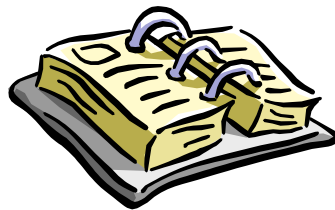


PHẦN MỀM XỬ LÝ TIẾNG VIỆT UNICODE/TCVN 6909
(CHO CÁC GIAO DỊCH VĂN BẢN ĐIỆN TỬ)

Hướng dẫn sử dụng Vietkey



© 2002, Vietkey Group, Vietkhang JSC
32 Hàng Cót, Hà Nội, +84(4) 8267593
155 Lô B4, Sư Vạn Hạnh (Nối dài), Q10, Tp HCM, +84(8) 8640401
<http://www.vietkey.net> , unicode@vietkey.net

MỤC LỤC

I. GIỚI THIỆU CHUNG	4
I.1. Đặt vấn đề	4
I.2. Phần mềm hỗ trợ xử lý tiếng ViệtUnicode.....	5
I.3. Yêu cầu cho việc triển khai TCVN 6909	6
I.3.1 Yêu cầu về thiết bị phần cứng	6
I.3.2 Yêu cầu về phần mềm	6
II. BỘ GỖ TIẾNG VIỆT ĐA NĂNG VIETKEY	8
II.1. Giới thiệu chung về bộ gõ	8
II.2. Cài đặt Bộ gõ tiếng Việt VIETKEY	9
II.2.1 Cài đặt bộ gõ tiếng Việt tối thiểu.....	9
II.2.2 Cài đặt bộ gõ đầy đủ bao gồm cả các font chữ.....	9
II.3. Nhập thông tin (gõ) tiếng Việt bằng vietkey.....	13
II.3.1 Chọn font chữ tiếng Việt trong các ứng dụng	13
II.3.2 Chọn bảng mã tiếng Việt trong Vietkey 2000	14
II.3.3 Chọn kiểu gõ tiếng Việt	16
II.3.4 Quy tắc gõ tiếng Việt	17
II.4. Chuyển chế độ gõ tiếng Việt, tiếng Anh:.....	21
II.5. Thiết lập gõ tắt – AutoText	22
II.6. Tự tạo biểu tượng cho Vietkey.....	24
II.7. Đặt chế độ tự khởi động Vietkey	24
II.8. Thay đổi giao diện hiển thị tiếng Việt trong bảng điều khiển của Vietkey.	25
II.9. Khởi động nhanh Vietkey vào thanh Taskbar hay OnTop.....	26
II.10. Đặt chế độ của MS Word, Excel để gõ được tiếng Việt TCVN 6909.....	27
III. BỘ XỬ LÝ TIẾNG VIỆT VĂN PHÒNG VIETKEY OFFICE	30
III.1. Giới thiệu chung về Vietkey Office	30
III.2. Cài đặt Bộ xử lý tiếng Việt trong MS Office	31
III.3. Kiểm tra chính tả tiếng Việt.....	33
III.4. Công cụ Sắp xếp tiếng Việt.....	35
III.5. Chuyển đổi chữ hoa/chữ thường	39

III.6.	Chuyển mã văn bản cho các văn bản word /excel /powerpoint	40
IV.	CHUYỂN MÃ CHO CÁC TRANG WEB	46
IV.1.	Giới thiệu chung	46
IV.2.	Cài đặt công cụ chuyển mã tiếng Việt cho các trang Web	46
IV.3.	Hướng dẫn sử dụng chuyển mã các trang Web.....	46
V.	BỘ THƯ VIỆN LẬP TRÌNH XỬ LÝ TIẾNG VIỆT	49
VI.	CÁC CÂU HỎI THƯỜNG GẶP	50
VI.1.	Làm gì khi gặp khó khăn trong việc dùng Vietkey?.....	50
VI.2.	Các bản nâng cấp và các thông tin mới về sản phẩm có ở đâu?	50
VI.3.	Tại sao khi gõ tiếng Việt, chữ có dấu đôi khi bị trên 1 ký tự trống?.....	50
VI.4.	Tại sao trong một số chữ tiếng Việt chữ ‘i’ biến thành chữ ‘I’?	51
VI.5.	Lỗi máy in khi in chữ Unicode?	51
VI.6.	Nếu máy tính không có ổ đọc đĩa CD cần phải làm thế nào?	53
VI.7.	Trong thư tín điện tử muốn dùng tiếng Việt Unicode cần phải làm gì?.....	54
VI.8.	Cách soạn thảo các trang Web Unicode tiếng Việt.....	54
VI.9.	Thiết lập chế độ hiển thị Unicode cho các trình duyệt Web?	55
VI.10.	Soạn thảo tiếng Việt Unicode trong một số phần mềm khác?	56
VI.11.	Unicode trong các môi trường lập trình	57
VI.12.	Unicode là gì?	57
VI.13.	UTF-8 là gì?.....	58
VI.14.	Tiếng Việt được bố trí như thế nào trong chuẩn Unicode?	58
VI.15.	Tại sao tiếng Việt lại không được xếp vào một vùng liên tục?	59
VI.16.	Ưu điểm chính của Unicode là gì?	60
VI.17.	Nhược điểm chính của Unicode là gì?	60
VI.18.	Các file Unicode có kích thước từ 2 đến 256 lần so với 8-Bit?.....	61
VI.19.	Tại sao không thể gõ được Unicode trong một số ứng dụng?.....	61
VI.20.	Muốn sử dụng Unicode phải nâng cấp máy tính lên Windows2000?.....	61
VII.	PHỤ LỤC	63
VII.1.	Địa chỉ liên hệ.....	63
VII.2.	Tác giả	64

I. GIỚI THIỆU CHUNG

I.1. ĐẶT VẤN ĐỀ

Trong thời gian vừa qua, việc không thống nhất sử dụng một bộ mã trong cả nước (miền Bắc dùng Font ABC, miền Nam Font VNI) đã gây khó khăn không nhỏ trong việc trao đổi thông tin dữ liệu giữa các đơn vị với nhau, đồng thời cũng gây nhiều trở ngại cho việc triển khai ứng dụng trên mạng Internet và mạng điện rộng.

Ngày 24/9/2001, Bộ Khoa học Công nghệ và Môi trường đã ban hành tiêu chuẩn Việt Nam: TCVN 6909:2001 dựa trên bảng mã Unicode và ISO 10646 và Quyết định số 72 của Thủ tướng Chính phủ về việc thống nhất sử dụng bảng mã tiếng Việt trong khối cơ quan Nhà nước là những cơ sở pháp lý và điều kiện thuận lợi cho việc thống nhất một bảng mã tiếng Việt duy nhất trong cả nước.

Ưu điểm chính của bộ mã tiếng Việt Unicode/TCVN 6909:2001:

- Có đầy đủ chữ hoa và chữ thường tiếng Việt.
- Cho phép tiếng Việt hội nhập với tất cả các ngôn ngữ chính khác trên thế giới.
- Không bị các lỗi mất chữ hoặc tranh chấp với các ký tự điều khiển.
- Đảm bảo tính tuân thủ các tiêu chuẩn ISO 10646 và UNICODE, là những chuẩn được áp dụng trong các công nghệ tiên tiến. Được các hãng cam kết hỗ trợ trong các sản phẩm mới.

1.2. PHẦN MỀM HỖ TRỢ XỬ LÝ TIẾNG VIỆT UNICODE

Vietkey Group /Vietkhang JSC là đơn vị có hơn 10 năm kinh nghiệm xây dựng các công cụ xử lý tiếng Việt cho các máy tính PC. Ngày nay Vietkey Group có phần mềm xử lý tiếng Việt cho rất nhiều môi trường khác nhau: MS Windows, Unix/Linux, Pocket PC, Palm...Các công cụ xử lý tiếng Việt của Vietkey Group rất đa dạng, phong phú và hoàn chỉnh, từ bộ gõ tiếng Việt cho máy Palm cho đến hệ điều hành thương hiệu Việt Nam hoàn chỉnh Vietkey Linux. Trong tài liệu này chúng tôi chủ yếu đề cập đến các công cụ xử lý tiếng Việt trong môi trường Windows:

- Công cụ nhập thông tin từ bàn phím (bộ gõ tiếng Việt Vietkey)
- Công cụ chuyển đổi chữ hoa/chữ thường.
- Công cụ kiểm tra chính tả tiếng Việt.
- Công cụ sắp xếp tiếng Việt.
- Các công cụ chuyển mã văn bản, dữ liệu từ các bộ mã/bộ Font cũ sang tiêu chuẩn mới, chuyển mã các tệp văn bản (Word, Excel), các trang Web và cơ sở dữ liệu.
- Bộ thư viện lập trình xử lý tiếng Việt cho các môi trường lập trình.

VietKey đã được đăng ký bản quyền theo số 1671 VH/BQ ngày 22 tháng 12 năm 1997 tại Cục Bản Quyền Tác Giả - Bộ Văn Hóa. Bản quyền thuộc về tác giả Đặng Minh Tuấn, trưởng nhóm Vietkey Group. Vietkhang JSC là công ty được uỷ quyền kinh doanh và phân phối các sản phẩm họ Vietkey. Mọi sửa đổi không được phép đối với các file chương trình, logo, hướng dẫn (Help) của VietKey đều vi phạm luật bản quyền.

Khi bộ sản phẩm Vietkey được bán ra, theo quy định, người sử dụng chỉ được phép cài đặt trên từng máy đơn lẻ. Cài đặt trên mạng cục bộ hoặc cho số lượng lớn trong một công ty cần phải có giấy bản quyền sử dụng từ phía nhà phân phối.

1.3. YÊU CẦU CHO VIỆC TRIỂN KHAI TCVN 6909

1.3.1 YÊU CẦU VỀ THIẾT BỊ PHẦN CỨNG

- Máy tính từ 386 trở lên.
- Bộ nhớ RAM: 16 MB trở lên.
- Ổ cứng HDD: 300 MB (để có thể cài tối thiểu được Win95 và Office 97).
- Không gian đĩa cho bộ gõ Vietkey: 150KB, Font Unicode và bộ Font ABC, VNI: 10MB, bộ Vietkey Office: 1MB.

1.3.2 YÊU CẦU VỀ PHẦN MỀM

- **Hệ điều hành:** Windows 95/98, Windows NT, Windows 2000, Windows XP, Pocket PC 2002 hoặc Linux có kernel 2.0 trở lên.
- **Soạn thảo văn bản:** Microsoft Office 97/2000/XP, Open Office.
- **Trình duyệt Web:** Internet Explore 4.0 trở lên, Netscape 4.7 trở lên.

- **Thư tín điện tử:** Outlook Express 5 trở lên, Netscape Mail 4.7 trở lên.
- **Cơ sở dữ liệu:** MS Access 2000, SQL Server 6.5 trở lên, ORACLE...

II. BỘ GỠ TIẾNG VIỆT ĐA NĂNG VIETKEY

II.1. GIỚI THIỆU CHUNG VỀ BỘ GỠ

Bộ gõ Vietkey là bộ gõ được ra đời từ rất sớm, phiên bản 1.0 phát hành năm 1993. Vietkey là một trong những bộ gõ tiếng Việt 32-Bit đầu tiên và cũng là một trong những bộ gõ đầu tiên hỗ trợ Unicode. Bộ gõ Vietkey có nhiều tính năng ưu Việt:

- Nhỏ gọn, chỉ cần File duy nhất.
- Chạy trên tất cả các môi trường Windows 32-Bit, bao gồm: Windows 95/98, Windows NT, Windows 2000, Windows XP. Đối với MS Pocket PC cần phải có phiên bản khác là Vietkey CE.
- Hỗ trợ gõ nhiều ngôn ngữ phổ thông: Việt, Nga, Anh, Pháp, Đức...
- Hỗ trợ các bảng mã tiếng Việt phổ thông: ABC, VNI...Unicode/ TCVN 6909:2001. Riêng với Unicode, hỗ trợ các biểu diễn UTF-8, UCS2 (cho các môi trường Web), C-Style: Unicode cho môi trường lập trình C, C++.
- Hỗ trợ nhiều kiểu gõ thông dụng ở Việt Nam cũng như ở nước ngoài: TELEX, VNI, VIQR.
- Có nhiều tính năng nâng cao: gõ tắt các từ mục không hạn chế, tự động chuyển chế độ Anh-Việt với kiểu gõ Telex, cho phép gõ lặp dấu-khử dấu là tính năng cũng do Vietkey Group cải tiến và đưa ra đầu tiên ở Việt Nam (1991).

II.2. CÀI ĐẶT BỘ GỠ TIẾNG VIỆT VIETKEY

II.2.1 CÀI ĐẶT BỘ GỠ TIẾNG VIỆT TỐI THIỂU

Bộ gõ tiếng Việt **Vietkey** có kích thước 150K, không cần cài đặt, chỉ cần copy tệp VKNT.EXE vào một thư mục bất kỳ trên ổ cứng (ví dụ thư mục C:\vietkey) và nhấp chuột vào biểu tượng của Vietkey, bộ gõ tiếng Việt sẽ được kích hoạt. Với kích thước 150 KB, Vietkey dễ dàng chuyển mang trên mạng hoặc copy trên đĩa mềm để cài đặt vào các máy mới.

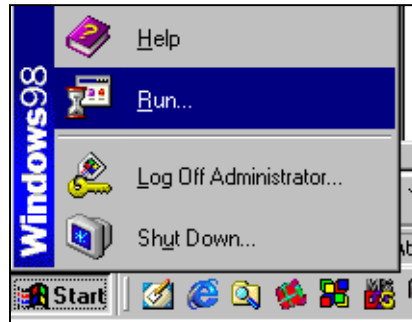
II.2.2 CÀI ĐẶT BỘ GỠ ĐẦY ĐỦ BAO GỒM CẢ CÁC FONT CHỮ

Bộ Vietkey đầy đủ (Full Version) có hơn 7 đĩa mềm hoặc được lưu trên đĩa CD có các thư mục con là disk1, disk2, disk3, disk4, disk5, disk6, disk7. Bộ đầy đủ bao gồm các font chữ Unicode /TCVN 6909 và các bộ font cũ (được dùng trong thời gian quá độ) : ABC và VNI.

