

Unicode

Biến-Dạng qua Internet

Hoàng Ngọc Ân

Bài này viết v/v Convert ra Unicode các Attachment đã dùng những Font khác nhau VNI, TCVN, VISCII, VPS ...và VIQR, nhận được từ các E-mail của AH khắp nơi.

Viết tiếng Việt có đầy-đủ dấu rất được phổ-biến kể từ ngày có Universal Code, gọi là Unicode. Tuy nhiên, các Font có thể dùng viết được Unicode rất giới-hạn. Computer ra đời sau Window 98 SE mới có sẵn các Font cần-thiết để hỗ-trợ cho Unicode là: Arial, Times New Roman, Courier New, Tahoma và Verdana.

Khi biết các Font ấy có sẵn trong Computer, muốn viết được Unicode, chúng ta cần phải có các Software và Program để viết. Có rất nhiều Software khác nhau, nhưng vẫn cùng một mục-tiêu là viết ra được Unicode, có thể đọc được rõ-ràng như đọc các sách-báo Việt-Ngữ.

Thường lệ, chúng ta hay viết thư theo Font Unicode nơi Microsoft Word xong, mới Copy qua chỗ viết E-mail và gửi đi theo lối “ Send as Unicode “. Qua Internet thì chữ viết theo Unicode đã bị biến-dạng, không thể nào

đọc được, trừ trường-hợp Computer của người nhận có khả-năng giống như người gửi.

Nếu viết một bài dài, chúng ta thường gửi trong Attachment, chuyển theo E-mail, nó cũng biến-dạng, người nhận phải biết cách để Convert nó ra Unicode mới đọc được .

Riêng tôi, không rành v/v xử-dụng Computer, nhưng rất may-mắn đã được AH Từ Minh Tâm cố-vấn, chỉ cho một Website Address để cho Attachment vào đó thì nó Convert những bài dùng các Font VNI, VPS, VIQR, VISCII hay TCVN ra Unicode. Tôi sẽ nói rõ-ràng hơn cách Convert trong Phần A dưới đây.

Sau đó, tôi có dịp đi thăm các cháu, và được biết thêm một cách khác để Convert ra Font Unicode nhờ bảng Hoán-Chuyển Clipboard của VPSKeys 4.3 và đã có kết-quả rất mau-lẹ, chính-xác với nhiều loại Font khác nữa. Xin xem cách Convert trong Phần B dưới đây.



Tôi chỉ học-hỏi và biết có 2 cách Convert, chữ A và B đó, nên sự hiểu biết rất giới-hạn, AH nào có nhiều kinh-nghiệm hơn xin vui lòng bổ-túc cho các chi-tiết còn thiếu-sót trong phần trình-bày của tôi.

Phần A : Chi-tiết v/v Convert ra Unicode

Xin nói rõ là tôi đã Download được Bảng này từ Website.

<http://www.vn.net/0webapps.vietconvert.php>

(Xin chú ý phần "0webapps" bắt đầu với số zero không phải chữ O.)

Khi nhận được Attachment thì Copy vào Word. Xong vào trong Internet, mở Website trên, chờ vài phút sẽ thấy một Webpage hiện ra, trong đó có chỉ dẫn rõ-ràng các việc phải làm... như "Chọn bảng mã" đó là loại Font mà người viết đã dùng để viết Attachment. Tiếp theo là làm Copy từ Word vô chỗ Webpage phía trên. Nhìn xuống phần dưới, chỗ "Chọn bảng mã" nhớ đổi nó ra Unicode, xong bấm vào chỗ Convert, sẽ thấy hiện ra Bài viết đã được hoán-chuyển. Bấm Select All và làm copy+paste trở lại vào chỗ nào mình muốn Save.



Phần B : Chi-tiết v/v Convert ra Unicode qua Bảng Hoán-chuyển VPSKeys 4.3

Xin nói rõ là tôi đã Download được Bảng này từ Website <http://www.vps.org>.

Cũng như trên, Save Attachment ở đâu thì phải ghi nhớ tên để khi cần dùng có thể Open cho mau lẹ. Mở Word ra; bấm File rồi Open. Tiếp theo, bấm Edit và Select All; bấm Copy.

