

Individuazione di parole con iniziali maiuscole

(i tutorial di Alessandro de Simone)

Copyright Alessandro de Simone 2009 (www.alessandrodesimone.net) - È vietato trascrivere, copiare, stampare, tradurre, riprodurre o divulgare il presente documento, anche parzialmente, senza l'autorizzazione scritta dell'autore. I siti Internet, le case editrici e le pubblicazioni di settore che intendano utilizzare questo documento possono contattare l'autore (info@alessandrodesimone.net) per gli accordi del caso.

Copyright Alessandro de Simone 2009 (www.alessandrodesimone.net) – No transcribing, no copyng, no reproducing, no translating, no printing, no publishing this document – even if partially – without author's written authorization. Websites and publishing house who wish to employ this document must write the author (info@alessandrodesimone.net).

Premessa

Il presente *Tutorial* nasce dalla richiesta di un lettore (**Paolo C.** che qui ringrazio) che voleva utilizzare il mio vecchio programma "Estrazione vocaboli", utile per individuare in un testo tutte le parole che iniziano con una lettera maiuscola. Il programma in questione – per come l'avevo strutturato – consentiva l'estrazione di tutte le parole, ma purtroppo convertiva tutti i caratteri in minuscolo.

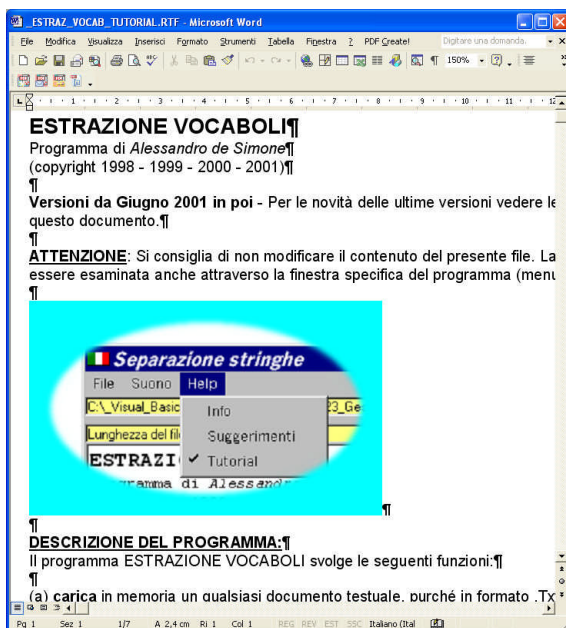
La richiesta del lettore è stata comunque uno spunto per suggerire una procedura, a dire il vero lunga e relativamente complessa, per risolvere il problema senza usare nessun altro programma al di fuori di un *Word Processor*, un *Text Editor* e uno *Spreadsheet*.

In questo tutorial ho usato Microsoft Word, Excel e Notepad (*Blocco note*), ma possono essere usati altri programmi universalmente noti e diffusi.

L'opportunità è di stimolo per creare macro istruzioni che possano velocizzare la procedura.

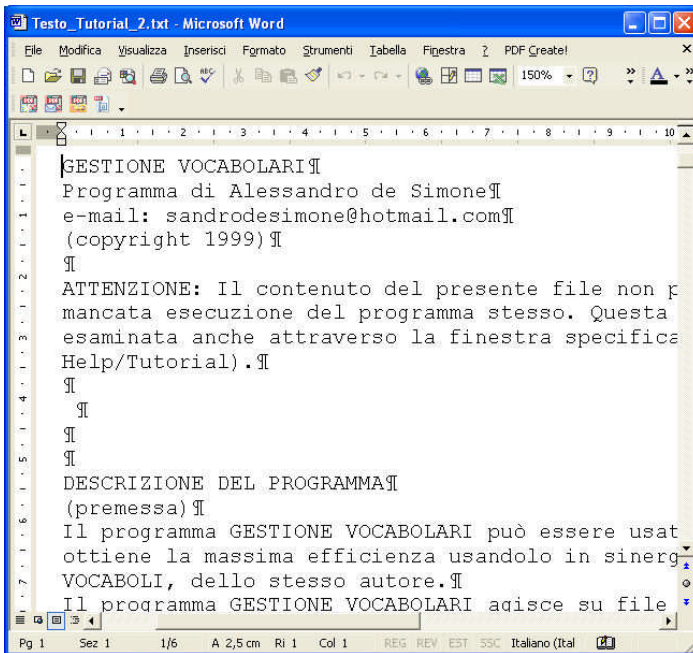
1 – Il testo

Usando il word processor (**Word**, cui mi riferirò d'ora in poi) aprite il testo che vi interessa, di qualunque formato esso sia, magari addirittura contenente figure:



Selezionate l'intero testo (menu *Modifica / Seleziona tutto*) e copiatelo nella clipboard (menu *Modifica / Copia*). Aprite il text editor (d'ora in poi mi riferirò a *Blocco Note*) e incollate (*Modifica / Incolla*) il testo. Noterete che tutte le immagini scompaiono, come pure le formattazioni. Sempre restando in *Blocco Note*, selezionate l'intero testo e copiatelo. Tornate

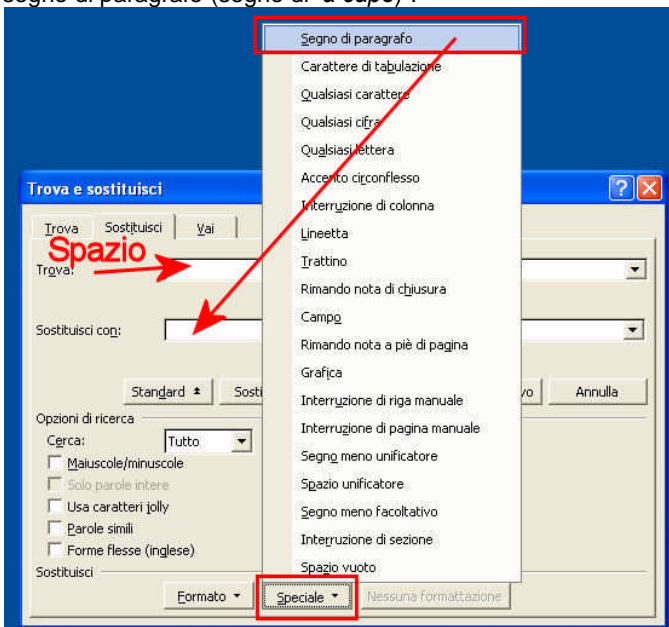
in *Word*, create un documento nuovo e incollate il testo (vedi figura). In questo modo avete lo stesso testo di prima, ma "depurato" dalle figure e dalle formattazioni.



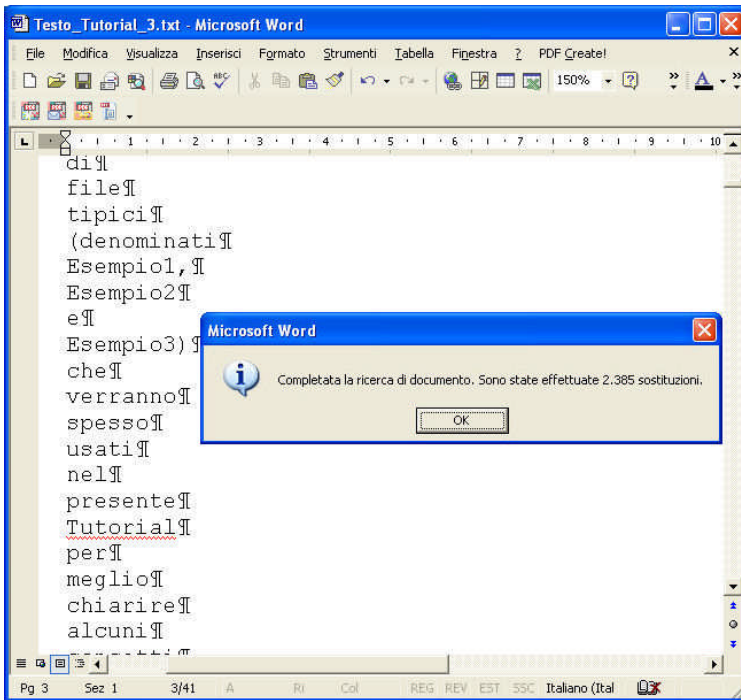
NB: l'indirizzo di posta elettronica che compare nelle figure non è più attivo.