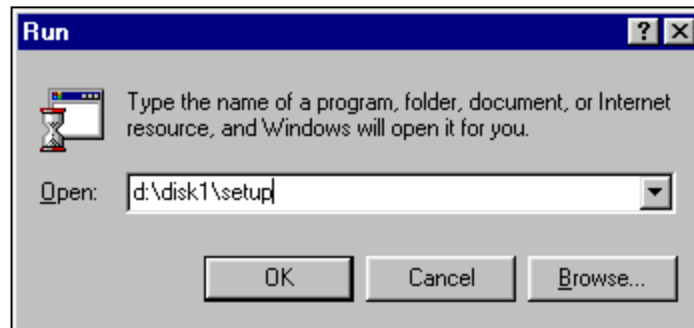
Các bước cài đặt:

1. Cho đĩa Vietkey vào ổ CD, hoặc cho đĩa Vietkey 1 (nếu dùng đĩa mềm), hoặc chọn vào thư mục Disk1 nếu đã copy trên đĩa cứng.

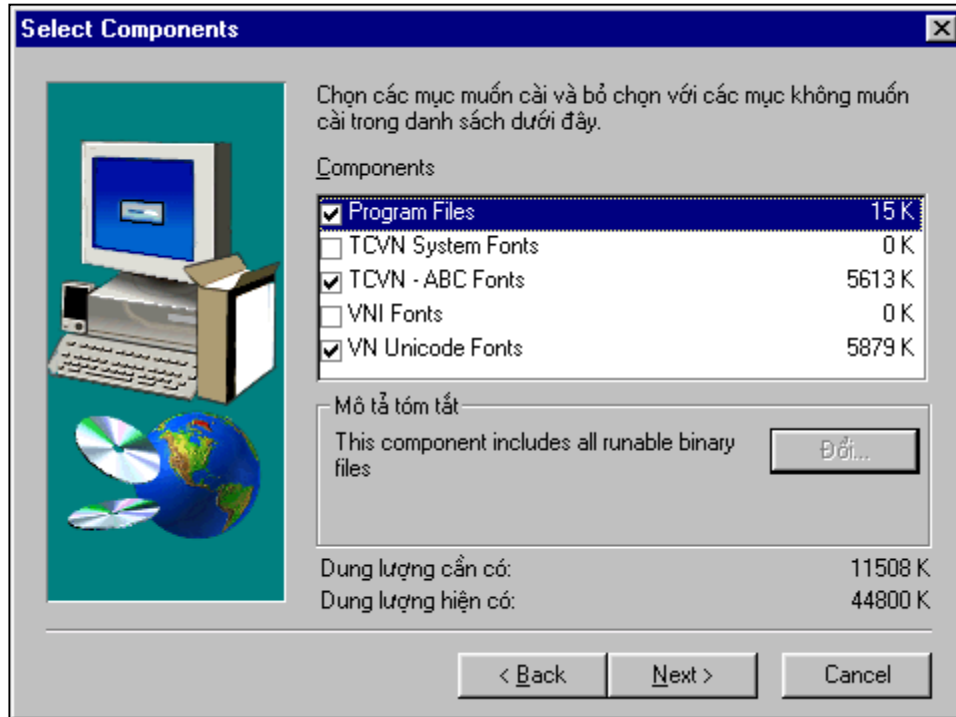
2. Nhấp chuột (Mouse) vào nút **start**, một cửa sổ xuất hiện, chọn vào mục **Run** như hình vẽ sau:



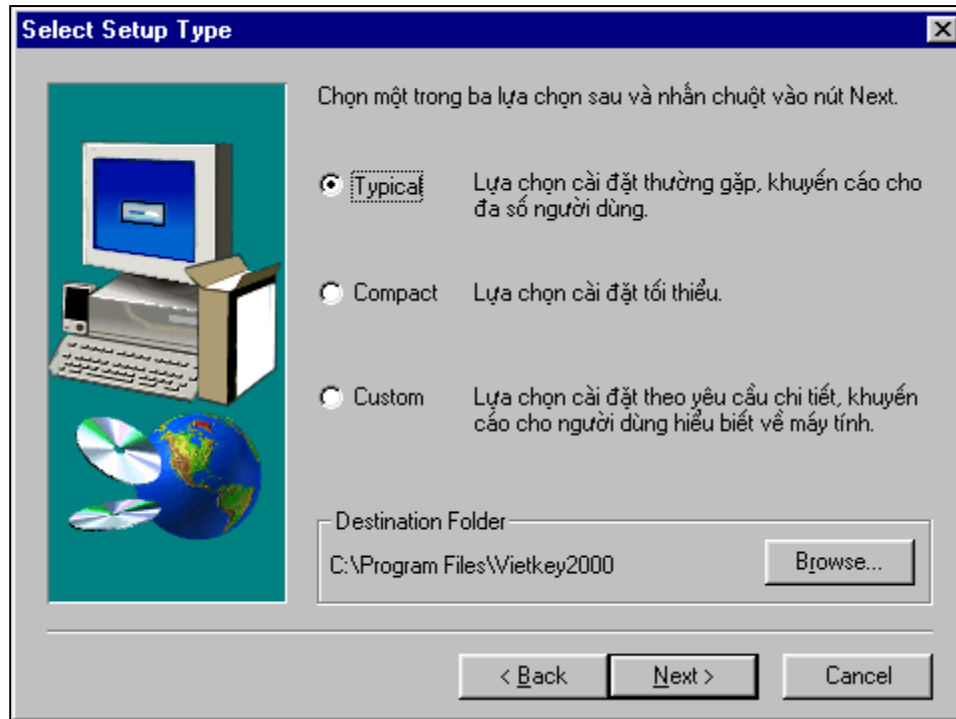
3. Sau khi cửa sổ **Run** xuất hiện, gõ trong ô text **d:\disk1\setup** nếu ổ CD được định nghĩa là ổ D, hoặc có thể nhấp chuột vào nút **Browse** để chọn đúng tệp **setup.exe** trong đĩa 1 của bộ cài đặt Vietkey nằm trên đĩa cứng hoặc trên đĩa mềm A như hình dưới đây:



4. Sau khi chương trình cài đặt **Setup.exe** chạy, nhấp chuột vào nút **Next** 2 lần cho đến khi hiện ra cửa sổ **Select Setup Type**, bình thường chương trình và các font chữ sẽ được cài đặt vào thư mục **c:\program files\vietkey2000**, để thay đổi thư mục khác có thể nhấp chuột vào nút **Browse**. Sau đó nhấp chuột vào mục Custom trong cửa sổ như hình sau:



5. Sau khi chọn nút **Custom**, cửa sổ **Select Component** sẽ xuất hiện như hình sau:



- Trong mục **Component**, nhấp chuột vào các mục font cần cài đặt, chú ý cần nhấp chuột để có dấu tick (dấu hất hình chữ V) vào mục **VN Unicode Fonts** để có được font chữ Unicode trên các máy Windows 95, Windows 98, Windows NT. Với các hệ điều hành Windows 2000, Windows ME hoặc Windows XP, không cần chọn vào mục **VN Unicode Fonts** vì các Font này đã có sẵn trên hệ điều hành.

- Chú ý: chương trình cài đặt này cũng có thể cài đặt font hệ thống tiếng Việt cho Windows (font hệ thống là các font được dùng trong Menu, các thông báo...) tuy nhiên Font hệ thống tiếng Việt này mới chỉ hỗ trợ bộ mã tiếng Việt cũ là TCVN 5712, và font hệ thống này sẽ không có tác dụng đối với Windows 2000 hoặc Windows XP. Trong một số phiên bản Windows bản địa như Windows tiếng Nga, Windows tiếng Hoa, tiếng Pháp... sau khi cài đặt font hệ thống tiếng Việt (TCVN System Fonts) sẽ làm hỏng hết các thông báo bằng tiếng bản địa, vì vậy chỉ chọn mục này khi thực sự cần thiết.
6. Nhấp chuột vào các nút **Next** sau đó và cuối cùng là nút **Finish**.
 7. Khởi động lại máy và quá trình cài đặt font chữ và bộ gõ tiếng Việt sẽ được kết thúc.

II.3. NHẬP THÔNG TIN (GỠ) TIẾNG VIỆT BẰNG VIETKEY

Để nhập được tiếng Việt và hiển thị đúng trong các ứng dụng cần phải:

- *Chọn đúng Font chữ tiếng Việt trong các ứng dụng,*
- *Chọn đúng bảng mã (Character Set) trong Vietkey,*
- *Chọn đúng kiểu gõ trong Vietkey.*

II.3.1 CHON FONT CHỮ TIẾNG VIỆT TRONG CÁC ỨNG DỤNG

Trong các ứng dụng MS Word, MS Excel, MS PowerPoint, Outlook, FrontPage, Access 2000... với bộ mã UNICODE /TCVN 6909 cần phải chọn tên Font trong các hộp font chữ là một trong các font sau đây:

Arial, Times New Roman, Tahoma, Courier, Platino Linotype, Verdana...

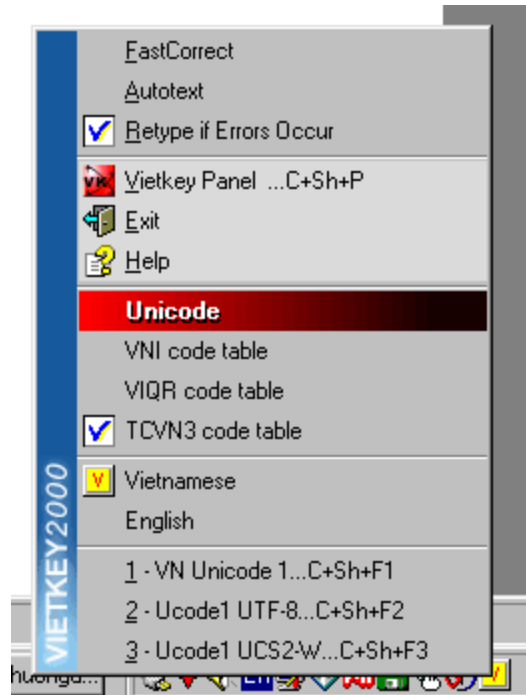
Chú ý: với các hệ điều hành Windows 2000, Windows ME, Windows XP thì các font nói trên mặc nhiên là Unicode, đối với các hệ điều hành cũ hơn như Windows 95/98 thì cần phải cài lại font chữ như mục II.1.

II.3.2 CHON BẢNG MÃ TIẾNG VIỆT TRONG VIETKEY 2000

- Khởi động **Vietkey**, bằng cách nhấp chuột vào biểu tượng **Vietkey**, cửa sổ điều khiển của Vietkey sẽ xuất hiện.
- Nhấp chuột vào nút **Char Sets** (Bảng mã).
- Sau đó chọn vào mục **Unicode Pre-Compound, TCVN 6909:2001 (Unicode dựng sẵn)**.



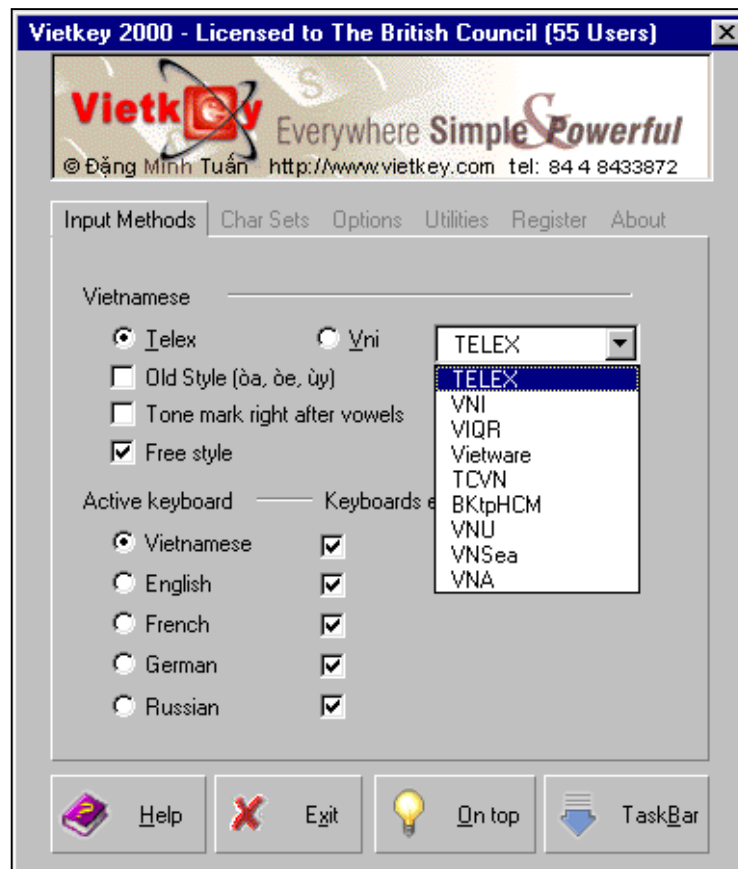
- Có thể chọn nhanh bảng mã tiếng Việt khi **Vietkey** đang nằm ở thanh **Taskbar** bằng cách nhấp chuột phải vào biểu tượng **Vietkey** trên thanh **Taskbar**, khi ấy một cửa sổ thực đơn xuất hiện và nhấp chuột vào dòng chữ **Unicode**.