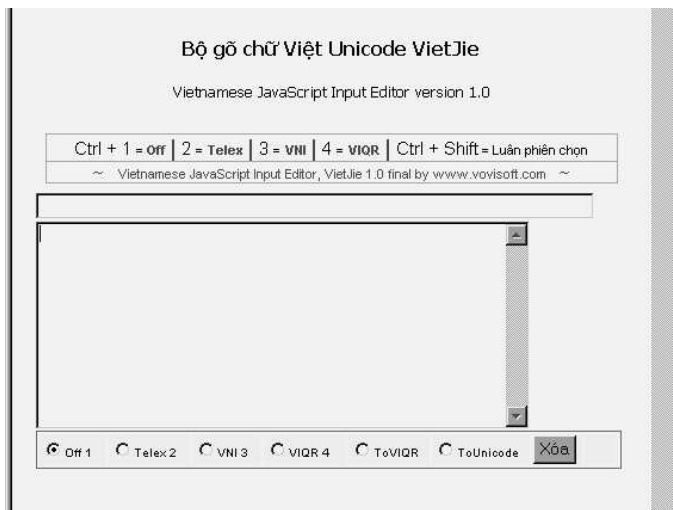
Bấm VPSKeys 4.3 ra. Bấm mục Hoán-chuyển, nhớ để ý chỗ "Từ ... Auto" và "Sang ... Unicode"; xong bấm luôn Hoán-chuyển Clipboard (Nghe có một tiếng...mà vẫn không thấy gì cả). Bấm tiếp chữ OK. Sau đó, trở lại Word và bấm Paste thì sẽ thấy Bài viết đã đổi ra Unicode rồi. Thay đổi Font size và Save as... với tên mới.



Một Đề-Nghị v/v Viết E-mail theo Mã VIQR

Cách đây 2 năm, tôi đã Download được một Software "Vietjie" của Vovisoft, dùng để gõ ra chữ Việt Unicode. Muốn có Software này, xin ghé qua Website:

<http://www.vovisoft.com/unicode/utilities/vietjie/default.asp>



Trong Website đó, có một nhu liệu mang tên là “ Vietnamese JavaScript Input Editor “ dùng để gõ một trong 3 kiểu “Telex, VNI hay VIQR” thì có ngay chữ theo Mã Unicode.

Telex dùng một chữ kèm sau Nguyên-âm để thay thế cho một dấu.

VNI dùng một con số kèm sau Nguyên-âm để thay thế cho một dấu.

VIQR dùng một dấu có sẵn nơi Keyboard để đánh dấu...nối đuôi...ngay sau Nguyên-âm.

Xin xem 4 ví-dụ dưới đây, tôi viết 1 câu (giống nhau) theo 4 Mã khác nhau để thấy sự sai biệt:

Đọc và viết thử bốn chữ “Việt-Nam thân thương” bằng Mã Unicode.

DDojc vaf vieest thuwr boosn chuwx “Vieejt-Nam thaan thuwowng” bawfng Max Telex.

D9o5c va2 vie61t thu73 bo61n chu74 «Vie65t-Nam tha6n thu7o7ng» ba82ng Ma4 VNI.

DDo.c va‘ vie^t thu+? bo^'n chu+~ «Vie^.t-Nam tha^'n thu+o+ng» ba(‘ng Ma~ VIQR.

AH nào quen dùng kiểu nào thì đánh máy theo kiểu ấy vào trong Webpage Input Editor sẽ thấy có chữ Việt theo Mã Unicode. Riêng tôi, nhận thấy Mã VIQR có dạng dễ đọc, gần giống Unicode hơn, nên thường dùng nó để viết E-mail .

Trong Input Editor, còn có một phần dùng để hoán-chuyển VIQR qua Unicode và ngược lại.

Như vậy, nếu quý vị muốn gửi e-mail thì nên gửi theo VIQR, và người nhận muốn đọc cho được thoải –mái theo Unicode thì đưa vào đó để Hoán-chuyển.

Tôi nghe nói VIQR chỉ cần dùng 7 Bits và hệ-thống E-mail Internet các máy chịu chuyển thường chỉ có 7 Bits, do đó gửi E-mail dùng Mã VIQR rất tiện lợi vì nguyên văn của e-mail không bị biến-dạng qua Internet. Computer nào, đầu cũ hay mới cũng nhận được. Nếu cần Reply hay Forward, nguyên văn vẫn không bị biến-dạng.

