

Univerzita Karlova v Praze

Filozofická fakulta

Ústav informačních studií a knihovnictví

BAKALÁŘSKÁ PRÁCE



Dana Prchalová

Zkoumání čtivosti - srovnání způsobů měření obtížnosti textu

Exploring readability - a comparison of methods of text difficulty measuring

Vedoucí práce:

Praha 2013

doc. PhDr. Richard Papík, Ph.D.

Děkuji vedoucímu práce doc. PhDr. Richardovi Papíkovi, Ph.D. za cenné rady a připomínky k práci. Poděkování patří také pracovníkům oslovených knihoven a nakladatelství za trpělivost a Ing. Pavlovi Střížovi, Ph.D. z Univerzity Tomáše Bati ve Zlíně, zaměstnancům Oddělení dalšího vzdělávání Univerzity Palackého v Olomouci, pracovnícím Pražské pedagogicko-psychologické poradny, spol. s r.o., Mgr. Marcele Drnkové ze ZŠ Lupáčova v Praze a Mgr. Haně Novákové ze ZŠ T. Stolzové v Kostelci nad Labem za velkou pomoc a ochotu.

„Podle mého názoru dá člověku kontakt s knihou nejvíc.“

Irena Šormová, vedoucí oddělení metodiky
Městské knihovny v Praze

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 27. 7. 2013

Dana Prchalová

Abstrakt: Tato bakalářská práce se skládá ze dvou hlavních částí. V první části jsou po stručném přehledu historie problematiky představeny nejpoužívanější způsoby měření čtivosti textu v českém a anglickém jazyce a nastíněny přístupy v některých jazycích dalších. U každého způsobu je rozebrán vzorec pro výpočet čtivosti a instrukce k jeho použití. V druhé, praktičtější části se práce zaměřuje na používání či nepoužívání těchto vzorců v praxi, uvádí problémy, které jsou s nimi spojeny, a diskutuje jejich spolehlivost za pomoci vlastních měření autorky. Nové je také využití vzorců ke sledování vývoje čtivosti textu v rámci literárního díla.

Klíčová slova: čtivost, obtížnost textu, vzorce, dynamika textu, národní korpus

Abstract: This bachelor thesis consists of two main parts. Following a short historical overview in the field, the most used readability formulæ for Czech and English languages are presented in the first part, with peeks to couple of solutions in other languages as well. Each readability measure is presented with an equation and instructions for its use. In the more practical second part, the work focuses at using or not using the formulæ on everyday basis, presents problems related to them and discusses their reliability supported by author's own measurements. Measuring the trend of readability within a literary work represents a novel use of the formulæ.

Keywords: readability, text complexity, formulas, dynamics of text, national corpus

Obsah

Obsah	5
Předmluva	7
Úvod.....	8
1. Vývoj problematiky vzorců pro měření čtivosti	9
2. Přehled vzorců	11
2.1 Vzorce pro český jazyk	12
2.1.1 FOG Index	12
2.1.2 Mistríkův vzorec	13
2.1.3 Komplexní míra obtížnosti textu	14
2.2 Vzorce pro anglický jazyk	17
2.2.1 Fryův graf	17
2.2.2 FOG Index	18
2.2.3 SMOG.....	19
2.2.4 Flesch Reading Ease Readability Formula	21
2.3 Vzorce pro jiné jazyky	24
3. Použití v praxi.....	27
3.1 Ukázka výpočtu vybraných vzorců.....	27
3.1.1 Ruční výpočet pro český jazyk	27
3.1.2 Ruční výpočet pro anglický jazyk	28
3.1.3 Počítačové programy pro anglický jazyk.....	30
3.1.4 Počítačové programy pro český jazyk	33
3.2 Vzorec jako prostředek	35
3.2.1 Hodnocení knih.....	35
3.2.2 Dynamika textu.....	37
3.3 Vzorec jako cíl	38
3.4 Spolehlivost.....	40
3.4.1 Vzorové texty.....	40

3.4.2	Terénní výzkum	41
4.	Informační podpora tématu.....	44
4.1	Organizace zabývající se čtením.....	44
4.2	Odborné databáze.....	46
4.2.1	ERIC	46
4.2.2	ProQuest.....	48
4.2.3	EbscoHost	48
4.3	Kvalifikační práce	49
5.	Závěr	50
	Seznam obrázků.....	51
	Seznam tabulek	51
	Seznam grafů	51
	Seznam použité literatury	52
	Příloha A. Výzkumný dotazník.....	55

