CHAPTER 17

International Language Support

With the growth of the world-wide PC market, the use of Microsoft Windows and Windows–based applications has made PCs easier to use around the globe. Windows and Windows–based application are sold and used worldwide. This poses some unique problems for both Microsoft as an operating system vendor, and ISVs as application developers.

When a new software application or operating system intended for a world market is developed, efforts must be undertaken to *localize* the software to a given country and written language in which it will be used. In many cases, this is as simple as changing the names of menus, menu items, and strings displayed by the software to match the foreign language used in the locale. However, as the features and functionality of a software product grow, so does the complexity required to tailor the application to characteristics of the native country. Since the start of the design work for the Windows NT operating system, Microsoft has been adding to the level of support for international languages and cultural conventions in the 32-bit editions of the Windows family of operating systems.

This section discusses the localization plans for Windows 95, the built-in international support for using Windows 95 on a world-wide basis, and the special provisions that Windows 95 provides for enhancing existing or developing new applications that can be used in different parts of the world.

Summary of Improvements over Windows 3.1

Support for using the Windows operating system on a global basis is improved in Windows 95. Windows 95 offers benefits to both end-users and to software developers, which can be summarized as follows:

Benefits for Users

- **Windows 95 makes it easy to use multiple language fonts and character sets, and easily switch between the different keyboard layouts required to support them.**

  With the Eastern European version of Windows 3.1, a user can directly switch between only two keyboard layouts; for example, Russian and English. With the standard Latin versions of Windows, a user cannot easily switch between different keyboard layouts directly—the user has to go to Control Panel for each language switch.

  With Windows 95, users can easily switch between all available languages and corresponding keyboard layouts configured on their system by using the Alt+Shift key combination—making it easy for users to integrated information into a multilingual document.

- **Windows 95 substitutes fonts when switching between different languages if the original font is not present on the system.**

  When switching between different languages, matching fonts for the new language are substituted if the original font is not available. This allows users to be able to read and use the text for a similar character set, even if they don't have the same font that the original information was created in.

- **Proper sorting and formatting rules are used corresponding to the presently configured locale.**

  Different locales and cultures have different rules for interpreting information. For example, cultures use different sequence algorithms for sorting information, use different comparison algorithms for finding or searching for information, and use different formats for specifying time and date information. Win32–based applications that use the National Language Support (NLS) APIs allow users to easily exchange information on a global basis, while preserving the integrity of the information.

Benefits for Developers

- **It’s easy for application developers to add international language support to their applications.**

  Developers can now use the Win32 NLS APIs for sorting, searching, and manipulating information in a locale-independent way. NLS services in Windows 95 ensures that information is handled properly for the given culture or locale. The correct national format is automatically supplied based on the international settings specified by the user in Control Panel. For example, to obtain the current date format information to match the current locale, a developer can simply make the application call an NLS API and the system will return the proper format. Likewise, to sort information in the proper sequence in French, Norwegian, or Spanish, the application calls a corresponding culture-independent NLS API.

- **Windows 95 provides services for application vendors to automatically switch between the proper fonts and keyboard layouts as users navigate through a multilingual document.**

  For users who create or edit multilingual content in their documents (for example, translators), a Win32–based application that uses the international services in Windows 95 automatically activates the correct fonts and corresponding keyboard layouts for the edit point in the text. This allows easy editing of information contained within multilingual documents.

- **Information can be passed through the Clipboard to other applications and preserve language-specific attributes.**

  Windows 95 provides additional services for application vendors to easily exchange information between internationally-aware applications, while preserving all language formatting characteristics.

- **Multilingual-aware applications can change the given language of the system under program control.**

  Windows 95 provides services that application vendors can use to automatically switch the language that the system uses to match attributes in a document. For example, as a user scrolls through a multilingual document, the application can automatically switch the system language to match the format of the information contained within the document.

- **Language-relevant information can be saved in Rich Text Format (RTF) from a multilingual-aware application.**

  Extensions are provided to the RTF specification to support storing international language information in RTF files.