2- Prima gestione del testo

Dobbiamo ora posizionare ogni singola parola del testo su un singolo rigo (vedremo che ci sarà utile quando dovremo elaborarlo con *Excel*). Per fare questo ricorreremo alla funzione del menu *Modifica / Sostituisci*. Nel campo *Trova* della finestra di dialogo che compare digiteremo lo spazio (premendo la barra spaziatrice) mentre nel campo *Sostituisci con* inseriremo il segno di paragrafo (vedi figura). Si tratta della coppia di caratteri **^p** che *Word* interpreta, appunto, come segno di paragrafo (segno di "a capo").

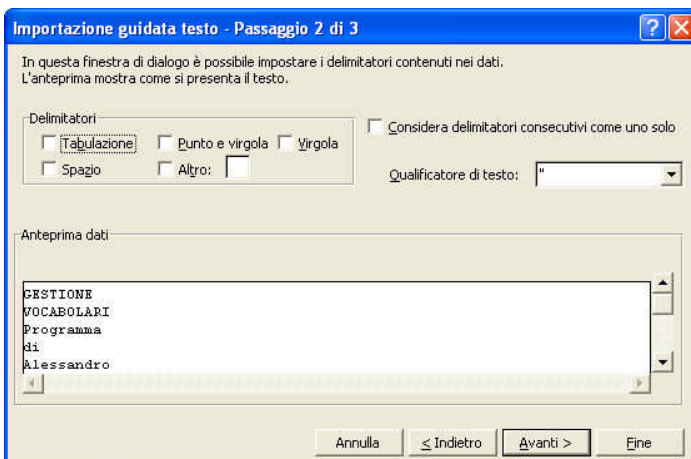
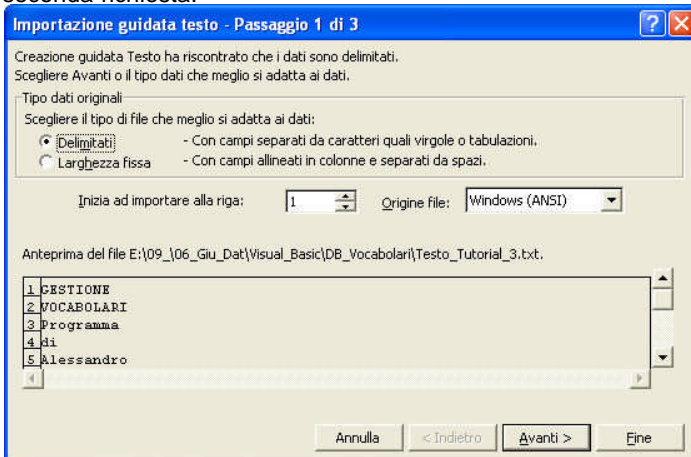


In seguito a tale comando, si verificheranno, verosimilmente, **migliaia** di sostituzioni, ma alla fine si otterranno singole parole su altrettanti rigi. Salvate il testo così ottenuto in formato puro ASCII, cioè con il suffisso TXT. Ignorate il probabile messaggio di *Word* che mette in guardia contro la possibile perdita di eventuali formattazioni: è proprio ciò che vogliamo ottenere!