- Vietkey cho phép chuyển nhanh các bảng mã bằng tổ hợp phím nóng. **Control+Shift+F1**, **Control+Shift+F2**, **Control+Shift+F3** Vietkey sẽ chuyển sang 3 bảng mã lần cuối cùng được chọn.

II.3.3 CHON KIỂU GỠ TIẾNG VIỆT

Trong cửa sổ của Vietkey, nhấp chuột vào mục **Kiểu gõ/Input Methods**, chọn kiểu gõ tương ứng là **TELEX** hay **VNI** tùy theo thói quen người dùng, ngoài 2 kiểu gõ chính phổ biến ở Việt Nam hiện nay, Vietkey còn cung cấp cho người dùng 9 kiểu gõ khác nhau là VIQR, Vietware, TCVN, BKtpHCM, VNU, VNSea, VNA.



II.3.4 QUY TẮC GỖ TIẾNG VIỆT

II.3.4.1. Quy tắc chung

1. Các dấu mũ và dấu trăng, dấu râu phải gõ trực tiếp vào nguyên âm, dấu (huyền, sắc, hỏi, ngã, nặng) có thể gõ trực tiếp hoặc sau nguyên âm cần bỏ dấu từ 1 đến 3 ký tự tức là có thể bỏ dấu Việt ngay tại nguyên âm hoặc ở cuối từ (bỏ dấu tự động). Nên bỏ dấu tự động để tránh bỏ sai dấu hoặc tránh lỗi 1 từ có nhiều dấu và với cách bỏ dấu tự động, dấu bao giờ cũng thống nhất do đó việc tìm kiếm, sắp xếp tiếng Việt mới có thể thực hiện được.
2. Trong trường hợp gõ sai dấu có thể gõ ngay dấu khác không cần phải xoá chữ để gõ lại.
3. Các phím dấu chỉ có tác dụng theo ngữ cảnh tức là nếu không có nguyên âm nào trong vùng tác dụng thì nó vẫn hiển thị như trong chế độ tiếng Anh, ví dụ phím ‘f’ nếu đi sau chữ ‘a’ thì sẽ thành chữ ‘à’, còn nếu gõ riêng nó vẫn hiện chữ f, tương tự như vậy trong kiểu gõ VNI, các phím số vẫn hiện là số nếu nó không đi sau nguyên âm có khả năng có dấu.
4. Với kiểu bỏ dấu tự động các phím dấu chỉ có tác dụng khi đằng sau nguyên âm có dấu chỉ có những phụ âm hợp lệ: <c,g,h,m,n,p,t>.

II.3.4.2. Quy tắc gõ tiếng Việt kiểu Telex**Qui ước, ý nghĩa của các phím:**

Phím	Tác động dấu
f	Huyền
s	Sắc
r	Hỏi
x	Ngã
J	Nặng
w	Dấu trắng trong chữ ‘ ă ’, dấu móc trong các chữ ‘ ơ ’, ‘ ư ’. Gõ phím ‘ w ’ đơn lẻ sẽ ra chữ ‘ ư ’
z	Khử dấu (xoá dấu) đã đặt
aa	â
aw	ã
ee	ê
oo	ô
dd	đ
ow,]	ơ
w, uw, [ư

Lập dấu:

Phím	Kết quả
ddd	dd
ooo	oo
eee	ee
[[[
]]]

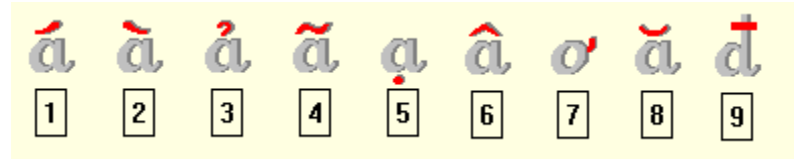
Cặp chữ **ươ** rất hay gặp trong tiếng Việt, để gõ nhanh có thể gõ 2 phím][gần nhau và theo chiều hướng vào tâm bàn phím nên đạt được tốc độ cao do giảm được một nửa số thao tác của các phím cách xa nhau uwow.

Ví dụ:

Gõ dòng chữ	Bằng dãy các phím
Nước chảy đá mòn	- Nwowes chayr ddas monf - N][cs chary ddas mofn
Thòong	- Thooongf - Tho<control>ongf
Đường	- DDUwowngf - DDwowngf - dD][ngf

II.3.4.3. Quy tắc gõ tiếng Việt kiểu VNI

Qui ước, ý nghĩa của các phím:



Phím	Tác động dấu
1	Sắc
2	Huyền
3	Hỏi
4	Ngã
5	Nặng
6	Dấu mũ của các chữ ‘â’, ‘ê’ và ‘ô’
7	Dấu râu của chữ ‘ơ’ và ‘ư’
8	Dấu trắng của chữ ‘ă’
d9	đ
0	Khử dấu (xoá dấu)

Ví dụ:

Gõ dòng chữ	Bảng dãy các phím
Nước chảy đá mòn	- Nu7o71c cha3y d9a1 mo2n - Nu7o7c1 chay3 d9a1 mon2
Thoòng	- Thoong2
Đường	- D9u7o7ng2
đã	- d9a4

Dùng phím <control> để gõ các chữ số và các ký tự <!,@.#,(,)...> sau các nguyên âm.

Ví dụ:

A! = **A<control>!**

II.4. CHUYỂN CHẾ ĐỘ GÕ TIẾNG VIỆT, TIẾNG ANH:

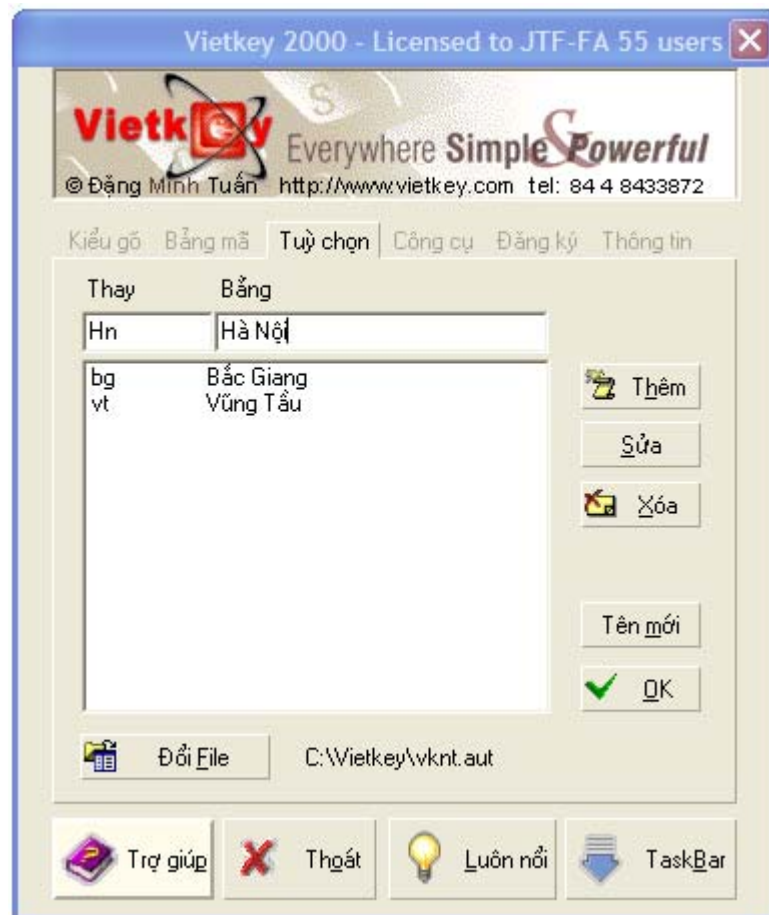
1. Có thể chuyển chế độ gõ tiếng **Việt** sang chế độ gõ tiếng **Anh** (hay Nga, Pháp, Đức) bằng tổ hợp phím **ALT-Z**. Tổ hợp phím chuyển này có thể thay bằng tổ hợp phím khác <**Control-Control**> hoặc <**Control-Shift**> trong mục **Tùy chọn (Option)**. Thay đổi tổ hợp phím chuyển chế độ nhằm tránh các phím tranh chấp trong một số phần mềm chuyên dụng như Photoshop...
2. Nhấp nút phải chuột vào biểu tượng của **Vietkey** trên thanh **Taskbar**, chọn chế độ gõ **Việt-Anh**.

II.5. THIẾT LẬP GỖ TẮT – AUTOTEXT

- Trong mục **Tùy chọn (Option)** nhấp chuột để bật chế độ **Cho phép gõ tắt (Enable AutoText)**.



- Nhấp chuột vào nút **Định nghĩa gõ tắt (AutoText Items)** để thiết lập các mục từ gõ tắt.



- Gõ tắt là chế độ cho phép thay một chuỗi ký tự dài bằng một chuỗi ký tự ngắn hơn. Gõ tắt cho phép tăng tốc độ gõ, nhập văn bản, đồng thời có thể chuẩn hoá được dữ liệu, thống nhất cho một cơ quan. Ví dụ chuỗi “hn” có thể thay bằng chuỗi “Hà Nội”, nhờ việc gõ tắt có thể chuẩn hoá được khâu nhập dữ liệu, tránh các tổ hợp khác nhau như “hà nội”, “Hà nội”. Việc không chuẩn hoá, không thống nhất sẽ gây nhiều khó khăn trong việc tìm kiếm, thống kê...

- Vietkey cho phép định nghĩa không hạn chế các mục gõ tắt **thêm, sửa, xoá (Add, Modify, Delete)**, đồng thời cũng cho phép lưu các mục gõ tắt ra các File khác nhau **Tên mới (Save as)**.
- Nhờ việc **Đổi File(Change File)** Vietkey cho phép có thể thay đổi các bộ gõ tắt một cách nhanh chóng và tiện lợi, tùy theo lĩnh vực đang xử lý văn bản. Các File gõ tắt này có thể được lưu trữ, trao đổi cho nhiều người dùng chung một quy ước gõ tắt nhờ vậy có thể góp phần thống nhất khâu nhập dữ liệu trong phạm vi cơ quan tổ chức.. Ví dụ “tp” trong một số trường hợp có thể là gõ tắt của “Thành phố”, nhưng trong File gõ tắt khác có thể là “thị phần”.

II.6. TỰ TẠO BIỂU TƯỢNG CHO VIETKEY.

- Vào mục **Utilities (Công cụ)** chọn **Create Vietkey Icon on Desktop (Tạo biểu tượng Vietkey trên Desktop)**.

II.7. ĐẶT CHẾ ĐỘ TỰ KHỞI ĐỘNG VIETKEY

- Vào mục **Utilities (Công cụ)** chọn **Run Vietkey at Boot Time (Luôn chạy Vietkey khi khởi động PC)**.



II.8. THAY ĐỔI GIAO DIỆN HIỂN THỊ TIẾNG VIỆT TRONG BẢNG ĐIỀU KHIỂN CỦA VIETKEY.

- Vào mục **Tùy chọn (Option)** chọn **Vietnamese interface (Hiển thị tiếng Việt)**.

II.9. KHỞI ĐỘNG NHANH VIETKEY VÀO THANH TASKBAR HAY ONTOP.