Localization of the Windows 95 Operating System

As a result of success of Microsoft Windows around the world, Windows and Windows–based applications have been localized into many different languages. Microsoft Windows 3.1 was localized into more 25 major languages. However, this process took as long as 18 months, thus delaying the availability of Windows 3.1 for some language versions. The Windows 95 development team has been working on international localization issues concurrently with the development of the domestic U.S. version of the operating system. To better support a global market, Microsoft plans to localize Windows 95 into at least 29 different language versions. The localized versions of Windows 95 will be released on a planned development schedule that does not exceed 120 days. The planned localized product versions for Windows 95 include German, French, Spanish, Swedish, Dutch, Italian, Norwegian, Danish, Finnish, Portuguese, Japanese, Chinese, Korean, Russian, Czech, Polish, Hungarian, Turkish, Greek, Arabic, Basque, Hebrew, Thai, Indonesian, and Catalan (as well as several other variations of these languages).

International Language Issues

For an operating system to be used effectively in a world-wide environment, the localization is only a small part of the solution to offer global customers a solution. A world-wide operating system release must also provide services to support the use of international applications and support a global market by making the application developer’s job easier.

Here are some international language issues which international end-users and applications developer face:

From the End-User’s Perspective

Some users need to use more than one language in a document. For example, they might be translating from English to Russian or they might be writing an instruction document for a product in many different languages. This causes series of obstacles for the user. For example, the user must repeatedly switch to another keyboard layout on-the-fly so that he or she can continue writing text in a different language. Likewise, when using a database, the user faces the problem of sorting the information in the proper order for a given language.

From the Developers Perspective

When localizing a product into different languages, a developer is faced with several questions, such as: What is the correct sorting order for French? How is a date represented in Germany? Do the Swedes really need to have the ability to use the characters: Å, Ä, Ö? If a document contains text in more than one language, is there some way for the software to “know” which part of the document is in which language? Can information in a multilingual document be passed to another application via the Clipboard?

It is of course not expected that every developer knows how to address these issues, but many try natively as part of their application and come a little short, creating problems for the end-user, their support organization, and their own development team.

In Windows 95, Microsoft has set out to offer international language support at the operating system and API level to add functionality which provides solutions to using software and exchanging documents around the world.

International Language Support

Since the mainstream Windows operating system platform has not previously offered international language support as an operating system service, many application vendors have hard-coded global characteristics into their applications. This allows their applications to be used in a given locale, but prevents them users from easily using the application in a different cultural environment. Therefore, the user is dependent upon the application vendor for providing a version of the application that matches their locale attributes.

Providing international language support services in the Windows 95 operating system makes it easier for application developers to solve international language issues related to presenting or manipulating information in their applications.

Date and Time Formats

Date and time information needs to be represented in different formats depending on the locale where the information is being used. For example, date information presented in English places the day between the month and year as in “March 9, 1994,” whereas a different locale may represent the same date as “9 March 1994.”

Sorting and Searching Support

International language issues are much more complex than simply representing date and time information in the proper format. Sorting and searching algorithms in applications must correspond to the proper language rules for the locale in which the information is being used and manipulated.

Here are some examples of subtle differences between different language rules:

- In French, diacritics are sorted right to left instead of left to right as in English.
- In Norwegian, some extended characters follow the Z character as they are considered unique characters rather than characters with a diacritic.
- In Spanish, CH is a unique character between C and D, and Ñ is a unique character between N and O.

In Windows 3.1, many developers developed their own sorting routines for different languages and hard-coded this functionality into their code. This makes their applications inflexible to support the numerous right sorting tables required for all the languages in which they want to localize their application in.