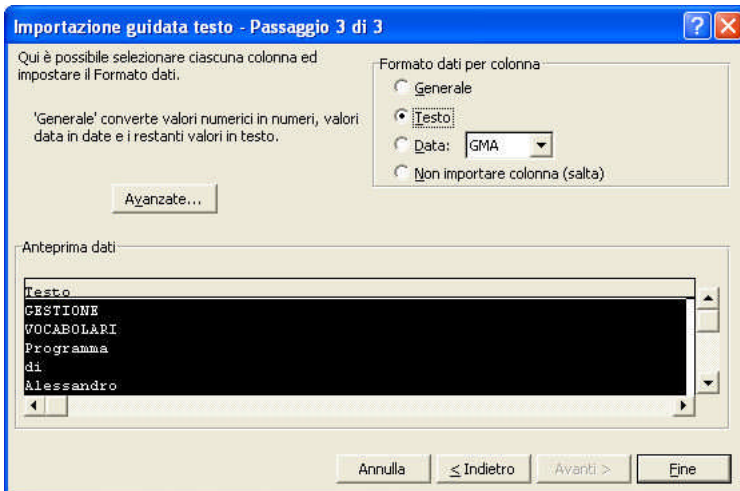


3- Importazione del testo in Excel

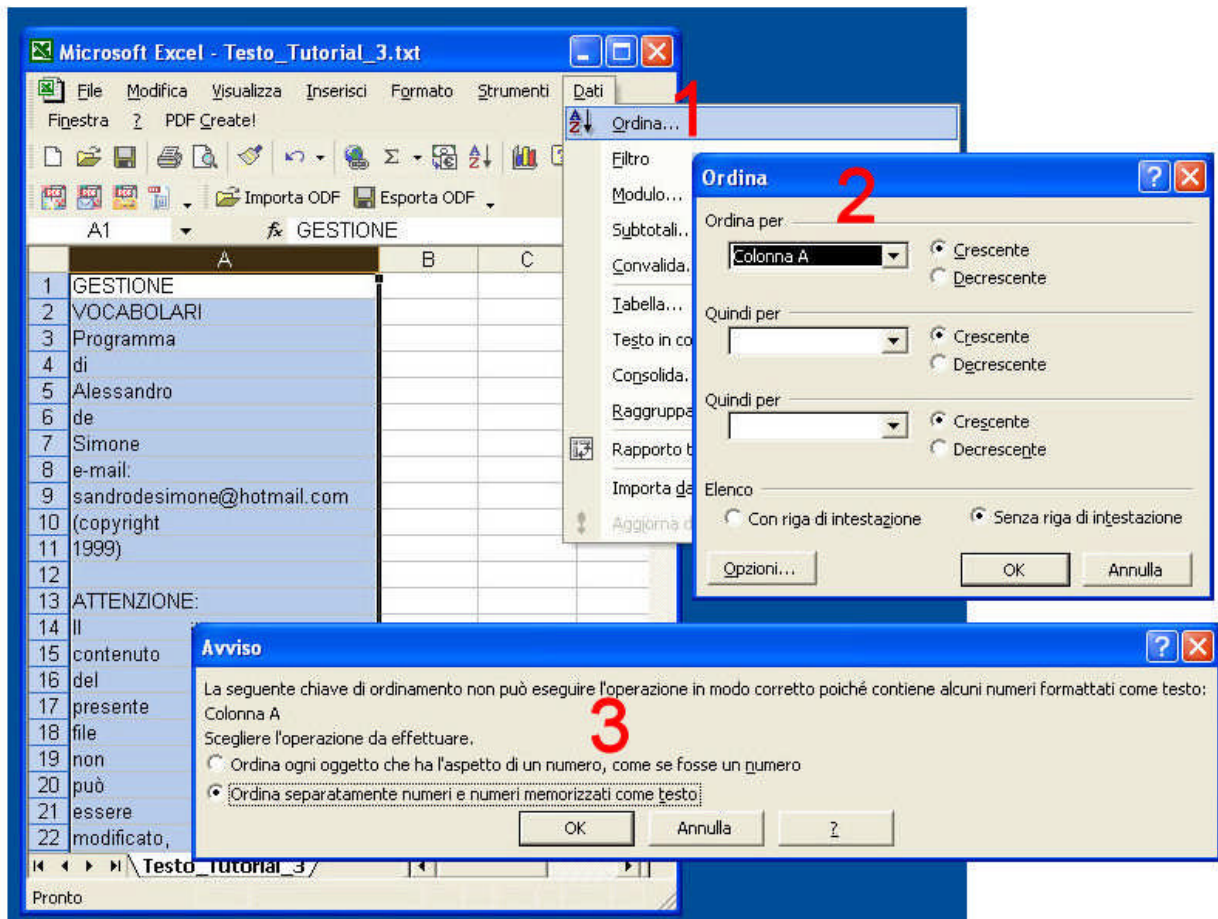
A questo aprirete lo spreadsheet (**Excel**, d'ora in poi) e importate il file di testo appena salvato con *Word*. Dal momento che *Excel* legge un formato "insolito", chiederà all'utente alcune informazioni. Accettate il *default* della prima e della seconda richiesta:



Per quanto riguarda la terza richiesta, selezionate *Testo* come *Formato dati per colonna*.