1. Vào mục **Tùy chọn (Option)** chọn **Khởi động nhanh (FastLoad)**

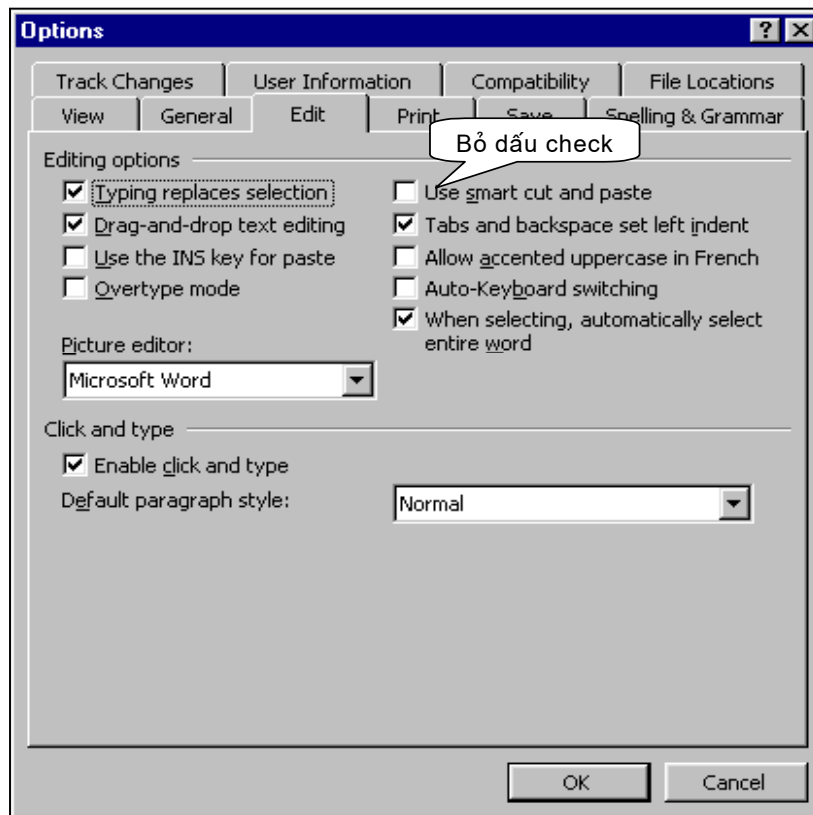


2. Nhấp chuột vào nút **Taskbar** hoặc **Luôn nổi (OnTop)**.

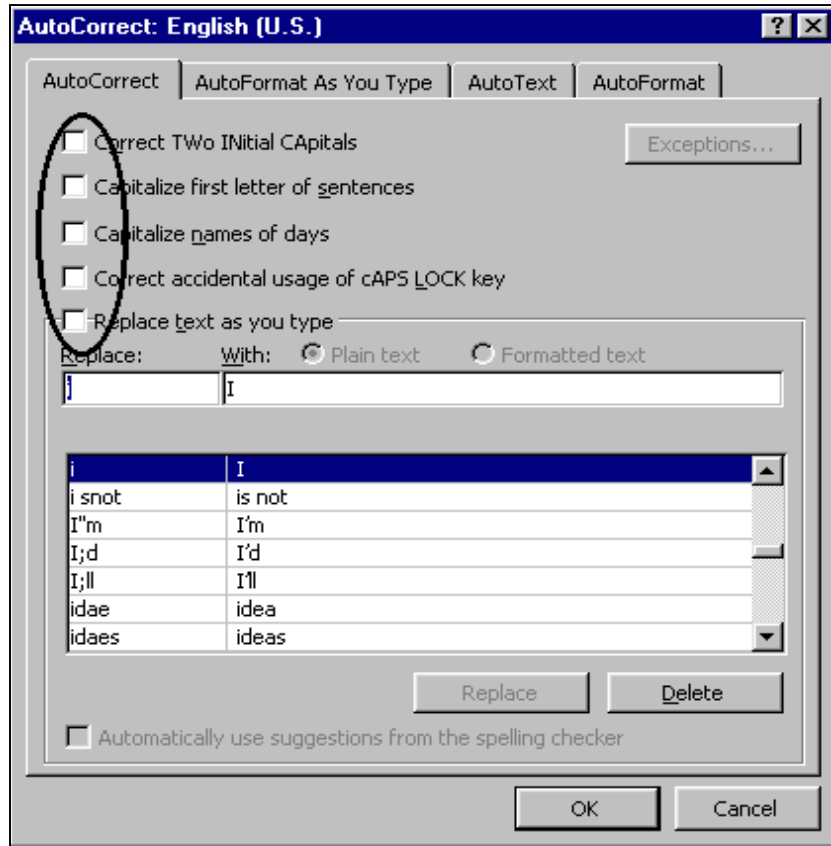
Bình thường khi khởi động, bảng điều khiển của **Vietkey** sẽ luôn xuất hiện trên màn hình, dùng tính năng này để **Vietkey** bỏ qua bước hiển thị bảng điều khiển, vào ngay thanh Taskbar hoặc luôn nổi biểu tượng trên màn hình.

II.10. ĐẶT CHẾ ĐỘ CỦA MS WORD, EXCEL ĐỂ GỠ ĐƯỢC TIẾNG VIỆT TCVN 6909

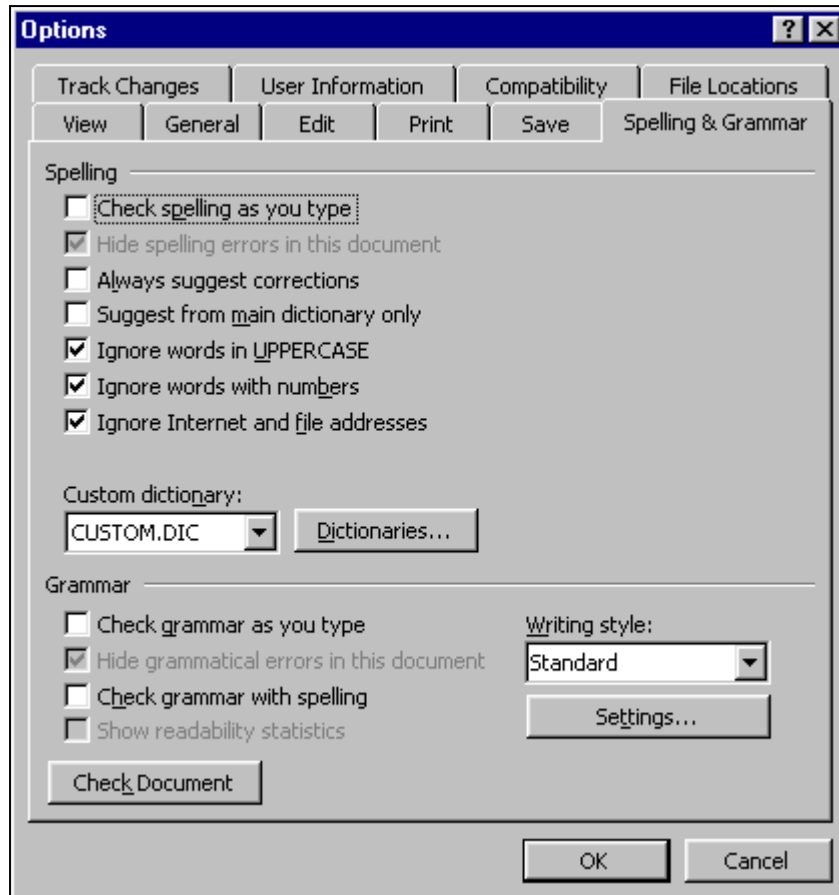
Do các phần mềm Word, Excel, PowerPoint... được thiết kế để phù hợp với ngôn ngữ không phải tiếng Việt (tiếng Anh, Nga, Pháp...) một số chức năng tự động hiệu chỉnh, hoặc kiểm tra chính tả của các ngôn ngữ này sẽ không tương hợp với ngôn ngữ tiếng Việt nên có thể gây ra một số trục trặc khi nhập dữ liệu tiếng Việt. Sau đây là cách đặt một số chế độ để tránh sự không tương thích này:



- Trong cửa sổ của MS Word chọn **Tools/Option/Edit** bỏ chức năng **Use smart cut and Paste**. Bỏ chức năng này để tránh hiện tượng cách dấu trống với các chữ có dấu.
- Trong cửa sổ của MS Word chọn **Tools/AutoCorrect** bỏ hết các tính năng trong cửa sổ **AutoCorrect**.



- Trong cửa sổ của MS Word chọn **Tools/Option/Spelling & Grammar** bỏ các tính năng: **Check Spelling as You Type, Check Grammar as You Type, Check Grammar with Spelling**. Như hình vẽ dưới đây.



III. BỘ XỬ LÝ TIẾNG VIỆT VĂN PHÒNG VIETKEY OFFICE

III.1. GIỚI THIỆU CHUNG VỀ VIETKEY OFFICE

Vietkey Office là một bộ công cụ xử lý tiếng Việt được tích hợp vào các phiên bản MS Office 97, MS Office 2000, MS Office XP. Sau khi cài đặt Vietkey Office, các hệ soạn thảo văn phòng (thường là các bản cho tiếng Anh, Pháp...) sẽ biến thành các hệ soạn thảo văn phòng có hỗ trợ xử lý tiếng Việt. Các chức năng cơ bản của Vietkey Office (tương tự như đã có trong bộ MS Office với tiếng Anh):

- Kiểm tra chính tả (spelling) tiếng Việt,
- Sắp xếp (sorting) tiếng Việt,
- Chuyển đổi chữ hoa/chữ thường (Change case) tiếng Việt.
- Ngoài ra Vietkey Office còn có chức năng rất quan trọng là chuyển mã các văn bản tiếng Việt.
 - o Bộ chuyển mã Vietkey Office (VKO) cho phép chuyển mã các văn bản được soạn thảo trong các chương trình MS Word 97/2K/XP, Excel, PowerPoint (trong bộ MS Office) từ 44 bảng mã tiếng Việt khác nhau sang Unicode và ngược lại.
 - o So với các bộ chuyển mã khác thì VKO cho phép chuyển đổi mã tiếng Việt mà vẫn giữ được hầu như nguyên vẹn tất cả các thuộc tính cũ của font chữ, kể cả tên, họ font tương ứng (mặc dù font đích và font nguồn có tên hoàn toàn khác nhau). Tất cả các đối tượng như bảng biểu, đồ họa vẫn được giữ nguyên, đặc biệt VKO còn cho phép chuyển đổi các text cả trong các Text Box, Header, Footer.

- VKO còn có chức năng tự động đoán mã tiếng Việt đang dùng giúp cho người dùng khỏi lúng túng khi chọn một trong số hơn 40 bảng mã khác nhau. VKO còn cho phép chuyển mã toàn bộ văn bản hay từng vùng cục bộ để chuyển mã cho các văn bản có sử dụng nhiều loại font chữ khác nhau.
- Các văn bản Word, Excel, PowerPoint sau khi chuyển hầu như không cần phải làm bất kỳ thao tác nào để hiệu chỉnh và có thể sẵn sàng nhập số liệu mới theo tiêu chuẩn mới (Unicode). Việc chuyển đổi đã được tự động hoá cao tránh nhiều thao tác mất thời gian (chuyển sang text only hoặc chuyển qua RTF như một số công cụ chuyển mã tiếng Việt khác) và trung thành một cách tối đa văn bản gốc.

III.2. CÀI ĐẶT BỘ XỬ LÝ TIẾNG VIỆT TRONG MS OFFICE

1. Thoát hết các ứng dụng ra khỏi bộ nhớ (kết thúc các chương trình) đặc biệt là các chương trình Word, Excel, PowerPoint cần phải kết thúc trước khi cài đặt các công cụ này.
2. Thực hiện các bước 1,2 giống như mục II.2 tức là nhấp chuột vào nút **Start** và chọn mục **Run** và gõ vào ô chữ : **d:\vkosetup.exe** hoặc nhấp chuột vào nút **Browse** để chọn **vkosetup.exe** nếu chương trình này nằm ở đường dẫn hoặc ở ổ đĩa khác.
3. Chọn vào nút **Đồng ý** để tiếp tục, một cửa sổ như hình dưới đây sẽ xuất hiện, nhấp chuột vào các mục MS Word, MS Excel, MS PowerPoint (đặt các dấu tick vào các ô vuông tương ứng).
4. Bình thường các công cụ này sẽ được cài đặt vào các thư mục tương ứng và thư mục **c:\vietkey**, nếu muốn cài vào thư mục khác thì có thể gõ đường dẫn khác vào ô text trong cửa sổ này.

5. Nhấp chuột vào nút **Install** để cài đặt.

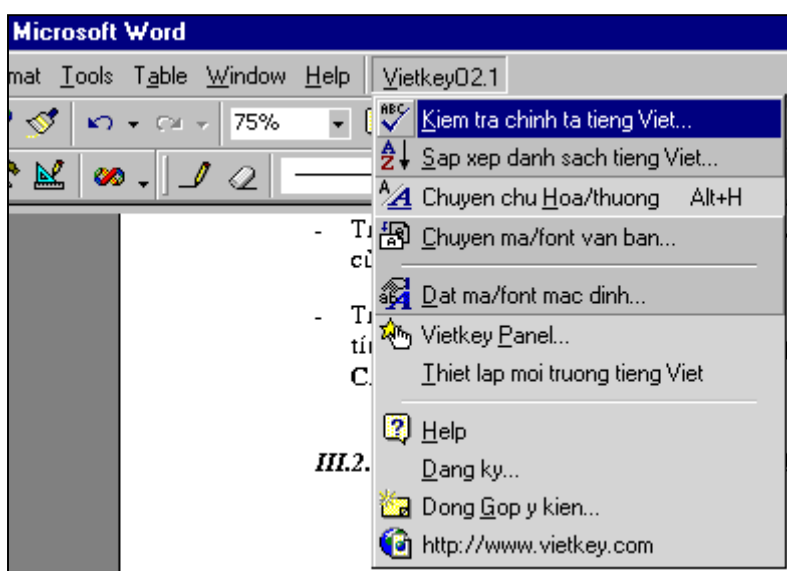


Lưu ý: Cần phải gỡ bỏ các bản Vietkey Office cũ, đặc biệt cần phải xoá các file: **VKLIB.DLL, VKSPELL.DLL** trong thư mục Windows (WinNT) trước khi cài đặt.

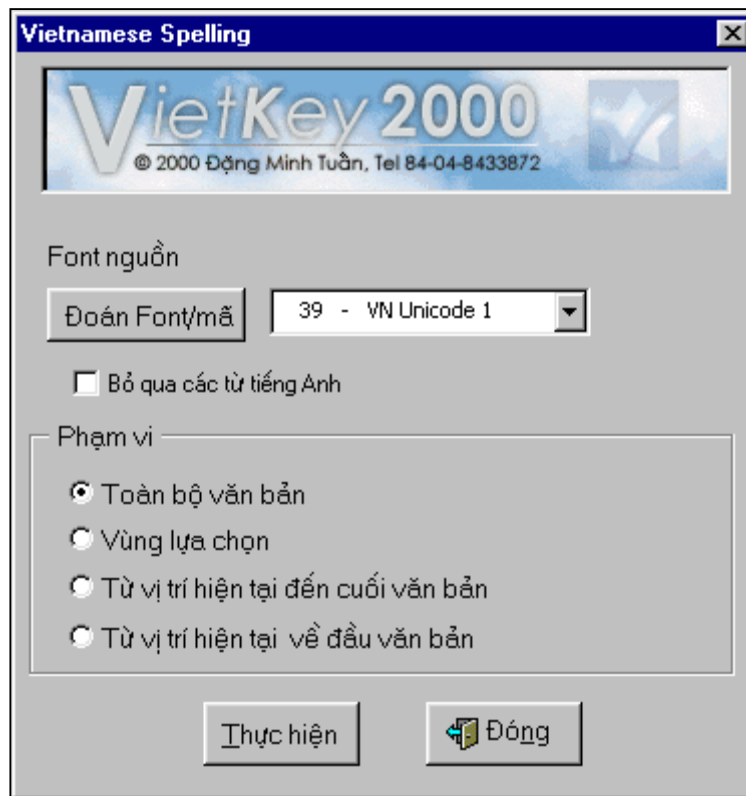
III.3. KIỂM TRA CHÍNH TẢ TIẾNG VIỆT

Sau khi cài đặt xong **Vietkey Office**, các công cụ kiểm tra chính tả tiếng Việt sẽ được tích hợp và Menu **VietkeyOF** sẽ được nối thêm vào Menu của các chương trình MS Word, MS Excel và MS PowerPoint, các Menu của Vietkey Office sẽ thống nhất xuyên suốt cả 3 chương trình này.

