

Accademia “Giuseppe Aliprandi – Flaviano Rodriguez”

Sessione culturale 2017: Le Radici e la Memoria fondano il presente e aprono il futuro

Abstract Relazione dott. Alessio Popoli

“Paragone tra le diverse lingue delle Gare Internet di Intersteno”

L'associazione Intersteno, ogni anno tra aprile e maggio, organizza una gara di produzione del testo al PC, ovvero l'Intersteno Internet Contest, i cui partecipanti sono chiamati a copiare un testo alla massima velocità per una durata di 10 minuti. Sono al momento disponibili 17 lingue. I partecipanti possono competere in ognuna di queste lingue; i punteggi ottenuti in ognuna di queste lingue sono sommati per ottenere il punteggio finale. Vengono poi stilate, in tempo reale, tre classifiche - lingua madre, multilingua (esclusa la lingua madre) e generale.

Chiaramente, la conoscenza delle lingue in cui si compete aiuta molto il concorrente, che però va incontro a un'altra difficoltà: quasi tutte queste lingue - fa eccezione soltanto l'inglese - usano almeno un segno diacritico (quelli che noi spesso chiamiamo “accenti”, generalizzando), alcuni dei quali sono peculiari di una data lingua, come ad esempio il doppio accento acuto " dell'ungherese o il kroužek (anello) ° del ceco. Inoltre, quindici lingue utilizzano l'alfabeto latino (più o meno modificato), mentre due, il russo e il giapponese, fanno affidamento a sistemi di scrittura diversi; il russo è infatti scritto con l'alfabeto cirillico, mentre il giapponese usa un sistema composito di cui si parlerà a breve.

Per competere nelle lingue “diacritic-heavy” (che usano molte diacritiche o che le usano molto spesso), come per esempio il ceco e lo slovacco, è indispensabile utilizzare una configurazione (layout) della tastiera ottimale, che permetta di modificare rapidamente i caratteri che lo richiedono, per esempio usando le deadkey, ovvero dei tasti che, una volta premuti, non producono un output immediatamente visibile, ma modificano il prossimo carattere che sarà premuto. Un'alternativa è usare alcune scorciatoie, per esempio alt-gr + c per č. Ovviamente sono a disposizione i tasti utilizzati in italiano per le vocali accentate, ma questi non bastano (allegherò una tavola che mostra, lingua per lingua, tutte le diacritiche utilizzate); inoltre, sebbene in alcune lingue non siano utilizzate certe lettere, per esempio la B in finlandese o la V in polacco, non è prudente rimuoverle dalla configurazione della tastiera, perché i nomi stranieri potrebbero contenerle.

In giapponese vengono utilizzati due sillabari, il katakana e l'hiragana, e una serie di logogrammi - il termine "ideogramma" è errato, perché si riferisce ad un carattere ottenuto da un disegno del concetto a cui si riferisce, cosa che non avviene in giapponese - chiamati kanji. Viene usato, marginalmente, un quarto sistema, detto rōmaji, rappresentato dai caratteri latini. In generale, il katakana viene usato per trascrivere le parole non giapponesi, come per esempio i nomi di persone straniere; i kanji sono responsabili di trasmettere il significato primario della parte del discorso che rappresentano, mentre l'hiragana viene usato per le desinenze, le particelle che sostituiscono i casi grammaticali e per qualche altra funzione comunque legata agli aspetti grammaticali più che lessicali.

Per scrivere il giapponese alla tastiera, viene solitamente utilizzato un IME (metodo di input) che funziona grossomodo così: si scrive la traslitterazione fonetica dei caratteri che si vogliono rappresentare, per esempio “watashi” per 私, poi si batte lo spazio per alternare tra le possibili ri-traslitterazioni in giapponese (“watashi” è anche la traslitterazione di “渡し”, che ha però un altro significato). Si va però incontro a diversi problemi: in primo luogo, chi non è madrelingua giapponese non può certo conoscere tutti i caratteri che incontra, siccome gli stessi giapponesi continuano a imparare nuovi kanji fino all’università. A questo problema si può ovviare con dei software che producano una traslitterazione del testo di gara. In secondo luogo, bisogna ridurre le pressioni del tasto spazio al minimo per aumentare la velocità, ma spesso è difficile determinare il punto giusto in cui farlo per evitare errori; soltanto la pratica può aiutare a migliorare. Infine, esistono diversi modi di traslitterare un carattere, tutti ugualmente validi con l’IME Microsoft, ma quelli che richiedono meno caratteri sono solitamente quelli meno intuitivi (come per esempio “ti” al posto di “chi” per rappresentare il suono che in italiano scriveremmo come “ci”).

Esiste però un ultimo problema da non trascurare: il punteggio finale non è determinabile semplicemente contando i caratteri presenti nel testo, non solo per questa molteplicità di modi per scrivere un carattere, ma anche perché il giapponese non usa un alfabeto, bensì due sillabari e una serie di logogrammi, come spiegato in precedenza. Gian Paolo Trivulzio, ex presidente di Intersteno da poco defunto, e Takahiro Sumino, madrelingua giapponese, hanno a lungo lavorato per decidere come valutare i caratteri della lingua nipponica e sono giunti alla decisione di utilizzare un tasso di conversione, pari a <informarsi>, per cui moltiplicare il numero totale di caratteri effettivamente presenti nel testo per ottenere il punteggio lordo. Gian Paolo Trivulzio ha fatto inoltre da tramite tra lo stesso Takahiro Sumino, il programmatore di TAKI Marco Olivo e Alessio Popoli, il sottoscritto, chiamato a testare il software effettuando esercitazioni in lingua giapponese e provando diverse tipologie di errore. La soluzione del tasso di conversione ha superato i test, ed è dunque l’alternativa che viene utilizzata attualmente.