Những Font khác đều là 8 Bits, Unicode 16 Bits, khi gửi E-mail qua Internet (7 Bits) thì phải bị biến-dạng. Do đó, tôi đề-nghị các AH nên thông-đạt với nhau bằng cách dùng Mã VIQR để viết và gửi E-mail qua Internet.

Phần Bổ-Túc v/v Đọc và Viết theo Mã Unicode

Sau khi gửi bài trên này đến BPT, tôi được y/c bổ-túc thêm về những kinh-nghiệm đã học-hỏi “ Cách Đọc và Viết theo Mã Unicode “ của tôi như thế nào hầu có thể giúp cho các AH khác, nếu muốn, cũng có thể làm được việc này như mọi người đã biết .

Thật tình, từ ngày bắt đầu đến tuổi “ Cô Lai Hy“, tôi rất lười biếng, không thích nghe ai nói là nên học thêm...cái gì cả... Nhưng muốn viết thư E-mail, gửi đi khắp nơi, trong và ngoài nước...miễn phí, tôi phải học xử-dụng Computer . Rồi từ đó, muốn viết được tiếng Việt cho có đủ dấu, như viết theo Mã Unicode, tôi phải vào các Website để tìm các Bài Hướng-Dẫn về cách viết.

Nghe nói có nhiều Software giúp cho chúng ta viết ra Unicode đã có từ nhiều năm qua. Riêng tôi, chậm tiến nhất, mới học và biết được chỉ có vài năm mà thôi . Để bắt đầu,tôi Download từ Website của Vovisoft ra Software “Vietjie”

Address là:

<http://www.vovisoft.com/unicode/utilities/vietjie/vietjie.zip>

Trong đó, có một Webpage mang tên là : Bảng Gõ chữ Việt Unicode Vietjie

Tôi đọc kỹ và theo sự hướng-dẫn của họ ... tôi viết ra ngay Mã Unicode .

Gần đây, vào Website aihuucongchanh và Website <http://www.vps.org> , tôi được biết thêm một cách viết ra Unicode “ Tự-Do “ hơn, vì không bị gò-bó trong cái “ Bảng Gõ chữ Việt...” ấy . Chỉ cần

Download VPSKeys 4.3 từ chỗ Website nói trên, xong sẽ thấy một icon màu xanh dương hiện ra trên màn ảnh của Computer. Theo sự hướng-dẫn trong bảng Nhu-Liệu, chúng ta vào Mục kỹ-thuật để chọn “cách viết thêm dấu theo ý muốn của mình“. Xong thì bấm OK.

Kể từ đó, nếu dùng Keyboard đánh máy gõ ra bất cứ chữ gì cũng là viết theo Mã Unicode . Viết bài , viết E-mail, viết chỗ nào cũng là Unicode

Cũng nên biết là VPSKeys có rất nhiều loại, khác nhau theo Địa-Phương xử-dụng và tùy thuộc vào Năm, Tháng đã ra đời. Hiện nay, VPSKeys 4.3 là loại mới nhất có thể làm được nhiều việc như có thể dùng cho nhiều ngôn -ngữ khác nhau, riêng tiếng Việt, còn có thêm các chức năng sửa các lỗi chính tả, tự điền để biết cách sửa dấu hỏi ngã, và viết tốc-ký của vài nhóm chữ thường dùng.

Sau khi Download VPSKeys 4.3 xong và muốn dùng nó thì nên nhớ phải cho ngưng các Software đã có thể viết ra chữ Việt một cách khác, vì mỗi Software phải dùng một loại Font riêng.

Trong Mục Kỹ-Thuật của VPSKeys có phần Sửa Nút Dấu, nếu chọn theo kiểu nào thì phải nhớ để áp-dụng mà viết cho đúng. Ngoài ra, cũng nên lưu ý là khi viết chữ Việt thì chỉ đánh dấu tiếp theo sau Nguyên-Âm mà thôi. (Các chữ a, e, i, o, u, y).

Muốn biết cho đầy-đủ chi-tiết rõ-ràng hơn, quý AH nên tìm đọc Mục “ Thắc-Mắc về Nhu-Liệu“ (F.A.Q. hay là Frequently Asked Questions) trong Website <http://www.vps.org> . ☐