Předmluva

Na téma měření čtivosti textu jsem psala seminární práci ve 2. ročníku bakalářského studia a objevila jsem několik pro mne zajímavých myšlenek (kterým pravděpodobně nebyla věnována pozornost). Například jde o zkoumání dynamiky textu či spolehlivosti vzorců čtivosti pro český jazyk. V minulosti se na Ústavu informačních studií a knihovnictví FF UK (dříve Katedra vědeckých informací a knihovnictví) problematikou textu, jeho reprezentací v informačních systémech a vztahem k lingvistickým oborům zabývali např. Marie Königová¹ a Vladimír Smetáček.² Problematiky čtení a jeho psychologických a informačních aspektů si všímala ve svých pracích rovněž PhDr. Lidmila Vášová. Téma čtení se objevuje již více než 20 let v osnovách oboru Informační studia a knihovnictví v podobě předmětu Kurz rychlého čtení, který vede Richard Papík.³

Lze také tušit významná rozhraní a souvislosti mezi problematikou čtenářských průzkumů a výzkumů čtivosti. Když jsem proto mezi vypsány tématy bakalářských prací našla téma „Zkoumání čtivosti (readability, readability formulas) - historické a moderní způsoby ‘měření’ čtivosti“, byla to skvělá příležitost, jak tyto myšlenky dále rozvíjet. Bohužel v systému zapsaný vedoucí práce Mgr. Roman Chýla odjel do zahraničí, spolupráce s ním tedy nebyla možná. Proto jsem oslovila tehdejšího vedoucího ústavu Richarda Papíka s žádostí o pomoc při hledání nového vedoucího. Tímto bych mu chtěla poděkovat za velkou ochotu vést mou práci i přes jeho velké pracovní vytížení a za zájem, který o mé téma projevil.

¹ prof. PhDr. Marie Königová, CSc.

² doc. PhDr. Vladimír Smetáček, CSc.

³ doc. PhDr. Richard Papík, PhD.

Úvod

Tato bakalářská práce se zabývá metodami měření čtivosti a obtížnosti textu se zaměřením na vzorce pro výpočet obtížnosti textu. V teoretické části práce je shrnut zásadní vývoj vzorců pro měření čtivosti jak v České republice, tak v zahraničí, jsou vysvětleny hlavní parametry vyskytující se v jednotlivých vzorcích a představeny algoritmy používané pro jednotlivé jazyky, převážně češtinu a angličtinu. Do praktické části je zahrnuta ukázka výpočtu obtížnosti textu vybranými vzorci pro češtinu i angličtinu, provedený průzkum v knihovnách a nakladatelstvích a diskuze o spolehlivosti vzorců, včetně výzkumu provedeného ve 2. a 4. třídě základních škol. Zcela nový pohled pak nabízí zkoumání dynamiky textu.

Je nutné mít na paměti, že parametry textu používané ve vzorcích nejsou zdaleka jedinými faktory, které čtivost ovlivňují. Snadnost čtení ovlivňují i takové prvky, jakými jsou například velikost písma a jeho typ (font), délka odstavců, používání nadpisů, barva papíru a pozadí či množství volného prostoru kolem textu.

1. Vývoj problematiky vzorců pro měření čtivosti

Studium čtivosti textu ve smyslu srozumitelnosti jazyka má dlouhou historii sahající až k antice. Vychází z klasické rétoriky Platóna a Aristotela a z analýzy slovní zásoby Bible, prováděné ve starověku hebrejskými vědci. Přestože tyto zdroje obohatily naše chápání jazyka a efektivitu textu, pro tuto práci jsou relevantnější výzkumy čtivosti ve smyslu obtížnosti čtení samotného, které se začaly objevovat na počátku 20. století, a to především ve Spojených státech amerických.

