

4 DATA

4.1 SBĚR, ORGANIZACE A ZPRACOVÁNÍ DAT

4.1.1 SBĚR A TŘÍDĚNÍ SOUBORU DAT – POROVNÁNÍ TEXTŮ

Studna na konci světa

Byla jednou jedna stará vdova, která žila v malé chaloupce se svou jedinou dcerou, a to vám bylo tak pěkné děvče, že byla radost se na ni podívat.

Jednoho dne si stařenka vzala do hlavy, že upeče plnou mísu, event. pekáč lívanečků. A tak si na stůl připravila vál a potom šla do spíže a přinesla mísu mouky. Když ale chtěla přinést džbánky vody, aby ji smíchala s tou moukou, zjistila, že doma žádnou nemá.

The Well on the World's End

There was once an old widow woman, who lived in a little cottage with her only daughter, who was such a bonnie lassie that everyone liked to look at her.

One day the old woman took a notion into her head to bake a girdleful of cakes. So she took down her bakingboard, and went to the meal-chest and fetched a basinful of meal; but when she went to seek a jug of water to mix the meal with, she found that there was none in the house.

Zjisti, z kolika písmen se skládá každé slovo v českém (Č) i anglickém (A) textu, a vyplň následující tabulku.

Počet písmen ve slově	1	2	3	4	5	6	7	8	9	10 a víc
Počet slov – Č										
Počet slov – A	5	14	28	23	13	4	2	3		

a) Sestroj sloupcový graf znázorňující zastoupení slov podle počtu písmen v Č a A.

b) Porovnej rozdělení četností slov podle počtu písmen v Č a A. Výsledek zdůvodni.

c) Vyhledej v obou textech předložky a porovnej četnosti jejich výskytu v Č a A. Zdůvodni rozdíl.

✂ ----- ↓ PŘED KOPÍROVÁNÍM PRO ŽÁKY OD TOHOTO MÍSTA ZAKRÝT ↓ ----- ✂

VÝSLEDKY: Tabulka – Počet slov Č: 6, 15, 9, 13, 17, 7, 5, 4, 2, 1; Počet slov A: 5, 14, 28, 23, 13, 4, 2, 3, 1, 1;

b) největší četnost mají tří- a čtyřpísmenná slova v A (členy a zájmena), v Č jsou to dvoj-, čtyř- a pěti-písmenná slova (předložky, zvratná zájmena, částice, slovesa a podstatná jména);

c) předložky jsou v A výrazně početnější oproti Č. Čeština na rozdíl od angličtiny používá koncovky.

KOMENTÁŘ: Úloha od žáků vyžaduje pozornou práci s textem, dobrou organizaci evidence a třídění údajů pro doplnění tabulky, vizualizaci grafem. Ukazuje, jak statistika v souvislosti s gramatikou může dávat zajímavé výsledky. Dále je nutné rozhodnout o choulostivých případech. Například v českém textu rozhodnout, zda hláska *ch* je jedno písmeno, nebo dvě písmena (my jsme volili druhou možnost), nebo zda výraz *meal-chest* v angličtině brát jako dvě slova, nebo jako jedno slovo (my jsme volili druhou možnost).