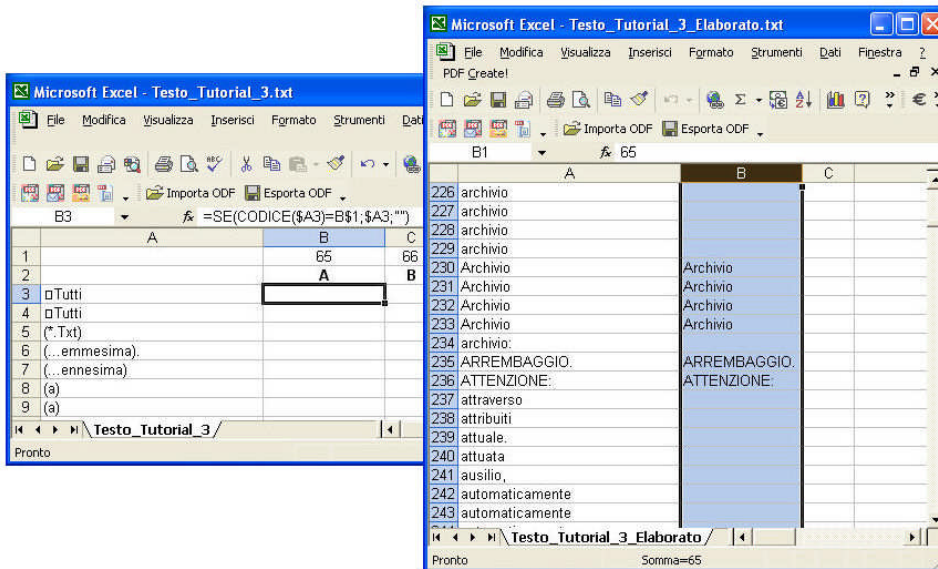
In questo modo abbiamo trasferito ogni parola (o assimilabile tale) in ciascuna cella di una sola colonna di *Excel*. A questo punto selezionate l'intera colonna **A** (in cui, cioè, sono presenti tutte le parole) e cliccate sul menu *Dati / Ordina*, seguendo le indicazioni di figura:



Nella colonna **A**, quindi, ritroviamo tutte le parole ordinate alfabeticamente. Ora bisogna individuare ed evidenziare tutte le parole che iniziano con una lettera maiuscola.

4- Estrazione parole con iniziali maiuscole

Alla destra della colonna **A** occorre ora creare tante colonne quante sono le lettere dell'alfabeto. In ciascuna di queste colonne verranno posizionate, usando specifiche funzioni di *Excel*, le parole di competenza. Non è questa la sede per descrivere il modo in cui operano le funzioni di *Excel*. Sul mio sito, tuttavia, ho messo a disposizione lo spreadsheet di queste figure (denominato **Estrazione_vocaboli_iniziali_maiuscole.XLS**), in modo che il lettore sia facilitato nella sua interpretazione ed eventualmente nella successiva manipolazione.



5- Trasferimento parole su Blocco Note

Occorre ora selezionare, una alla volta, le colonne di *Excel* relative alle singole iniziali maiuscole. Per fare questo, cliccate sull'intestazione della colonna in questione, per assicurarsi che *Excel* la selezioni per intero: le celle vuote presenti dopo l'ultima non verranno infatti prese in considerazione nell'operazione di *Copia* (menu *Modifica / Copia*). Dopo aver copiato il contenuto dell'intera colonna, aprite *Blocco Note* ed effettuate l'operazione inversa (menu *Modifica / Incolla*).

