

calculation of number of words / various words

set t "Se vi ho raccontato tanti particolari sull'asteroide B 612 e se vi ho rivelato il suo numero, e proprio per i grandi che amano le cifre."

set e "If I have told you so many details about asteroid B 612 and if I have revealed its number, it is precisely for adults who love numbers."

```
.t insert end "italian text\n$t\n"  
### only lower case letters  
set t [string tolower $t]  
### no punctuation  
regsub -all {[.,?!]+} $t " " t  
# regsub -all {[']+} $t " " t
```

```
set n 0  
.t insert end "\n\nAnalysis\n-----\n"  
set l [llength $t]  
.t insert end "italian text, lower case without punctuation \n-----\n$t\n"  
.t insert end "number of words $l\n"
```

```
##### find out how often words occur #####  
while {$n < $l} {  
set wort [lindex $t $n]  
### .t insert end $zeichen  
incr word($wort)  
incr n  
}  
set z [array get word]  
set l [llength $z]  
.t insert end "various words [expr $l/2]\n"
```

```
set n 0  
set zz ""  
while {$n < $l} {  
set a [lindex $z $n]  
incr n  
set b [lindex $z $n]  
set zz "$zz {$a $b}"  
incr n  
}  
# .t insert end "$zz\n\n"
```

```
set zz [lsort -ascii -index 0 $zz]  
set yy [lsort -integer -decreasing -index 1 $zz]  
# .t insert end "\n\n$yy\n\n"
```

```
##### print table #####  
.t insert end "\nword      number\n"  
.t insert end "=====\n"  
set l [llength $yy]  
set n 0  
while {$n < $l} {  
set a [lindex $yy $n]  
.t insert end [lindex $a 0]  
.t insert end "          [lindex $a 1]\n"  
incr n  
}  
.t see end  
#####
```

result

italian text

Se vi ho raccontato tanti particolari sull'asteroide B 612 e se vi ho rivelato il suo numero, e proprio per i grandi che amano le cifre.

Analysis

italian text, lower case without punctuation

se vi ho raccontato tanti particolari sull'asteroide b 612 e se vi ho rivelato il suo numero e proprio per i grandi che amano le cifre

number of words 26

various words 22

word	number
=====	
e	2
ho	2
se	2
vi	2
612	1
amano	1
b	1
che	1
cifre	1
grandi	1
i	1
il	1
le	1
numero	1
particolari	1
per	1
proprio	1
raccontato	1
rivelato	1
sull'asteroide	1
suo	1
tanti	1

calculation of number of charcters / various charcters

set t "Se vi ho raccontato tanti particolari sull'asteroide B 612 e se vi ho rivelato il suo numero, e proprio per i grandi che amano le cifre."

set e "If I have told you so many details about asteroid B 612 and if I have revealed its number, it is precisely for adults who love numbers."

```
.t insert end "english translation\n$t\n"
.t insert end "italian text\n$t\n"
```

```
set n 0
.t insert end "\n\nAnalysis\n-----\n"
##### 1 is the number of characters #####
set l [string length $t]
.t insert end "number of characters $l\n"
```

```
##### find out how often characters occur #####
while {$n < $l} {
set zeichen [string index $t $n]
if {$zeichen == " "} {set zeichen "spaces"}
incr letter($zeichen)
incr n
}
set z [array get letter]
set l [length $z]
.t insert end "various characters [expr $l/2]\n"
```

```
set n 0
set zz ""
while {$n < $l} {
set a [lindex $z $n]
incr n
set b [lindex $z $n]
set zz "$zz {$a $b}"
incr n
}
# .t insert end "$zz\n\n"
```

```
set yy [lsort -integer -decreasing -index 1 $zz]
# .t insert end "\n\n$yy\n\n"
```

```
##### print table of characters #####
.t insert end "\ncharacter number utf-8-code\n"
.t insert end "=====\n"
set l [length $yy]
set n 0
while {$n < $l} {
set a [lindex $yy $n]
set kod [format %4.4x [scan [lindex $a 0] %c]]
if {[lindex $a 0] == "spaces"} {set kod 0020}
.t insert end [lindex $a 0]
.t insert end " [lindex $a 1]"
.t insert end " \\u$kod\n"
incr n
}
.t see end
#####
```

result

english translation

If I have told you so many details about asteroid B 612 and if I have revealed its number, it is precisely for adults who love numbers.

italian text

Se vi ho raccontato tanti particolari sull'asteroide B 612 e se vi ho rivelato il suo numero, e proprio per i grandi che amano le cifre.

Analysis

number of characters 136

various characters 27

character	number	utf-8-code
spaces	25	\u0020
e	12	\u0065
i	12	\u0069
o	12	\u006f
r	11	\u0072
a	10	\u0061
t	7	\u0074
l	6	\u006c
c	5	\u0063
n	5	\u006e
p	4	\u0070
s	4	\u0073
u	3	\u0075
v	3	\u0076
h	3	\u0068
d	2	\u0064
m	2	\u006d
1	1	\u0031
2	1	\u0032
B	1	\u0042
S	1	\u0053
f	1	\u0066
6	1	\u0036
g	1	\u0067
'	1	\u0027
,	1	\u002c
.	1	\u002e

calculation number of syllables containing diphthongs and vowels

set t "Se vi ho raccontato tanti particolari sull'asteroide B 612 e se vi ho rivelato il suo numero, e proprio per i grandi che amano le cifre."