Tính năng kiểm tra chính tả tiếng Việt của **Vietkey Office** hoạt động tương tự như tính năng **Spelling** của MS Word. Ngoài việc phát hiện ra các lỗi đánh máy, các từ không có trong tiếng Việt sẽ được công cụ kiểm tra chính tả tiếng Việt phát hiện và có đề nghị các sửa lỗi thông minh.

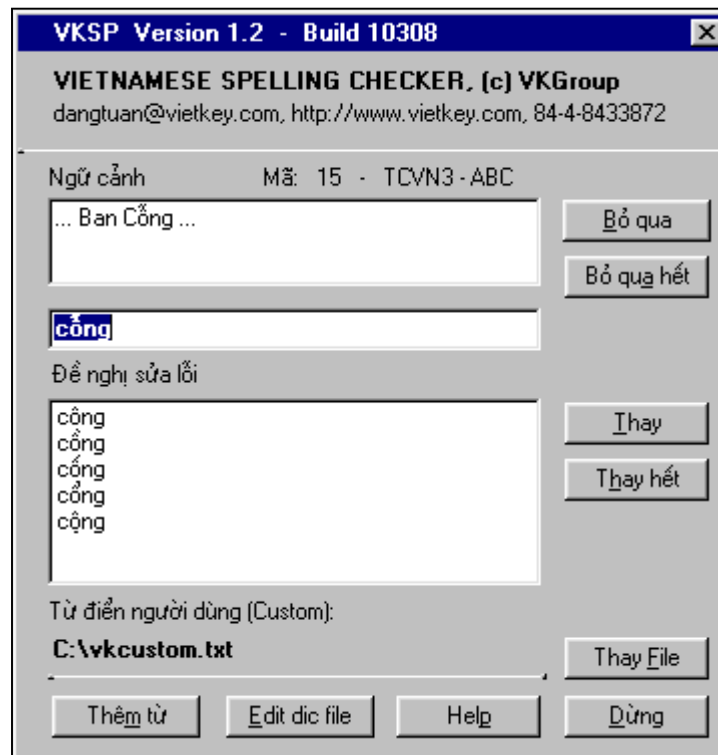


Chọn chức năng **Kiểm tra chính tả tiếng Việt** trong Menu VietkeyO2.1, chương trình hiển thị cửa sổ kiểm tra chính tả như sau:



Cửa sổ này cho phép chọn kiểm tả chính tả tiếng Việt cho toàn bộ văn bản hay trong vùng lựa chọn, hay từ một vị trí bất kỳ đến cuối hoặc đầu văn bản.

Khi phát hiện ra lỗi, công cụ kiểm tra chính tả tiếng Việt sẽ hiển thị cửa sổ sau để cho phép sửa lỗi hay chọn các từ đã được đề nghị sửa lỗi. Công cụ này cũng cho phép thêm vào từ điển người dùng một số từ không có trong tiếng Việt như 1 số từ gõ tắt. Ngoài ra công cụ này cũng cho phép bỏ qua 1 số từ hay thay 1 hay thay hết các từ tương ứng. Xem hình dưới đây:



III.4. CÔNG CỤ SẮP XẾP TIẾNG VIỆT

Thứ tự sắp xếp của bảng chữ cái tiếng Việt:

A ã Â B C D Đ E G H I K L M N O Ô Ớ P Q R S T U U V X Y

Thứ tự sắp xếp của các dấu thanh:

Không dấu, Huyền, Hỏi, Ngã, Sắc, Nặng

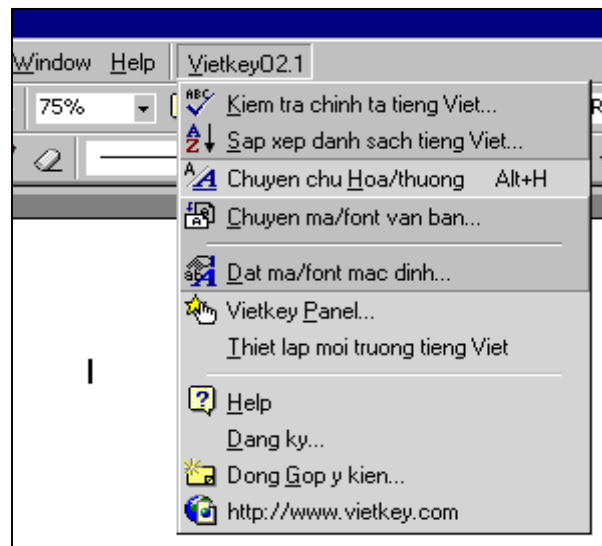
Ví dụ với chữ A: A Ì Á ã Á Ạ

Trong các ứng dụng của MS Word, muốn sắp xếp 1 danh sách tiếng Việt, trước hết cần phải tạo 1 bảng bằng cách vào Menu: **Table/Insert/Table**, điền số cột và số hàng (bỏ qua bước tạo bảng mới nếu đã có sẵn bảng). Nhấp chuột để con trỏ vào cột cần sắp (trong Word chỉ có thể sắp theo cột, trong Excel có thể sắp theo cột hoặc theo hàng).

SỐ TỰ TỰ	HỌ VÀ TÊN	NGHỀ NGHIỆP
1	Đặng Minh Tuấn	Lập trình viên
2	Đỗ Đức Dương	KS xây dựng
3	Phạm Tuấn	Phi công
4	Phạm Hải An	Tiếp thị

Trong MS Excel trước khi sắp xếp cần chọn vùng sắp xếp bằng cách bôi đen vùng cần sắp xếp.

Nhấp chuột vào thực đơn **VietkeyOF** trong cả Word hoặc Excel và chọn **Sắp xếp danh sách tiếng Việt** như hình dưới đây:





Sau khi chọn chức năng sắp xếp trong thực đơn, một cửa sổ **Sort By Vietnamese** sẽ xuất hiện. Trong cửa sổ này, có thể chọn:

- **Font nguồn:** là bộ font đang hiện hành, đang được dùng trong văn bản.
- **Thứ tự sắp xếp:** có thể chọn thứ tự sắp xếp hoặc tăng dần hoặc giảm dần.
- **Xếp theo tên:** công cụ này sẽ tự động cắt từ cuối là tên chính để xếp, nếu trùng sẽ tiếp tục xếp theo họ và đệm của tên. Nếu bỏ thuộc tính này thì công cụ sắp xếp sẽ sắp theo thứ tự từ trái qua phải.
- **Xếp cả Header:** công cụ này sẽ xếp hay bỏ qua hàng đầu tiên (tiêu đề), không xếp .

Riêng trong MS Excel, công cụ sắp xếp có giao diện khác, và cho phép xếp theo cột hoặc xếp theo hàng. Không những thế công cụ sắp xếp trong MS Excel còn cho phép xếp với 3 mức khác nhau.

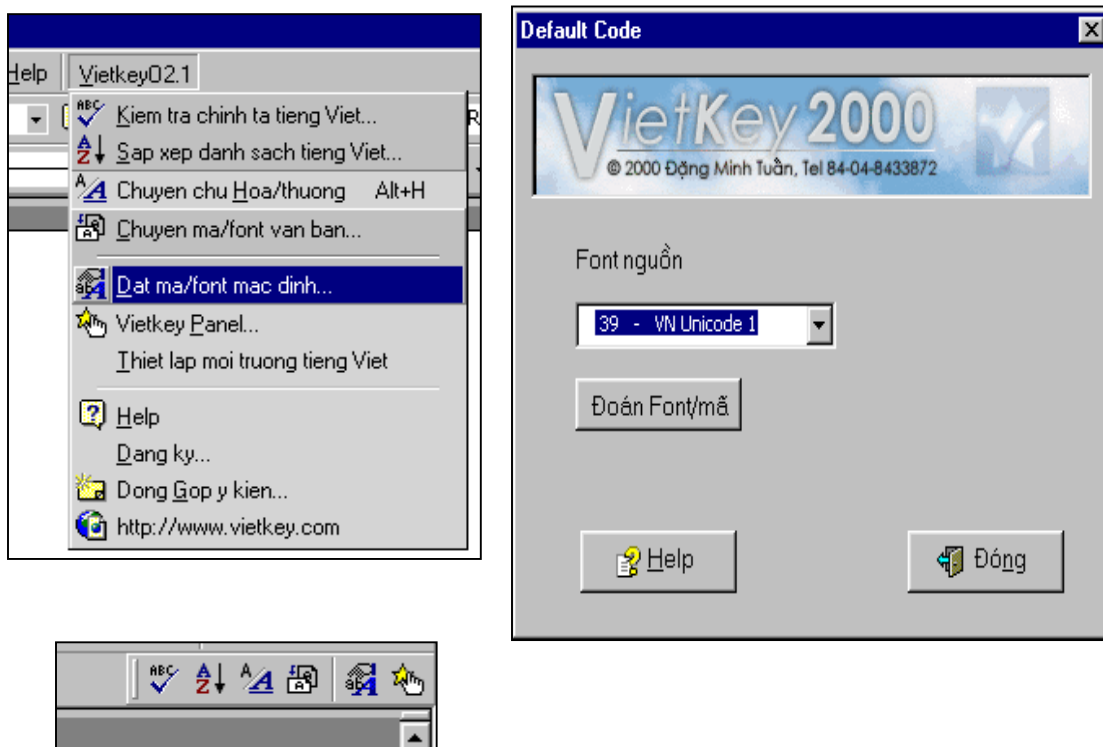



Sau khi sắp xếp bảng trên sẽ được xếp lại, các dữ liệu khác cùng hàng cũng sẽ được di chuyển theo như bảng sau.

SỐ T HỮ TỰ	HỌ VÀ TÊN	NGHỀ NGH ỆP
4	Phạm Hải An	Tiếp thị
2	Đỗ Đức Dương	KS xây dựng
3	Phạm Tuấn	Phi công
1	Đặng Minh Tuấn	Lập trình viên

III.5. CHUYỂN ĐỔI CHỮ HOA/CHỮ THƯỜNG

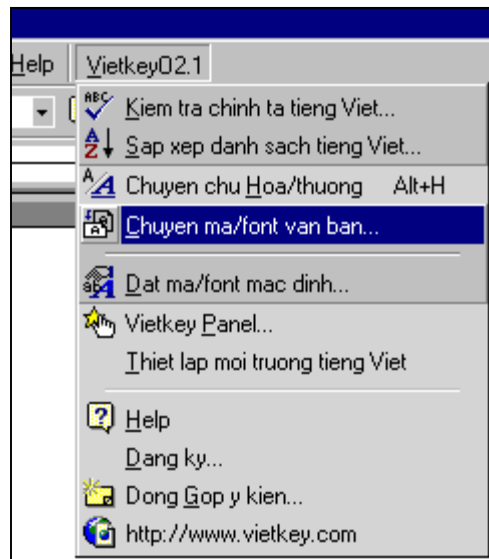
Vì mã của các chữ hoa tiếng Việt và chữ thường tương ứng trong bảng mã Unicode/TCVN 6909 không cách nhau 32 đơn vị nên không thể dùng các chức năng chuyển chữ hoa/chữ thường (**Change Case - Shift+F3**) trong các ứng dụng Word/Excel/PowerPoint. TCVN6909/Unicode cũng không có font riêng cho chữ hoa và font riêng cho chữ thường như font ABC, do đó nếu không có công cụ chuyển chữ hoa chữ thường thì bắt buộc phải xoá chữ thường để gõ lại chữ hoa và ngược lại. Tuy nhiên trong công cụ của **Vietkey Office**, công cụ chuyển chữ hoa và chữ thường đã được tích hợp vào Word/Excel/Powerpoint, để chuyển đổi chữ hoa và chữ thường, trước tiên **cần phải thiết lập bảng mã mặc định** sẽ dùng trong soạn thảo văn bản.

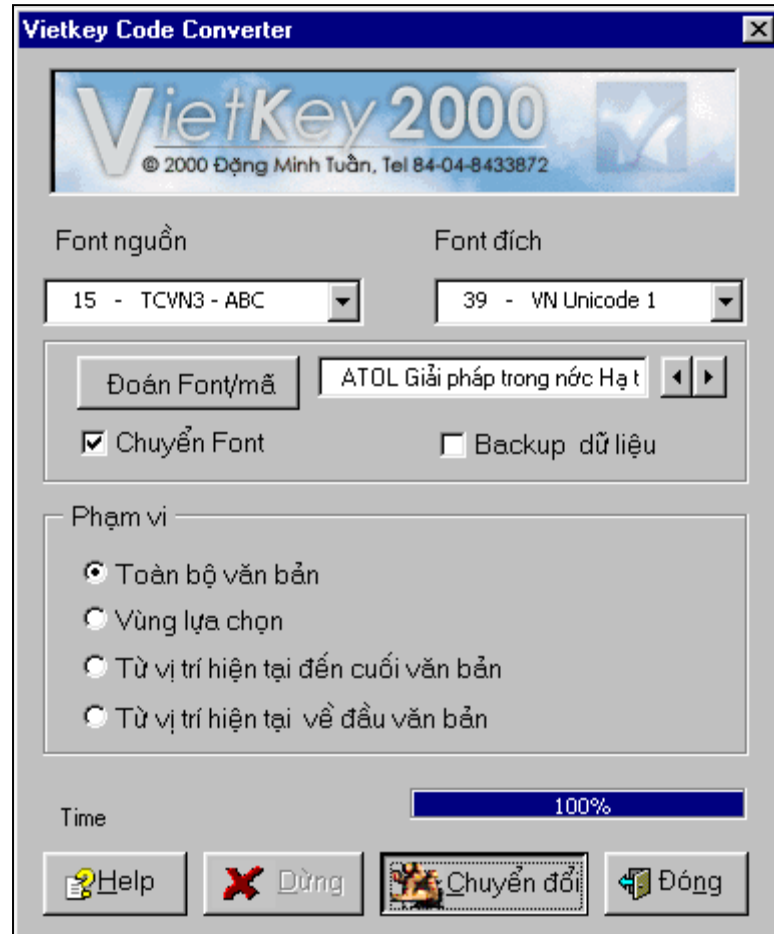


Muốn chuyển một đoạn text (văn bản) từ chữ hoa thành chữ thường hoặc ngược lại, cần phải chọn đoạn text đó (bôi đen bằng chuột hay bàn phím), sau đó nhấp chuột vào biểu tượng  hoặc chọn mục **Chuyen chu hoa/chu thường** trong menu VietkeyOffice, hoặc cũng có thể dùng tổ hợp phím nóng là **Alt-H**, đoạn chữ bôi đen sẽ được chuyển thành chữ thường, nếu làm tiếp lần nữa, đoạn chữ đó sẽ được chuyển thành chữ hoa.