As a further example, when sorting a database in Swedish with an English-language sort algorithm, names would be sorted like this:

English sorting	**Correct Swedish sorting**
Andersson	Andersson
Åkesson	Karlsson
Ärlingmark	Magnusson
Karlsson	Turesson
Magnusson	Åkesson
Turesson	Ärlingmark

The system treats the Å and Ä as an A and therefore sorts it as an A at the top of the list. But in the correct Swedish sort order, the Å and Ä are sorted at the end of the alphabet, after Z. The reason that the names starting with Å and Ö are sorted as last in the correct Swedish sorting is that the Å and Ä are separate vowels in the Swedish language and occur at the very end of the alphabet. A Swede looking for “Ärlingmark” would be quite confused to find it near the beginning, not the end, of a list of names, for example.

Support for Different National Character Set Support, Keyboards, and Fonts

In standard Windows 3.1 there is no way to use fonts native to the Eastern European countries such as Greece, Russia, or Turkey. For instance, if a user tries to install a Russian font on an English or French version of Windows 3.1, the characters appear unintelligible on the screen and the user is unable to use the font. Therefore, a special English Eastern European version of Windows 3.1 was designed for English users who needed to use the Eastern European fonts, including Russian Cyrillic or Greek. The English Eastern European version of Windows 3.1 offered the same capabilities as the true Russian or Turkish EE (Eastern European) version of Windows for displaying font and character information.

International Language Solution: Multilingual Content Support

Windows 95 resolves many problems related to international language issues by integrating multilingual content support in the core of the operating system. Windows 95 also offers national language support to software developers as a series of APIs that are part of the Win32 API set.

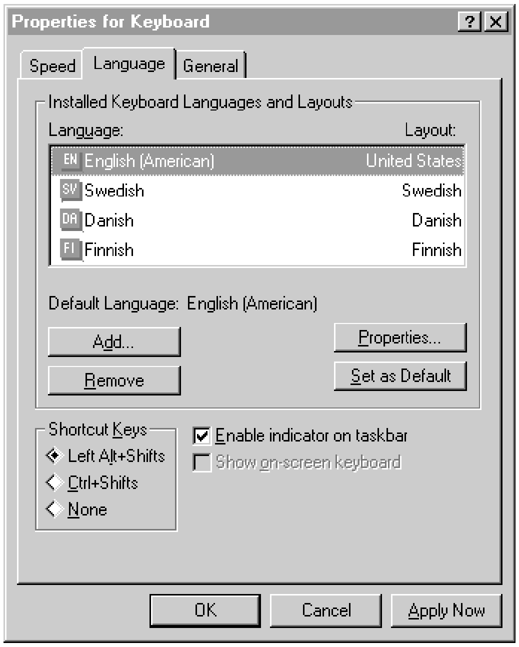
What is Multilingual Content?

Multilingual content support is the ability to display and edit text of various languages and scripts in a single document. Multilingual content support is a core feature in the Windows 95 product and will be supported in the next major release of Windows NT (code named “Cairo”).

Multilingual content support in an application has two major benefits. The first is that users who need to deal with content in multiple languages and scripts and exchange these documents with users on other language systems can do so. This is an important feature within the European Union, for example, where Greek- and Latin-based languages must coexist in documents. The second benefit is that an application which supports multilingual content will support the native content of any market into which it is sold. A multilingual application is a great application for the world.

Switching Between Languages and Keyboards the Easy Way

Windows 95 allows the user to add support for multiple keyboard layouts to match different international conventions. In Control Panel, the Keyboard icon provides the ability to configure the system to support the preferred keyboard layouts as shown in Figure 1.



Language:	Layout:
EN English (American)	United States
SV Swedish	Swedish
DA Danish	Danish
FI Finnish	Finnish

Default Language: English (American)