Syllables contain diphthongs or vowels or even sometimes only consonants like "krk" in czech language
It is not always simple to divide a word into syllables: e.g. (german word) Bäckerei formerly Bäk-ke-rei today Bäk-ke-rei but never Bäk-cker-ei
The syllables in the italian sentence contain the following diphthongs and vowels: io oi a e i o u

```
.t insert end "\n\n$t\n"
set n1 [regsub -all "oi" $t "xx" t]
.t insert end "\n$n1 x oi\n"
set n2 [regsub -all "io" $t "xx" t]
.t insert end "\n2 x io\n"
set n3 [regsub -all "a" $t "x" t]
.t insert end "\n3 x a\n"
set n4 [regsub -all "e" $t "x" t]
.t insert end "\n4 x e\n"
set n5 [regsub -all "i" $t "x" t]
.t insert end "\n5 x i\n"
set n6 [regsub -all "o" $t "x" t]
.t insert end "\n6 x o\n"
set n7 [regsub -all "u" $t "x" t]
.t insert end "\n7 x u\n"
.t insert end "-----\n"
.t insert end "sum [expr ($n1 + $n2 + $n3 + $n4 + $n5 + $n6 + $n7)] syllables\n"
#####
```

result

Se vi ho raccontato tanti particolari sull'asteroide B 612 e se vi ho rivelato il suo numero, e proprio per i grandi che amano le cifre.

```
1 x oi
1 x io
10 x a
12 x e
10 x i
10 x o
3 x u
-----
sum 47 syllables
```

calculation of word roots in Esperanto

in Esperanto the meaning of a word is composed by the meanings of the root-words it consists

e.g. plen'kresk'ul'o'j plen = completely kresk = grown-up ul = person o = substantive j = plural this means "adults"
global inhalt

definition of the list of root-words

```
set inhalt "a' al am' as' asteroid' da detal' e' est' ĝi' ĝust' is' j' kaŝ' kaj kiu' kresk' la mal' mi n' nombr' numer' o' por' plen' pri rakont' se ŝip' tio tiom ul' vapor' vi "
```

return the root, when it is found in the root-list

```
proc Wortwurzel { wurzel } {  
  global inhalt  
  set z [lsearch $inhalt $wurzel]  
  if {$z > -1} {set wurzel $wurzel}  
  if {$z == -1} {set wurzel ""}  
  return "$wurzel"  
}
```

find out wheather the part of the word is a root-word and separate the root-word from the rest of the word with a ""-character

```
proc Teil { wort } {  
  set vorto $wort  
  set wort "$wort {}"  
  set letztes [lindex $wort end-1]  
  set l [string length $letztes]  
  set n 0  
  # .t insert end Anfang\n  while {$n < 100} {  
    incr n  
    incr l -1  
  
    set wort [lreplace $wort end-1 end-1 [string range $letztes 0 $l]]  
    set wort [lreplace $wort end end [string range $letztes [expr $l + 1] end]]
```

```
# .t insert end "$n Wort $wort letztes $letztes l $\n"  
set gefunden [Wortwurzel [string range $letztes 0 $l]]
```

```
if {$gefunden != ""} {  
  # .t insert end "### gefunden $gefunden "  
  set letztes [lindex $wort end]  
  set l [string length $letztes]  
  set wort "$wort $letztes"  
  # .t insert end "# letztes $letztes xxx\n"  
  if {[lindex $wort end] == ""} {  
    set vorto "[string map {" " "} [lrange $wort 0 end-1]]"  
    ## .t insert end "$vorto "  
    break  
    set letztes "[lindex $wort end-1]"  
    set l [string length [lindex $wort end-1]]  
    incr l -1  
    # .t insert end "#### W $wort letztes $letztes l $\n"  
  }  
}
```

```
if {[lindex $wort 0] == ""} {  
  ## .t insert end "$vorto "  
  set n 111  
  break  
  if {[lindex $wort end-1] == ""} {  
    set wort [lreplace $wort end-1 end-1]  
    set letztes [lindex $wort end-1][lindex $wort end]  
    set l [string length [lindex $wort end-1]]  
    incr l -1  
    # .t insert end "Teilwort zu Ende  
    # wo $wort l $l letztes $letztes\n" }  
  .t see end  
}  
return $vorto  
}
```

```
set text "se mi rakontis al vi tiom da detaloj pri la asteroido b 612 kaj se mi malkaŝis ĝian numeron tio estas ĝuste por plenkreskuloj  
kiuj amas nombrojn"
```

```
set n 0  
set laenge [llength $text]
```

```
.t insert end "\n\ntext  
=====  
}
```

\$text\n\n"

.t insert end "\n\nlist of root-words:

=====

\$inhalt\n\n\n"

.t insert end "words divided into root-words

=====\n"

.t see end

devide the word into possible root-words

while {\$n < \$laenge} {

set wort [lindex \$text \$n]

set vorto "[Teil \$wort] "

.t insert end "\$vorto "

incr n

}

.t insert end \n

.t see end

#####

result

text

=====

se mi rakontis al vi tiom da detaloj pri la asteroido b 612 kaj se mi malkaŝis ĝian numeron tio estas ĝuste por plenkreskuloj kiuj amas nombrojn

list of root-words:

=====

a' al am' as' asteroid' da detal' e' est' ĝi' ĝust' is' j' kaŝ' kaj kiu' kresk' la mal' mi n' nombr' numer' o' por plen' pri rakont' se ŝip' tio tiom ul' vapor' vi

words divided into root-words

=====

se mi rakont'is al vi tiom da detal'o'j pri la asteroid'o b 612 kaj se mi mal'kaŝ'is ĝi'a'n numer'o'n tio est'as ĝust'e por plenkresk'ul'o'j kiuj am'as nombr'o'j'n