THE COMMONEST DOUBLE WORD SQUARES

M. DOUGLAS McILROY
Bernardsville, New Jersey

Draw a stock of n-letter words in decreasing order of frequency from Kucera and Francis's Computational Analysis of Present-Day American English until they yield a double word square. For n = 2, 3, 4, 5, word stocks of 15, 66, 326 and 1444 are needed, in good agreement with Chris Cole's theoretical predictions of 16, 63, 250 and 992 given in "Mathematics of Square Construction" in February 1993. The Kucera and Francis frequencies (in words per million) are given next to the words in each square.

IN  21420    MAN  1238    SHOW  291    GLARE  7
TO  26253    AGO  246    NONE  108    ROMAN  59
IT  8768     YET  420    OPEN  332    AVOID  58
NO  2219     MAY 1399    WEST  244    DENSE  9
            AGE  242    SNOW  62    EDGED 11
            NOT  4618    HOPE 178    GRADE 49
            age 453    ONES 116    LOVED 56
            man 1238    WENT 507    AMONG 370
            set 414    ROSE  89    RAISE  52
            are 453    EVIL  72    ENDED  61
            sets 66
            tree 62

Some runners-up include

AN  3745    SAM  79    REST  164    SCARS 10
TO  26253    ERA  30    OVER 1306    HABIT 23
           TEN 165    SITE  66    ANODE 77
AT  5383     SET  414    ELSE 176    ROUGE  7
NO  2219     ARE  453    ROSE  89    PETER 36
           MAN 1238    EVIL  72    SHARP  75
           sets 66
           tree 62
           about 1817
           ridge 18
           steer  9