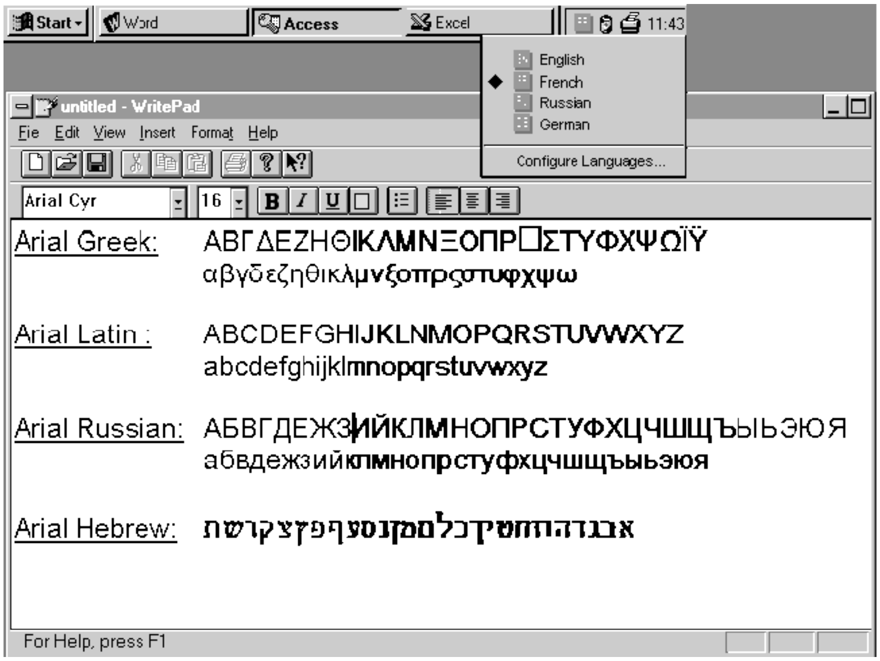
Shortcut Keys

● Left Alt+Shifts
○ Ctrl+Shifts
○ None

☑ Enable indicator on taskbar
☐ Show on-screen keyboard

Figure 1. Keyboard Properties Dialog Showing International Layout Support

Under Windows 3.1, to change the keyboard layout a user would go to Control Panel *each time* he or she wanted to change to a different keyboard format. Windows 95 makes this even easier. Figure 2 shows a sample legacy word processing document that illustrates the ability to integrate text using the Arial font in different languages within the same document. The language identifier in the status area of the taskbar allows the user to easily switch the system language between the available language options. A Windows 95 application that uses NLS APIs would incorporate the ability to switch the preferred language directly on the toolbar of the application.



Arial Greek: ΑΒΓΔΕΖΗΘΙΚΛΜΝΞΟΠΡ□ΣΤΥΦΧΨΩΪΫ
αβγδεζηθικλμνξοπρςστυφχψω

Arial Latin : ABCDEFGHIJKLNMOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz

Arial Russian: АБВГДЕЖЗИЙКЛМНОПРСТУФХЦЧШЩЪЫЬЭЮЯ
абвдежзийклмнопрстуфхцчшщъыьэюя

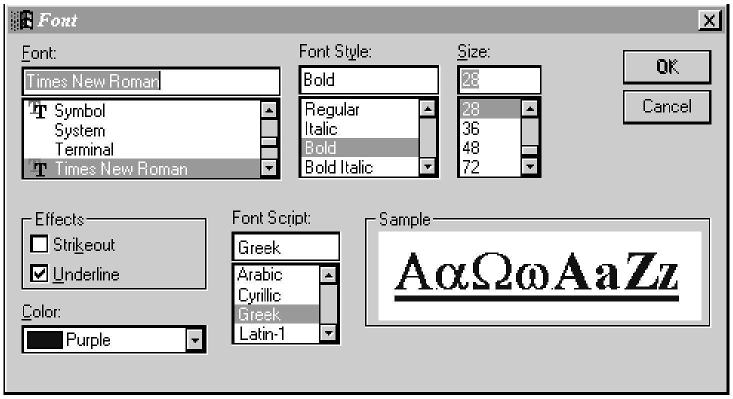
Arial Hebrew: אבגדהוזחטיךכלםמןנסעףפץצקרשת

Figure 2. Windows 95 Makes it Easy to Switch Between Different Languages to Create Multilingual Documents

Multilingual Extensions to the ChooseFont Dialog Box

The ChooseFont common dialog is now enhanced to include a list box showing the character set scripts supported by a particular font. This allows for proper representation of fonts for a given language.