Potřeba kvantitativně určovat obtížnost textů vznikla ve školském sektoru, zejména pro hodnocení učebnic a odborné literatury. Výzkumné práce se prováděly ve dvou odvětvích: regulace slovní zásoby a měření čtivosti. Regulace slovní zásoby se zabývala sadami slov, které by byly nejefektivnější při učení žáků z učebnic čtení. Zvláště byla sledována „nová slova“ v každé knize, počet jejich opakování a jejich obtížnost. Výzkumy v oblasti měření čtivosti vedly k procedurám a nástrojům snažícím se spolehlivě a účinně odlišit složitější text od lehčího, či texty podle obtížnosti ohodnotit. Počáteční léta výzkumu v této oblasti byla experimentální — vědci vyzkoušeli mnoho proměnných, než zjistili, že nejlepšími ukazateli pro určování čtivosti jsou délka věty a obtížnost slova. Pro novou generaci žáků nastupujících do škol byly stávající učebnice považovány za příliš obtížně srozumitelné, což vedlo k neustálému snižování jejich obtížnosti až do 60. let, kdy se proti výzkumům čtivosti a metodám určování obtížnosti zvedla vlna kritiky.

V roce 1939 byl rovněž zahájen výzkum měření čtivosti u jiných jazyků než angličtina pro potřeby studentů, kteří cizí jazyky studovali. Proto jedním z prvních jazyků, u kterého vědci v USA začali vyvíjet vzorce pro měření čtivosti, byla španělština. Během dalších let byly vytvořeny nástroje i pro ruštinu, němčinu, hebrejštinu, čínštinu či vietnamštinu. Tento výzkum velmi ovlivnil vytváření vzorců pro měření čtivosti po celém světě.

První vzorce pro měření obtížnosti textu v Evropě byly založeny na úpravě Flesch Reading Ease Formula, například pro francouzštinu vzorec Kandel & Moles z roku 1958, či vzorec Fernández Huerty pro španělštinu (1959). Během následujících desetiletí se tato oblast výzkumu rozšířila i do dalších zemí, například Dánska, Švédska, Německa nebo Nizozemí. [Zakaluk, 1988]

V České republice začíná vývoj měření čtivosti textu až v sedmdesátých letech 20. století. Průkopníkem výzkumu v tomto oboru je Vladimír Smetáček, který se této problematice začal věnovat v době, kdy ještě žádný vzorec pro měření čtivosti textu v českém jazyce neexistoval.

Ve své knize **ČTIVOST TEXTŮ PRO DĚTI** z roku 1973 se zabýval prvky, které by rozlišovaly texty dle věku předpokládaného čtenáře. Přestože z jeho výzkumů nebyl sestaven žádný vzorec pro měření čtivosti textů pro děti, této oblasti se věnuje i nadále.

Pro český jazyk se začal používat *Mistríkův vzorec*, vytvořený Jozefem Mistríkem⁴ v roce 1968 pro slovenský jazyk. Dále byl z angličtiny převzat nejjednodušší ze vzorců — *FOG Index*, který vyvinul Robert Gunning roku 1952. Roku 2002 pak byla publikována *Komplexní míra obtížnosti textu* autorů Nestlerové, Pluskala a Průchy.

⁴ Prof. PhDr. Jozef Mistrík, DrSc.

2. Přehled vzorců

Většina vzorců pracuje se stejnými parametry — délka věty, délka slova a počet rozdílných slov.

„Čím je věta delší, tím obtížnější je model skutečnosti přenášený textem. Dlouhé věty se člení do souvětí s větami nadřazenými a podřazenými, zachycujícími složité příčinné a časové vztahy. Proto může být délka věty indikátorem umožňujícím stanovit obtížnost textu. Délku věty vyjadřujeme počtem slov. Pro vysokou frekvenci některých větných členů je možné vyřadit z výpočtu nejfrekventovanější členy — spojky, předložky apod.“ [Jonák, 2000]

Charakteristika významu délky slova pro analýzu obtížnosti textu je podle autora článku **CO VŠE BY MĚL (MOHL) O KNIZE VĚDĚT ŠKOLNÍ KNIHOVNÍK**, PhDr. Zdeňka Jonáka, následující:

„Délka slova měřená v slabikách je rovněž příznakem obtížnosti textu. Čím delší slovo, tím obtížněji se vnímá. Jelikož platí, že dlouhá slova se častým používáním zkracují, je možné rovněž předpokládat, že mezi dlouhá slova budou patřit ta, která jsou v komunikaci méně používána.“ [Jonák, 2000]

A takto doktor Jonák charakterizuje poslední z kritérií:

„V delším textu se některá slova vyskytují s menší frekvencí, jiná častěji. Lze předpokládat, že slova méně častá budou při porozumění působit čtenáři větší obtíže než slova, s nimiž se v textu setká častěji. Předpokládáme proto, že čím větší je počet slov, které se vyskytují v textu pouze jednou nebo s minimální frekvencí, tím bude text obtížnější.“ [Jonák, 2000]

Tento fakt si zřejmě uvědomila i internetová encyklopedie Wikipedia, která články ze své anglické verze začala nabízet i v tzv. *Simple English*, využívající maximálně 3000 nejpoužívanějších anglických výrazů. V roce 2004, kdy projekt *Simple English* začal, měla anglická verze Wikipedie přes 150 000 standardních článků. K červnu 2012 k tomu nabízí přes 84 000 článků za použití jen toho nejnútnějšího z anglického jazyka. Tyto články jsou také kratší než jejich protějšek v angličtině nezjednodušené. Wikipedia tak činí informace přístupnější dětem, studentům nebo například cizincům, kteří se chtějí naučit anglicky.

2.1 Vzorce pro český jazyk

Pro český jazyk existují pouze tři vzorce pro měření čtivosti, tzv. *FOG Index*, *Mistikův vzorec* a *Komplexní míra obtížnosti textu*.

V dnešní době lze tuto analýzu zadat některému z počítačových programů, které se zabývají textovou analýzou a tuto funkci ovládají. Zatímco pro anglické texty je takových programů poměrně dostatek, na české texty neexistuje téměř žádný, přestože používané vzorce nejsou nijak složité.

Kromě toho má Česká republika významná pracoviště zabývající se českým jazykem a lingvistikou, ať už se jedná o Ústav českého jazyka FF MU, Ústav teoretické a počítačové lingvistiky FF UK, Ústav formální a aplikované lingvistiky MFF UK nebo Ústav pro jazyk český AV ČR. V čele s Ústavem českého národního korpusu FF UK se mimo jiné podílejí i na tvorbě Českého národního korpusu, což je soubor počítačově uložených textů, který slouží k jazykovému výzkumu. Český národní korpus je budován od roku 1994 a obsahuje texty čítající v souhrnu více než miliardu slov psané i mluvené češtinou z různých žánrů a stovek zdrojů. Data jsou každému zájemci k dispozici na adrese <http://www.korpus.cz/> pro nekomerční účely. Pracoviště se dále zabývají např. gramatickou analýzou textů, strojovými překlady, porozuměním textu a většinu svých výsledků a nástrojů zveřejňují a dávají k dispozici veřejnosti. Použité slovní druhy, tvary slov, šířka nebo hloubka stromu větného rozboru, četnost použitých slov nebo slovních spojení jsou všechno parametry, které mají velkou vypovídací hodnotu a jsou díky těmto ústavům snadno k dispozici. Je proto až překvapující, že nic z toho nebylo dosud využito ke zjišťování obtížnosti či čtivosti textu.

2.1.1 FOG Index

Tzv. *FOG Index* popsal Američan Robert Gunning v roce 1952 ve své publikaci **THE TECHNIQUE OF CLEAR WRITING** [Gunning, 1952] určené především pro spisovatele v anglickém jazyce. Tento vzorec se řadí mezi nejjednodušší způsoby výpočtu čtivosti, pracuje pouze s počtem slov, počtem vět a počtem „těžkých“ slov v úryvku textu o přibližně sto slovech.

Na internetových stránkách Centra distančního vzdělávání (CDV) Univerzity Palackého v Olomouci [Univerzita Palackého v Olomouci, 2005], v sekci pro autory, se využívá následující varianta výpočtu *FOG Indexu*:

1. Vyberte soubor textu s přesným počtem 100 slov.
2. Spočítejte počet vět v souboru a zaokrouhlete na celé číslo.
3. Dělte 100 počtem vět a dostanete průměrný počet slov ve větě.
4. Spočítejte počet slov v souboru, která mají 3 a více slabik. Opakuje-li se slovo, započítejte ho pokaždé.
5. Sečtěte průměrný počet slov ve větě a počet slov, která mají 3 a více slabik.
6. Vynásobte 4, vydělte 10 a přičtěte 5.

Vzorec pak vypadá takto:

$$FOG Index_{CDV} = 5 + 0,4 \cdot (A + L)$$

kde

A je průměrný počet slov ve větě,

L je počet tří a víceslabičných slov.