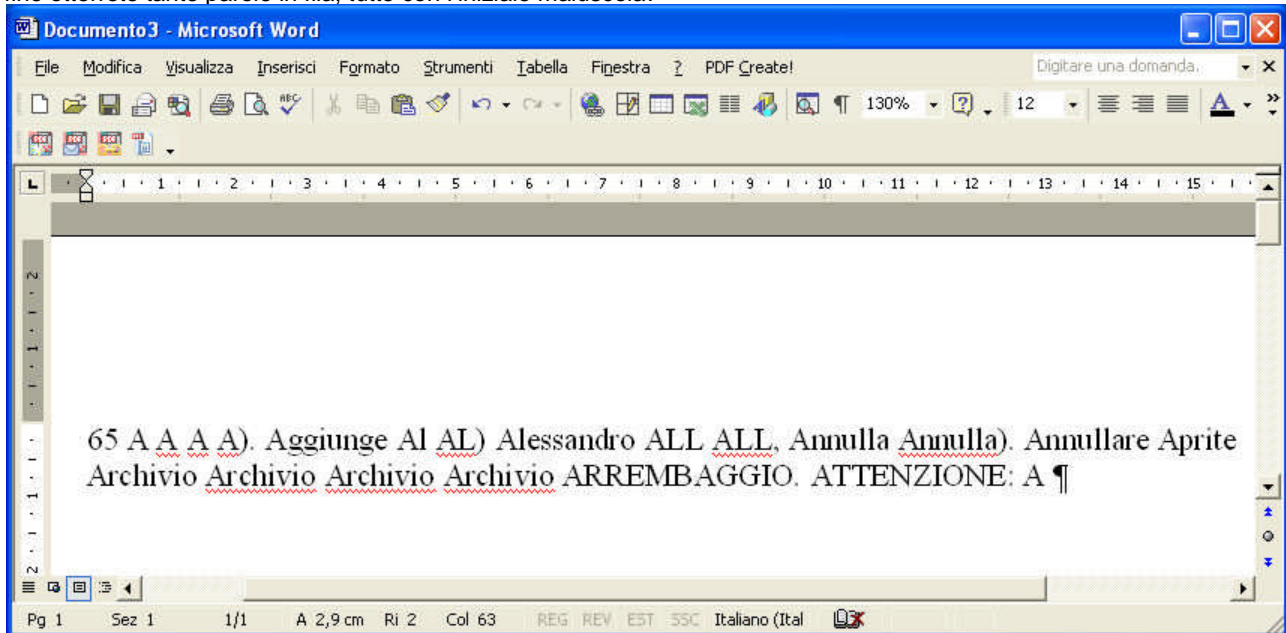


Nonostante sembri un'operazione superflua, nella finestra di *Blocco Note* selezionate l'intero testo (menu *Modifica / Seleziona tutto*) e copiate (menu *Modifica / Copia*). Aprite ora *Word* e, in un documento nuovo, effettuate l'operazione inversa (menu *Modifica / Incolla*). Se quest'operazione fosse stata compiuta subito dopo aver copiato i dati da *Excel* (senza cioè "passare" attraverso *Blocco Note*) avreste potuto incontrare alcune difficoltà (provare per credere!).

A questo punto, che cosa abbiamo ottenuto? Un documento di *Word* che contiene tutte le parole con iniziali maiuscole (anche se relative a una sola lettera dell'alfabeto) infarcite purtroppo di una quantità enorme di caratteri di paragrafo (**^p**, ricordate?): bisogna eliminarli. Per far questo li convertiamo dapprima in spazi (menu *Modifica / Sostituisci* e poi **^p** nel campo *Trova* e carattere *spazio* nel campo *Sostituisci con*).

Sembrebbe esser caduti dalla padella nella brace: abbiamo eliminato tutti i segni di paragrafo, ma sono stati sostituiti da un'infinità di spazi bianchi. Niente paura: basterà ricorrere più volte alla funzione *Sostituisci con* ricorrendo a un truccetto:

Nel campo *Trova* digitate **una dozzina** di spazi e nel campo *Sostituisci con* digitate **un unico** spazio. Ripetete l'operazione fino a che non verranno notificate **0** sostituzioni. A questo punto ripetete l'operazione digitando **quattro spazi** in *Trova* e sempre **uno solo** in *Sostituisci con*. Procedendo infine alla sostituzione di due spazi con uno solo, alla fine otterrete tante parole in fila, tutte con l'iniziale maiuscola.



6- Conclusioni

La procedura è un po'(!) complessa, ma sono sicuro che i miei lettori troveranno metodi per creare macroistruzioni, sia in *Excel* che in *Word*, per facilitare e snellire la procedura.

Questo Tutorial è stato redatto il giorno **29 giugno 2009**. Per verificare la disponibilità di eventuali aggiornamenti sul presente Tutorial, collegarsi al sito dell'autore: <http://www.alessandrodesimone.net>