in /etc/profile or similar initialization files).

While you’re using the locale command, you should try it with the -a option, which  
identifies all of the locales that are available to you:  
$ **locale -a**  
C  
en\_US.utf8  
POSIX  
In this example (from an Ubuntu system), very few locales are installed. Some systems  
may have many more; one of our computers has hundreds of locales available, for example.