Od obecného Gunningova vzorce se v rámci zjednodušení liší v počítání s přesně 100 slovy (počet vět se zaokrouhluje) a přičtením konstanty, jejíž smysl je diskutabilní, a působí spíše jako značka CDV. Interpretace výsledků je rovněž upravena:

20 – 24	materiál je jednoduchý ke čtení a pochopení, věty jsou krátké, je zde málo dlouhých slov
25 – 29	materiál je dosti obtížný, některé věty je třeba číst i vícekrát než jednou
30 – 39	odborná kniha
40 a více	těžký až nečitelný text

Tabulka 1 Stupnice *FOG Indexu* dle CDV ÚPOL

Před několika lety navrhl z vlastního zájmu Ing. Pavel Stříž z Univerzity Tomáše Bati ve Zlíně na základě absolvování kurzu Centra distančního vzdělávání základ aplikace pro výpočet *FOG Indexu* dle tohoto postupu. Centrum tento program přivítalo a dále rozšířilo a doplnilo, nyní ho nabízí zdarma účastníkům svých kurzů.

2.1.2 Mistríkův vzorec

Mistríkův vzorec je pojmenován podle svého tvůrce, Jozefa Mistríka, který ho vyvinul roku 1968 pro slovenský jazyk [Mistrík, 1968]. Tento algoritmus je na výpočet o něco složitější než

FOG Index, kromě průměrné délky vět a slov pracuje i s počtem rozdílných slov. Mistríkovi pokusy ukázaly, že zhruba 300 slov postačuje pro reprezentativnost vzorku.

$$R = 50 - S \cdot V \cdot \frac{L}{N}$$

S je průměrný počet slabik ve slově,

V je průměrný počet slov ve větě,

L počet různých slov ve vzorku,

N počet všech slov ve vzorku.

V případě *Mistríkova vzorce* je stupnice opačná:

50 – 40	velmi lehké texty, narativní, konverzační, odpočinkové a zábavné
40 – 30	texty průměrně obtížné, lehce srozumitelné, dají se číst plyně
30 – 20	texty výkladové, náročné, ale srozumitelné
20 – 10	texty stylizované, těžko srozumitelné, vhodné ke studiu, nikoli ke čtení
10 – 0	text na hraně srozumitelnosti

Tabulka 2 Stupnice Mistríkova vzorce

Podle Mistríka je teoreticky možno dosáhnout i záporných hodnot, ale v takovém případě by se nejednalo o „kontexty“ v pravém slova smyslu, ale o umělé, syntakticky nepřehledné texty [Mistrík, 1968]. Ve své práci se Mistrík zabýval také možným rozšířením vzorce o faktografické termíny F (např. čísla, vlastní jména) a hloubku syntaktické struktury C :

$$R = 50 - S \cdot V \cdot \frac{L}{N} \cdot \left(1 + \frac{N}{F}\right) \left(1 + \frac{C}{N}\right)$$

Ukázalo se však, že celkový výsledek se tím změní o méně, než je předpokládaná odchylka původního vzorce.

Přestože byl vzorec původně vyvinut pro slovenský jazyk, verifikován byl i na dalších jazycích. Podle empirických zjištění nevykazovaly slovanské jazyky odchylky, proto je *Mistríkův vzorec* použitelný pro český jazyk bez jakýchkoliv úprav [Mistrík, 1968].

2.1.3 Komplexní míra obtížnosti textu

Komplexní míra obtížnosti textu neboli míra obtížnosti didaktických textů je vzorec pro měření čtivosti didaktických textů (učebnic), který původně vyvinula německá psycholožka Kathrin Nestlerová v roce 1982. Algoritmus byl převzat a v některých částech upraven, aby vyhovoval pro české texty. Vznikla tak míra obtížnosti, která byla českými a slovenskými pedagogy a psychology mnohokrát aplikována na velký počet nejrůznějších učebnic, včetně