Figure 3 shows an early representation of the new ChooseFont common dialog box, illustrating the integration of font script selection options. The scripts list shows the script names for each of the character sets covered by the font selected in the Font control. The preview window displays a font sample that is dependent on the script selected, as well as the other font attributes. The sample preview string is specific to the character set chosen by the user, showing what each of the different scripts look like.



Font: Times New Roman

Font Style: Bold

Size: 28

Effects

☐ Strikeout

☑ Underline

Color: Purple

Font Script: Greek

Figure 3. Windows 95 Displays the Available Font Selections for a Given Font Script Chosen by the User

Internationally-aware applications can support multilingual font selection by using the ChooseFont common dialog box (thus allowing users to select fonts) and by recognizing the extensions to the ChooseFont data structures in Windows 95. Even Windows–based applications—which, though not originally designed for Windows 95, still support formatted text, but not multilingual messages—may gain some basic level of support for multilingual content. If an application uses the ChooseFont common dialog box, it benefits from the enhancements, allowing users to select from the full range of character sets and fonts configured in the system. As long as the application saves the complete logical font data structure representation for fonts, an existing Windows–based application can get by without being aware that the font change a user has made includes a possible change of character set. (Applications do generally save this data, at least when saving text in their native format. Fewer save this when writing to interchange formats such as RTF.)

Multilingual Support for Exchanging Information Via the ClipBoard

A good multilingual-aware application should can exchange multilingual content with other aware applications and can exchange appropriate flat text to others, within the limitations of the ASCII text formats. Windows 95 provides special support in the data exchange APIs to pass language information along with the rich text data.

Win32 National Language Support APIs

When they install Windows 95, users specify a locale preference. (This preference can be changed later via Control Panel.) The Win32 NLS APIs can use the user' default locale settings or a specific locale setting.

By using the Win32 NLS APIs, developers can easily integrate international language support in their Win32–based applications. These APIs are supported both on the Windows 95 and Windows NT operating system platforms (with limited support available with Win32s under Windows 3.1) and allow applications to properly retrieve regional and language settings, format date and time, sort lists according to cultural rules, compare and map strings, and determine character type information.

This means that a developer in the U.S. can be assured that the sorting order and date formats that Microsoft provides with the operating system are correct. Therefore, the developer has only to use the appropriate Win32 NLS APIs to sort or display information.

By using the Win32 NLS APIs, developers can more easily develop applications for new global markets. Using this API set also lowers development costs by eliminating the need for proprietary sorting methods, parsing the WIN.INI file or Registry, and locale-specific coding. Perhaps more importantly for developers, the API set provides a mechanism for accurate and consistent behavior on each 32-bit Microsoft Windows platforms.

End-users benefit because the API set ensures that information is handled and displayed properly for a given locale-specific format. In addition, users don't have to worry as to whether their international text is being sorted properly.

Try It!

Mouse

To be able to examine the improvements present in Windows 95 to switch between language types and keyboard layouts, you’ve got to Try It!

Support for Multilingual Content

To demonstrate the multilingual content support in Windows 95, try these simple procedures to install different language support.

- In Control Panel, open the Keyboard icon and click the Language tab. Add a couple of keyboards (for example, Swedish and French) and then choose OK. On the right corner of the taskbar, notice a small square is displayed in the status area to represent the active keyboard layout—two letters are displayed in the square to represent the language, for instance EN for English.
- Start a word processing application and try to create a document in which to try the multilingual content support. To hot-switch between different input languages, press ALT+SHIFT, and toggle through the available configured languages. Once you have switched to a new input language, type something (a multilingual-aware application would automatically switch the proper font). Of course you have to know where the keys are on that country-specific keyboard layout.

In a true multilingual-aware application, you can scroll the text and the application will automatically switch the current input language to match the proper language format when you scroll through different languages in the text.