Lưu ý: Trong MS Excel chỉ chuyển đổi chữ hoa, chữ thường trong từng ô (Cell).



III.6. CHUYỂN MÃ VĂN BẢN CHO CÁC VĂN BẢN WORD /EXCEL /POWERPOINT

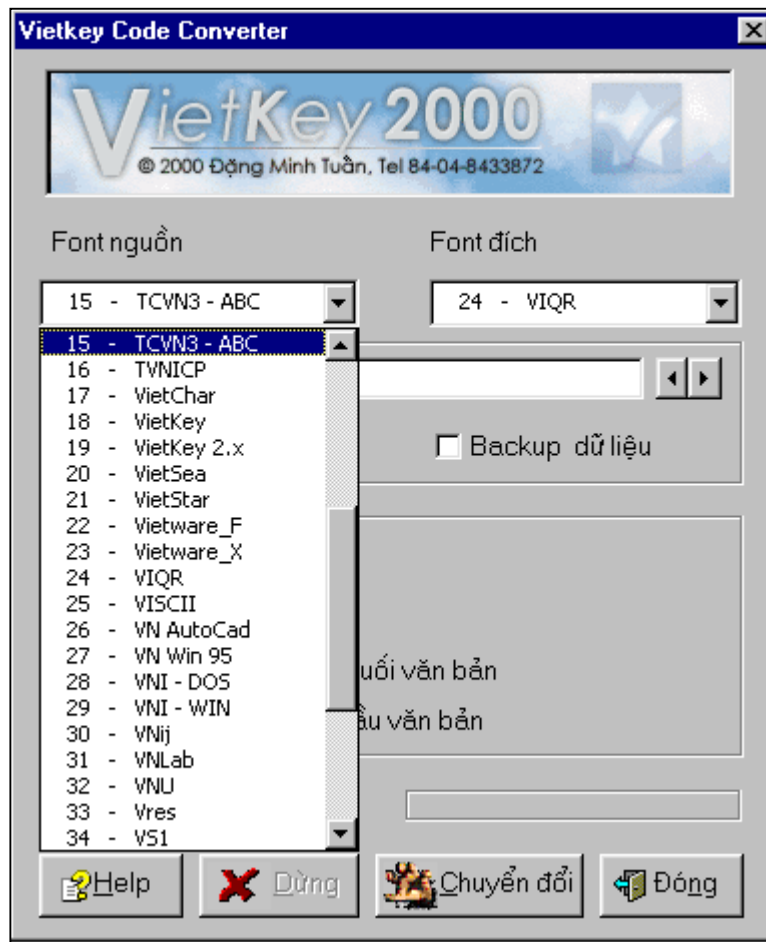




Chuyển mã văn bản là một chức năng quan trọng trong bộ Vietkey Office, bộ chuyển mã này cho phép chuyển đổi các văn bản được soạn thảo trong các ứng dụng Word, Excel, PowerPoint trong các bộ MS Office 97, 2000, XP. Ưu điểm của bộ chuyển mã VKO là nó cho phép chuyển đổi trung thành các văn bản được

soạn thảo từ hơn 40 bộ mã/Font chữ tiếng Việt khác nhau, giữ nguyên được các thuộc tính (đậm, nghiêng, gạch chân, kiểu dáng, kích thước font) của văn bản. Bộ chuyển mã cũng cho phép chuyển đổi các đoạn chữ trong 11 vùng khác nhau, kể cả đồ họa, bảng biểu, header/footer...

- Để thực hiện việc chuyển mã 1 văn bản, trước tiên cần mở văn bản đó bằng chương trình soạn thảo tương ứng (Word, Excel hoặc PowerPoint).
- Nên lưu văn bản gốc trước khi chuyển mã sang 1 thư mục khác để lưu trữ.
- Nhấp chuột vào mục **Chuyen ma/font van ban** trong Menu của VietkeyO2.1 hoặc nút  trong thanh công cụ của VietkeyO2.1, khi đó cửa sổ **Vietkey Code Conveter** xuất hiện.
- **Font nguồn:** là bộ mã hoặc bộ font tương ứng cần chuyển mã (Ví dụ TCVN3-ABC hoặc VNI...), khi bắt đầu khởi động, công cụ chuyển mã sẽ cố gắng đoán nhận bộ font/mã đang được dùng nhằm giúp người sử dụng khỏi lúng túng khi chọn đúng Font nguồn trong số hơn 40 bộ font khác nhau. Trong 1 số trường hợp, không đủ thông tin, công cụ chuyển mã sẽ đoán không đúng bảng mã đang dùng, khi ấy người dùng có thể chọn lại cho đúng trong danh sách Font nguồn bằng cách nhấp chuột vào nút  của danh sách này.



- Lưu ý: trong 1 số trường hợp có thể căn cứ theo tên font để đoán nhận ra bộ Font hoặc bộ mã nguồn:

- Các font bắt đầu bằng **.Vn** (Ví dụ .VnTime, .VnTimeH, .VnArial, .VnArialH) là bộ font **TCVN3-ABC**.
 - Các font bắt đầu bằng **VNI-** (ví dụ VNI-Times...) là bộ font **VNI-WIN**.
 - Các Font bắt đầu bằng **Vn** (Ví dụ VnTimes, VnArial...) thường là bộ font **BK. HCM 1**.
 - Các Font bắt đầu bằng **Vn** nhưng kết thúc là số **2** (Ví dụ VnTimes2, VnArial2...) thường là bộ font **BK. HCM 2**.
 - Các font **Ánh Min, Minh Quân, Heo may, Hoàng Yến, U hoài** là bộ Font **VISCII**.
 - Các font bắt đầu bằng **VN** (Ví dụ VNtimes new roman, VNmonotype corsiva Italic...) thường là bộ font **Vietware_X** hoặc font **VNU**.
 - Các font bắt đầu bằng **VN.** (Ví dụ VN.Time, VN.Arial...) là bộ font **Vietkey 2.x**.
 - Các font bắt đầu bằng **VPS** (Ví dụ VPS An Giang, VPS Nghe An...) là bộ font **VPS**.
- **Font đích:** là bộ mã/font cần chuyển đến, muốn chuyển về font Unicode/TCVN 6909 hãy chọn VN Unicode1.
- **Chuyển Font:** thuộc tính này cho phép chuyển luôn cả tên font thành tên font mới tương ứng trong văn bản đích.
- **Đoán mã/Font:** nút này cho phép tự động đoán mã/font của văn bản hiện hành (trong 1 số trường hợp như thiếu thông tin, chức năng này sẽ đoán không đúng).

- **Phạm vi:** trong mục này, công cụ chuyển mã cho phép chuyển toàn bộ văn bản, hay trong vùng lựa chọn (bôi đen bằng chuột hay bàn phím), hoặc từ 1 vị trí bất kỳ về đầu hoặc cuối văn bản.
- **Lưu ý:** trong 1 số trường hợp nếu có mục lục tự động, hoặc bảng biểu phức tạp, bộ công cụ chuyển mã có thể làm việc không ổn định thì hãy chuyển các đối tượng này sang 1 văn bản mới, sau khi chuyển xong phần chữ (text) rồi mới copy về văn bản đích.

IV. CHUYỂN MÃ CHO CÁC TRANG WEB

IV.1. GIỚI THIỆU CHUNG

IV.2. CÀI ĐẶT CÔNG CỤ CHUYỂN MÃ TIẾNG VIỆT CHO CÁC TRANG WEB

Công cụ chuyển mã cho các trang Web (cho 1 trang Web hay toàn bộ cả các thư mục) là **VKCNVR.EXE** có kích thước khoảng 160KB, cũng như bộ gõ Vietkey rút gọn, không cần phải cài đặt, copy chương trình về một thư mục bất kỳ và nhấp chuột vào biểu tượng của **VKCNVR.EXE**, chương trình sẽ được kích hoạt.

Công cụ này có khả năng chạy trên các hệ điều hành Windows 95/98, Windows NT, Windows 2000, Windows XP.

IV.3. HƯỚNG DẪN SỬ DỤNG CHUYỂN MÃ CÁC TRANG WEB

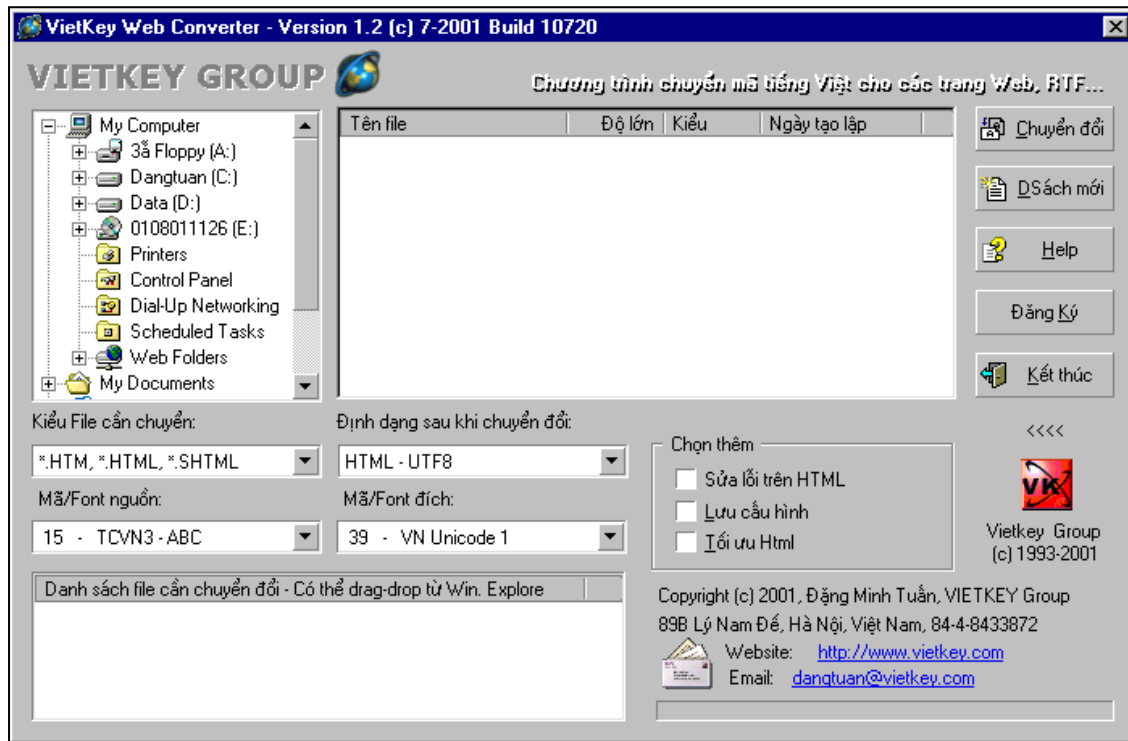
Công cụ chuyển mã cho các trang Web: **Vkconvr.exe**, có khả năng chuyển mã cho từng trang Web hay cùng 1 lúc chuyển nhiều trang Web trong 1 thư mục và có khả năng quét tiếp các thư mục con để chuyển mã.

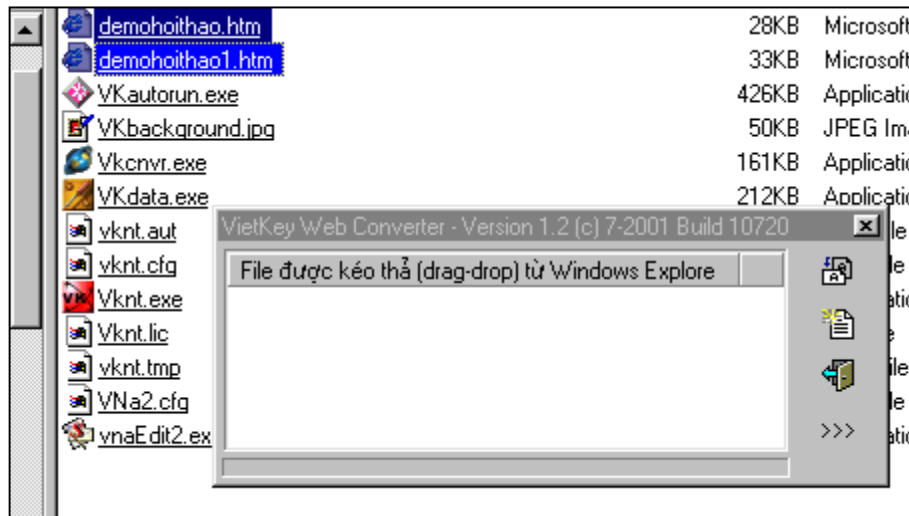
Chương trình có khả năng chuyển các tệp HTML (các trang Web) từ hơn 40 bảng mã tiếng Việt khác nhau sang bộ mã Unicode/TCVN 6909 mà vẫn giữ nguyên các định dạng và các đối tượng có trên trang Web. Đầu ra có thể chuyển thành các khuôn dạng UCS2 (mã có độ dài 16-Bit) hoặc dạng UTF-8 (dạng biến đổi các

ký tự 16-Bit thành chuỗi các ký tự 8-Bit), đồng thời cũng có thể lưu các trang Web thành dạng 7-Bit khác.

Trước khi chuyển mã, nên lưu các trang gốc ra 1 thư mục khác để có thể sử dụng về sau.

Chương trình có khả năng chọn file theo hình thức kéo-thả (drag-drop) bằng chuột khi nhấp chuột vào nút <<<<,





Sau đó có thể mở **Windows Explore** để chọn tệp (File) và dùng chuột kéo vào cửa sổ **Vietkey Web Converter**.

Lưu ý: Trong một số trường hợp nếu các trang Web không hiển thị đầy đủ tên phần mở rộng (htm, html) thì chương trình sẽ không chuyển mã đúng. Trong **Windows Explore->Tools->Folder Options->View** bỏ chức năng **Hide Extension for know file type**.

V. BỘ THƯ VIỆN LẬP TRÌNH XỬ LÝ TIẾNG VIỆT

- Có trên 20 hàm (API) cơ bản cho phép đáp ứng hầu như tất cả các nhu cầu xử lý tiếng Việt.
- Hỗ trợ cho hầu như tất cả các ngôn ngữ lập trình: Visual Basic, Visual C, Foxpro, SQL Server, MS Access, Delphi...
- Không cần cài đặt, đi kèm theo bộ Vietkey Office.

```
long VK_Convert(char *str_source, char *str_dest);
Byte VK_Code2Code(Byte id_source, Byte id_dest);
long VK_VN2unicode(char *str_source, LPWSTR str_dest);
long VK_unicode2VN(LPWSTR str_source, char *str_dest );
Bool VK_isuniCharup (long tch);
long VK_loweruniChar(long tch);
long VK_upperuniChar(long tch);
Byte VK_defaultCode (Byte id_code);
long VK_upperStr (char *str_source);
long VK_lowerStr (char *str_source);
long VK_vn2str7bit(Byte mode, char *str_source,
                  char *str_dest);
long VK_strcmp(char *str_source, char *str_dest);
long VK_convertFont(char *font_source, char *font_dest);
Byte VK_run (Byte cmd);
```

VI. CÁC CÂU HỎI THƯỜNG GẶP

VI.1. LÀM GÌ KHI GẶP KHÓ KHĂN TRONG VIỆC DÙNG VIETKEY?

- Nghiên cứu tài liệu hướng dẫn sử dụng.
- Tìm thêm thông tin trên trang Web: <http://www.vietkey.net>
- Tìm các vấn đề trong phần các câu hỏi thường gặp, trong diễn đàn của <http://www.vietkey.net> .
- Gửi Email cho nhóm phát triển Vietkey Group: unicode@vietkey.net, viengkhang1@hn.vnn.vn, viengkhang2@hcm.vnn.vn hoặc trong trường hợp rất cần thiết có thể gửi cho tác giả dangtuan@hn.vnn.vn.
- Gọi điện thoại cần tư vấn hỗ trợ cho các chi nhánh tại Hà Nội và Tp Hồ Chí Minh, địa chỉ cụ thể có trên hộp sản phẩm hoặc trên Website.

VI.2. CÁC BẢN NÂNG CẤP VÀ CÁC THÔNG TIN MỚI VỀ SẢN PHẨM CÓ Ở ĐÂU?

Trên Website <http://www.vietkey.net> hoặc Website của công ty Vietkhang.

VI.3. TẠI SAO KHI GÕ TIẾNG VIỆT, CHỮ CÓ DẤU ĐÔI KHI BỊ CHÈN 1 KÝ TỰ TRỐNG?

Nguyên nhân: Trong các ứng dụng của Microsoft (Word), có một số tính năng được dùng cho ngữ pháp tiếng Anh nên khi gõ tiếng Việt sẽ có sự xung đột về 2 ngôn ngữ này, do đó cần phải tắt các tính năng tự động hiệu chỉnh theo ngữ pháp tiếng Anh:

Khắc phục: Mở **MS Word**, vào **Tools/Option/Edit** bỏ dấu check ở mục **use smart cut & paste_** để khắc phục lỗi dấu cách.

VI.4. TẠI SAO TRONG MỘT SỐ CHỮ TIẾNG VIỆT CHỮ ‘I’ BIẾN THÀNH CHỮ ‘I’?

Nguyên nhân: Trong tiếng Anh, đại từ ngôi thứ nhất là ‘I’ luôn để ở dạng viết hoa, nếu viết chữ ‘i’, Word tự động hiệu chỉnh theo ngữ pháp tiếng Anh thành ‘I’.

Cách khắc phục: Mở **Word**, vào **Tools/AutoCorrect** và bỏ dấu check trong mục này **Replace text as you type** hoặc trong xóa mục i->I trong danh sách thay thế.

VI.5. LỖI MÁY IN KHI IN CHỮ UNICODE?

Nguyên nhân: Một số máy in có Driver không tương thích với Unicode, phần xử lý in ấn của một số phiên bản của các phần mềm này không tương hợp với Unicode. Một số máy in dưới đây có thể có vấn đề in ấn với văn bản Unicode:

Canon BubbleJet BJ-C600, 4000, 4200, 4500, 4550, v 3.40

Canon Multipass 2500 3.40

Epson Color 500

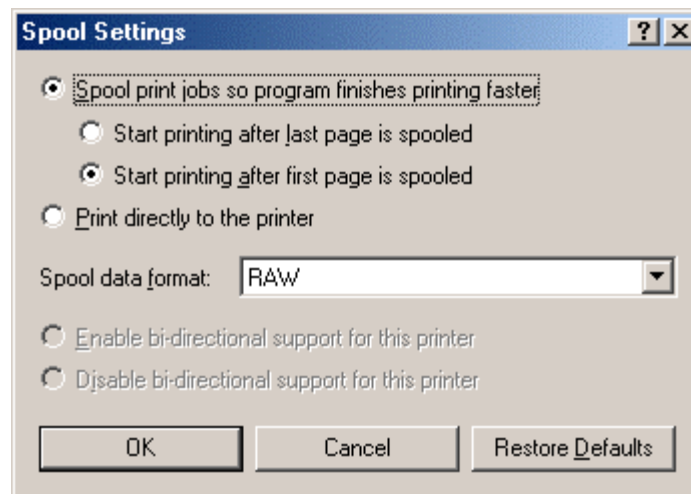
Epson Stylus Color, v 2.x

Epson Stylus Pro / XL v. 2.11BE

HP 1100, 1200
HP Color LaserJet 5 PCL, printer driver version F 1.300
HP DeskJet 1600C, printer driver version 4.20
HP LaserJet 4 PCL, printer driver version 3.78
HP LaserJet 6P, standard printer driver
Okidata 4 laser printer ...

Khắc phục:

- Nâng cấp Printer Driver mới nhất có thể được từ Website của nhà sản xuất ví dụ <http://www.hp.com>.
- Đối với một số chủng loại máy in (HP1100):
 - o vào **Start->Setting->Printer->Máy in của bạn**.
 - o Chọn máy in, vào mục **Properties -> Details -> Spool Setting...**



- o Chọn **Spool Setting... -> Spool data format: RAW**.
- Máy in Postcripts:

- Vào **Properties**, chọn **download soft font** or **download outline font**.
- Máy in HP DeskJet 1600C:
 - Vào **Properties->Detail->Setup->Colorsmart Setup->Advanced** chọn **Use LaserJet III font scaling->OK**.
- Một số máy in khác (HP LaserJet 4 Series PCL...):
 - **Properties -> Fonts ->** chọn "**Always use True Type Fonts**"
 - Hoặc chọn **Print True Type Fonts as graphic**.
 - Có thể dùng phần mềm **FinePrint**, vào <http://www.google.com> gõ từ FinePrint để tìm ra các địa chỉ tải xuống.
- Đối với MS Office 97 có thể khắc phục bằng cách sửa đổi giá trị trong Registry (Word và Excel):
 - Start->Run chạy regedit
 - Trong mục

HKEY_CURRENT_USER\Software\Microsoft\Office\8.0\Excel\Microsoft Excel thêm mục mới và gõ **NoWideTextRendering** trong ô giá trị **4**.

- Trong mục:

HKEY_CURRENT_USER\Software\Microsoft\Office\8.0\Word\Options thêm mục mới **NoWideTextPrinting** và trong **Value Data** nhập giá trị **1**.

VI.6. NẾU MÁY TÍNH KHÔNG CÓ Ổ ĐỌC ĐĨA CD CẦN PHẢI LÀM THẾ NÀO?

- Đến một máy tính khác có ổ đĩa CD và ổ đĩa mềm 1.44, copy từng thư mục Disk1,...Disk7 trong đĩa CD vào từng đĩa mềm. Sau đó tiến hành cài từ ổ đĩa mềm.
- Tải xuống (Download) từ địa chỉ <http://www.vietkey.net> hoặc <http://www.vietkey.com>.

VI.7. TRONG THƯ TÍN ĐIỆN TỬ MUỐN DÙNG TIẾNG VIỆT UNICODE CẦN PHẢI LÀM GÌ?

- Đối với trình soạn thư điện tử phổ thông Outlook Express hoặc MS Outlook, khi soạn một thư mới cần làm các bước sau:
 - o Mở Outlook chọn **New**,
 - o Vào **Format** chọn **HTML**,
 - o Vào **Format->Encoding** chọn **Unicode(UTF-8)**.
- Với thư nhận được hiển thị tiếng Việt Unicode không đúng:
 - o Vào **Format->Encoding** chọn **Unicode(UTF-8)** hoặc **Western European (ISO** hoặc **Windows)**.

VI.8. CÁCH SOẠN THẢO CÁC TRANG WEB UNICODE TIẾNG VIỆT

- Các trình soạn thảo Web phổ thông hỗ trợ Unicode tiếng Việt: FrontPage 2000, FrontPage XP.
- Các trình soạn thảo Web khác như Dreamweaver tuy không hỗ trợ Unicode nhưng vẫn có thể soạn thảo được các trang Web tiếng Việt Unicode. Dùng bộ Vietkey đầy đủ, chọn mã tiếng Việt là **Unicode1 UTF-**

8 (số 41) hoặc **Unicode1 UCS2 (mã số 42)** tuy hiển thị trong trình soạn thảo không đúng tiếng Việt, nhưng các file HTML kết quả vẫn hiển thị đúng tiếng Việt.

- Nên sử dụng định dạng **Unicode UTF-8** vì định dạng này cho phép tương thích tốt hơn với các môi trường 8-Bit, nếu dùng định dạng UCS2, hoặc Windows 1252 thì có thể một số trình duyệt, hoặc một số môi trường Unix sẽ không đọc được.
- Trước khi soạn thảo một trang Web mới, bước đầu tiên nên làm là chuyển định dạng về UTF-8. Trong **FrontPage**, mở trang mới bằng lệnh **New** nhấp chuột vào nút **HTML** ở dưới để xem source của trang Web:

```
<meta http-equiv="Content-Type" content = "text /html;  
charset = windows-1252">
```

- o Trong dòng trên thay **Windows-1252** bằng **UTF-8**.
- o Sau khi chế độ này được thiết lập, khi chuyển văn bản từ MS Word sang FrontPage, đoạn text sẽ được chuyển tự động sang dạng UTF-8.

VI.9. THIẾT LẬP CHẾ ĐỘ HIỂN THỊ UNICODE CHO CÁC TRÌNH DUYỆT WEB?

- Trước tiên muốn hiển thị được tiếng Việt Unicode trong máy tính cần phải cài bộ Font chữ Unicode (có trong bộ đĩa cài đặt).
- Trình duyệt Web phải có các phiên bản Internet Explore 5.0 trở lên hoặc Netscape 4.7 trở lên.
- Thông thường trình duyệt Web sẽ tự động hiển thị đúng tiếng Việt Unicode có bảng mã đúng theo sự chỉ định của trường **Meta charset** trong trang Web. Tuy nhiên trong một số trang Web hoặc không có trường chỉ

định này hoặc chỉ định sai. Do đó trong những trường hợp này cần phải chỉ định lại bằng tay: Mở IE, **View->Encoding** chọn **Unicode(UTF-8)** hoặc **Western European (ISO** hoặc **Windows)**.

VI.10. SOẠN THẢO TIẾNG VIỆT UNICODE TRONG MỘT SỐ PHẦN MỀM KHÁC?

- Emacs 21.2:

- Vietkey chọn bảng mã UTF-8.
- Trong Emacs đặt **keyboard-coding-system** là **UTF-8**.
 - Bấm tổ hợp phím: **C-x RET k utf-8**. Cách này chỉ có tác dụng trong một phiên làm việc của emacs.
 - Chạy lệnh customize của emacs: **M-x customize**. Sau đó chọn **environment - i18n - Mule - Keyboard Coding System**, đặt **Keyboard Coding System** là **UTF-8**. Lựa chọn này sẽ luôn luôn có tác dụng khi khởi động emacs.

- GVim:

- Vietkey chọn bảng mã tiếng Việt là UTF-8.
- Vào file **_vimrc** ở thư mục gốc của **gVim** (thường là **c:\vim**) thêm các dòng lệnh:

```
Setencoding=Unicode
```

```
set guifont=courier_new
```


- Các phần mềm đồ hoạ:
 - o PhotoShop từ phiên bản 7.0 hỗ trợ tiếng Việt Unicode.
 - o CorelDRAW hiện nay chưa hỗ trợ Unicode.

VI.11. UNICODE TRONG CÁC MÔI TRƯỜNG LẬP TRÌNH

- Bộ Visual Studio 97 chưa hỗ trợ Unicode một cách hoàn chỉnh. Trong các trình soạn thảo không cho phép gõ trực tiếp Unicode. Tuy nhiên đối với ngôn ngữ lập trình C, C++, có thể dùng Vietkey để nhập các đoạn Text Unicode bằng cách chọn bảng mã số 43 (Unicode C Style).
- Các Control chuẩn (Menu, Combo box, Listbox...) thường không lưu trữ và hiển thị được các dữ liệu text Unicode. Đặc biệt trong môi trường Windows 95/98 là môi trường phổ biến ở Việt Nam thì việc hỗ trợ lập trình ứng dụng Unicode rất hạn chế.
- Các Controls trong môi trường Windows NT (2000, XP) hỗ trợ Unicode tốt tuy nhiên các ứng dụng Unicode dịch trên môi trường này lại không hoạt động được trên Windows 9x. Để giải quyết vấn đề này có thể dùng bộ Vietkey Controls cho phép hiển thị Unicode trong các Control chuẩn trong 2 môi trường Win-32, kể cả trên hệ điều hành Windows 9x (xem phần mềm ATOL : www.vietkey.net).

VI.12. UNICODE LÀ GÌ?

Unicode là viết tắt của từ Universal Code, tức là bộ mã vạn năng, có thể dùng để mã hoá tất cả các ngôn ngữ chính trên toàn thế giới (trong đó có tiếng Việt), được dùng chủ yếu trong trao đổi hiển thị dữ liệu trong các hệ thống công nghệ thông tin. Unicode còn là cộng đồng của một số công ty hàng đầu trong lĩnh vực công

nghe thông tin như Microsoft, IBM, Sun... được thành lập từ năm 1991 nhằm tạo ra một bộ mã dùng chung cho toàn thế giới. Song song với tổ chức Unicode còn có tổ chức ISO(Tổ chức chuẩn quốc tế chính thống) cũng nghiên cứu một bộ mã đa ngữ dùng trong CNTT là ISO /IEC 10646. Unicode và ISO từ năm 1993 đã thống nhất cùng nhau phát triển và đồng nhất 2 bộ mã ở miền 16-Bit. Unicode là bộ mã 16-Bit (có 65.536 ô mã), còn ISO 10646 là bộ mã 32-Bit (có 4.294.967.296 ô mã). Phiên bản mới nhất của UNICODE là phiên bản 3.2 (trong đó có 143 ký tự tiếng Việt Quốc ngữ và 9299 chữ Nôm), Phiên bản mới nhất của ISO là ISO/IEC 10646-1:2000.

VI.13. UTF-8 LÀ GÌ?

UTF-8 là viết tắt của cụm từ Unicode Transformation Format 8-Bit. Là dạng chuyển đổi Unicode dùng để chuyển ký tự 16-Bit Unicode thành chuỗi ký tự 8-Bit tương đương. Vì đa phần các ứng dụng, các môi trường hiện hành (kể cả Unix-Linux) thường chỉ hỗ trợ các dữ liệu 8-Bit, nên để tăng khả năng tương thích với các hệ thống cũ trong nhiều trường hợp người ta lưu đoạn dữ liệu Unicode thành dạng 8-Bit.

UTF-8 được dùng rất phổ biến trong Internet, hệ điều hành Unix-Linux và các hệ quản trị cơ sở dữ liệu như Oracle...

VI.14. TIẾNG VIỆT ĐƯỢC BỐ TRÍ NHƯ THẾ NÀO TRONG CHUẨN UNICODE?

Từ năm 1993, tiếng Việt đã được bố trí đầy đủ tất cả các con chữ dựng sẵn, tổ hợp và cả các ký tự tổ hợp dấu thanh và tổ hợp dấu nguyên âm. Một con chữ tiếng Việt trong một số trường hợp có thể được biểu diễn ở 3 dạng khác nhau: dựng sẵn, tổ hợp, chính tắc (Canonical). Dạng tổ hợp là dạng của các nguyên âm tổ hợp với các ký tự dấu thanh, dạng chính tắc là dạng tổ hợp của các nguyên âm

(a,e,u,o) với các ký tự dấu nguyên âm (râu, trắng, mũ) và ký tự dấu thanh. Ví dụ chữ ‘ừ’ có thể được biểu diễn như sau:

	Tổ hợp chữ	Chuỗi mã Hex
Dạng sẵn	ế	1EAB
Tổ hợp	ư + `	01B0 + 0300
Chính tắc	u + ' + `	0075 + 031B + 0300

Trong bảng mã Unicode các con chữ Việt không được xếp liên tục liền nhau trong một vùng mà nằm rải rác trong 5 vùng cách nhau: Latin-1, Latin mở rộng A, Latin mở rộng B, Vùng các ký tự tổ hợp, Latin mở rộng thêm. Chữ hoa và chữ thường trong một số trường hợp được xếp cạnh nhau 1 cách tương ứng mà không cách nhau 1 khoảng 32 giá trị như trong các bảng mã ISO 8859 và trong nhiều trường hợp lại cách nhau với các khoảng có giá trị hoàn toàn khác nhau.

VI.15. TẠI SAO TIẾNG VIỆT LẠI KHÔNG ĐƯỢC XẾP VÀO MỘT VÙNG LIÊN TỤC?

Một trong những nguyên tắc của Unicode là tính tường minh và tính không dư thừa, tức là một hình chữ chỉ có một mã duy nhất để bộ mã có tính nhất quán, tính đơn trị và cũng để tiết kiệm không gian mã. Có nhiều hình chữ có thể xuất hiện trong nhiều ngôn ngữ, ví dụ chữ ‘è’ có trong cả tiếng Việt, tiếng Pháp và tiếng Tiệp, như vậy các ngôn ngữ này phải dùng chung một mã, do đó không thể có 3 chữ ‘è’ cho 3 ngôn ngữ khác nhau. Đây chính là nguyên nhân tại sao các chữ Việt lại không được xếp liên tục cùng nhau.

Về vấn đề sắp xếp, nếu các chữ được xếp liên tục thì tưởng chừng việc sắp xếp sẽ được dễ dàng hơn, tuy nhiên trên thực tế không phải như vậy, tiếng Việt không thể chỉ dựa vào giá trị mã của con chữ để sắp xếp mà phải dùng thuật toán tương đối phức tạp để sắp, do đó một khi đã phải dùng thuật toán riêng thì việc các con chữ có ở liên nhau hay không, cũng không ảnh hưởng lắm đến thuật toán sắp xếp. (ví dụ ‘a’ đứng trước ‘à’ nhưng ‘à’ lại đứng trước ‘an’).

VI.16. ƯU ĐIỂM CHÍNH CỦA UNICODE LÀ GÌ?

Ưu điểm chính của Unicode là không gian mã rộng hơn 256 lần so với không gian mã của các bộ mã 8-Bit, do đó nó cho phép mã hoá tất cả các ngôn ngữ chính trên thế giới, cho phép tiếng Việt hội nhập với tất cả các ngôn ngữ khác trên thế giới. Tiếng Việt trong Unicode có đầy đủ tất cả các con chữ, cả hoa và thường, đồng thời khắc phục được triệt để các lỗi tranh chấp với ký tự điều khiển của các bộ mã 8-Bit, ví dụ khắc phục hiện tượng mất chữ ‘ư’, ‘ơ’, ‘ả’...

Các công ty lớn hàng đầu thế giới về CNTT đều cam kết hỗ trợ Unicode, Unicode/ISO1646 hiện và đang là xu thế của các chuẩn và công nghệ mới.

VI.17. NHƯỢC ĐIỂM CHÍNH CỦA UNICODE LÀ GÌ?

Do là bộ mã 16-Bit nên không gian nhớ và trên đĩa cứng của các văn bản được soạn bằng Unicode sẽ lớn hơn các bộ mã 8-Bit.

Nhược điểm chính là các ứng dụng hỗ trợ Unicode chưa nhiều, tuy nhiên trong tương lai không xa các ứng dụng sẽ được chuyển sang hỗ trợ Unicode, cũng giống như việc chuyển từ 7-Bit sang 8-Bit từ những năm 80. Và trên thực tế các ứng dụng phổ thông và các hệ điều hành phổ thông nhất đã đều hỗ trợ Unicode.

VI.18. CÁC FILE UNICODE CÓ KÍCH THƯỚC TỪ 2 ĐẾN 256 LẦN SO VỚI 8-BIT?

Không đúng, chỉ trong trường hợp tất cả các con chữ đều được mã hoá bằng 1 số 16-Bit thì file Unicode sẽ lớn hơn gấp đôi và dung lượng chuyển tải trên đường truyền sẽ lớn hơn gấp đôi. Tuy nhiên trên thực tế các văn bản được soạn bằng Word hay các trang Web Unicode, chỉ các chữ tiếng Việt có dấu mới được lưu ở dạng 16-Bit trong khi các chữ ở vùng 7-Bit thì thường được giữ nguyên do đó không phải bao giờ kích thước cũng tăng lên gấp đôi. Theo các phép thống kê thì thông thường các file Unicode có kích thước tăng lên khoảng 25% so với bộ mã 8-Bit TCVN3 và hầu như không tăng so với bộ mã VNI.

VI.19. TẠI SAO KHÔNG THỂ GÕ ĐƯỢC UNICODE TRONG MỘT SỐ ỨNG DỤNG?

(AutoCad, Foxpro, CorelDRAW, QuarkXpress, VB...)

Một trong những điểm khác biệt lớn nhất giữa bộ font Unicode và các bộ font 8-Bit hiện hành là chỉ có ứng dụng nào hỗ trợ Unicode thì mới hiển thị và nhập được chữ Unicode. Các ứng dụng nêu ở trên do không hỗ trợ Unicode nên không thể gõ được tiếng Việt Unicode trong các ứng dụng này. Tuy nhiên trong tương lai không xa các ứng dụng này sẽ được nâng cấp để có thể hỗ trợ Unicode.

VI.20. MUỐN SỬ DỤNG UNICODE PHẢI NÂNG CẤP MÁY TÍNH LÊN WINDOWS2000?

Không nhất thiết, nếu sử dụng bộ gõ Vietkey2000 Full thì có thể gõ tiếng Việt Unicode với 9 kiểu gõ phổ thông: Telex, Vni, BkTPHCM...với Windows 95 và MS Office97.

Các hệ điều hành mới: Windows 2000/ME/XP đều đã có sẵn các font Unicode tiếng Việt, sau khi cài bộ MS Office 2000, IE5 thì các font Unicode cũng được tự động cài đặt trong hệ thống ngay cả đối với hệ điều hành Windows9x. Đối với các hệ điều hành cũ thì cần phải cài font Unicode thông qua bộ cài đặt Vietkey2000 Full Version.

VII. PHỤ LỤC

VII.1. ĐỊA CHỈ LIÊN HỆ

Mọi thắc mắc, hỏi đáp, đóng góp ý kiến cho các công cụ xử lý tiếng Việt Unicode/TCVN 6909 xin liên hệ với nhóm phát triển phần mềm ứng dụng:

Vietkey Group, Vietkhang JSC.

Hà Nội: 32 Hàng Cót, Hà Nội, +84(4) 8267593,

TP Hồ Chí Minh: 155 B4 Sư Vạn Hạnh (nối dài), Quận 10, +84(8)8640401.

Email:

unicode@vietkey.net
vietkhang1@hn.vnn.vn
vietkhang2@hcm.vnn.vn

Website:

<http://www.vietkey.net>

Hoặc có thể download phần mềm trên các Website

<http://www.itvietnam.gov.vn>
<http://www.moste.gov.vn>

VII.2. TÁC GIẢ

KS. Đặng Minh Tuấn, Trưởng phòng tích hợp hệ thống, Viện Tự động hoá.

- Tác giả bộ gõ ABC, Vietkey, Vietkey Office.
- Biên soạn dự thảo đầu tiên cho tiêu chuẩn TCVN 6909.
- Trưởng nhóm **Vietkey Group** với các sản phẩm khác:
 - **VKDataTools**: Bộ chuyển mã tiếng Việt (Unicode) cho các cơ sở dữ liệu phổ thông: MS Access, Foxpro, SQL Server, Oracle.
 - **Vietkey Linux**: Hệ điều hành thương hiệu Việt Nam, 1 đĩa CD duy nhất bao gồm hệ điều hành và các ứng dụng văn phòng, giao diện và hướng dẫn sử dụng Online bằng tiếng Việt Unicode/TCVN 6909.
 - **Vietkey4Palm, Vietkey4CE**: Bộ gõ tiếng Việt cho các máy PDA cầm tay PALM và PocketPC, hỗ trợ tiếng Việt Unicode và TCVN 5712 (15 Font chữ mỹ thuật).
 - **AnyTimeOnline**, giải pháp mạng môi trường cộng tác cho quản lý điều hành cho doanh nghiệp và các tổ chức nhà nước, hạ tầng cơ sở cho thương mại điện tử và chính phủ điện tử (www.vietkey.net).
 - **T3c** - Hệ quản lý các đối tượng trên nền bản đồ số (GIS), (đô thị, mạng cáp, quản lý dự án...)
 - Các phần mềm quản lý hành chính: **nhân sự, công văn, vật tư, tài chính**...phục vụ các công tác quản lý hành chính.
 - Tư vấn, xây dựng các hệ thống **an toàn bảo mật thông tin**...
 - Hệ thống thương mại điện tử B2B: www.xvmart.net

- Thiết kế các hệ thống phần mềm và phần cứng nhúng dùng trong đo lường điều khiển các quá trình tự động hoá.