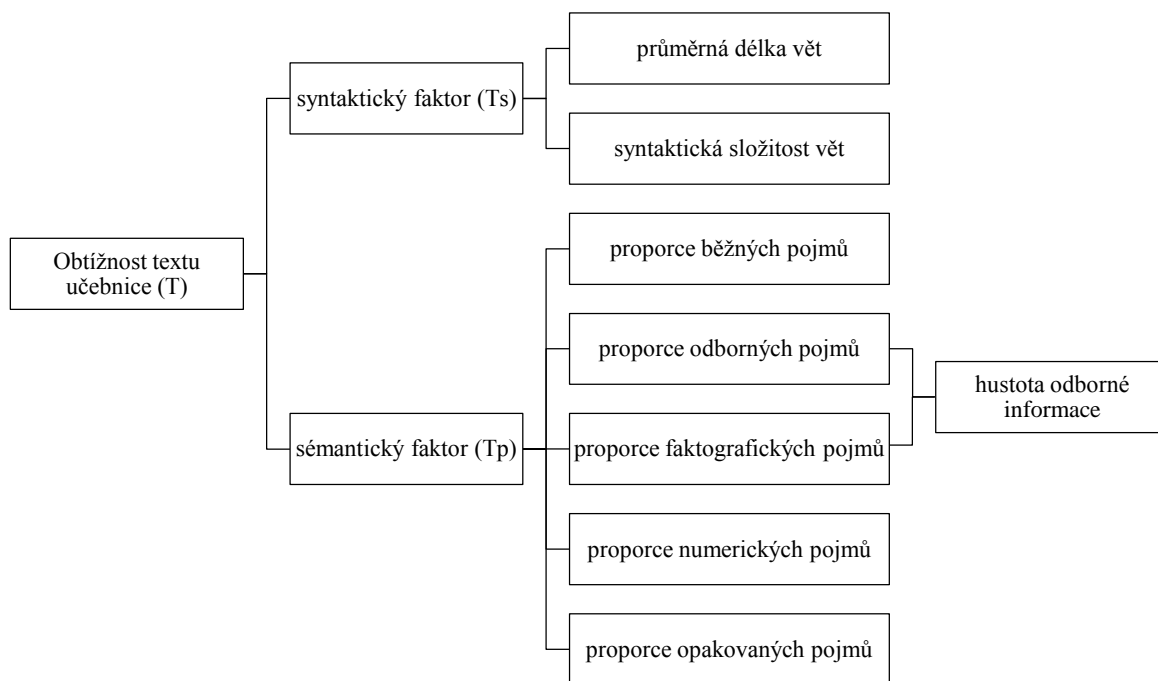
vysokoškolských. Míra byla nadále zdokonalována a je používána i při evaluaci nově vznikajících učebnic např. na Slovensku.

Na rozdíl od předchozích dvou vzorců se *komplexní míra obtížnosti* specializuje na hodnocení odborných textů, nikoliv krásné literatury. Základem je posouzení jak syntaktické, tak sémantické obtížnosti textu.

$$T = T_s + T_p$$

kde T_s (syntaktická obtížnost) a T_p (pojmová, sémanticko-lexikální obtížnost) se vypočítávají dle samostatných vzorců a výsledkem je počet bodů. Původně byla míra založena na dvou parametrech odrážejících syntaktickou složitost textu a osmi parametrech vyjadřujících sémantickou složitost.

V devadesátých letech 20. století byla míra zdokonalena M. Pluskalem [Pluskal, 1997]. Hlavní inovace spočívá v tom, že se rozšířil počet parametrů reflektujících sémantickou obtížnost textu, takže v současné době lze zdroje míry obtížnosti didaktických textů schematicky popsat takto:



Obrázek 1 Schematický popis zdrojů Komplexní míry obtížnosti [PRŮCHA, 2013]

Parametry jsou následující:

$$T_s = 0,1 \cdot \bar{V} \cdot \bar{U}$$

kde

\bar{V} je průměrný počet slov ve větě,

\bar{U} je průměrný počet slov ve větných úsecích.

$$T_p = \frac{P}{N} \cdot \frac{P_B + 3P_O + 2P_F + 2P_C + Z}{N} \cdot 100 \%$$

kde

N je celkový počet slov ve vzorku,

P je celkový počet pojmů ve vzorku, tj. jmen,

P_B je počet běžných pojmů ve vzorku,

P_O je počet odborných pojmů ve vzorku,

P_F je počet faktografických pojmů ve vzorku,

P_C je počet číselných pojmů ve vzorku,

Z je počet opakovaných pojmů ve vzorku,

přičemž za odborné pojmy se považují jména, která mají platnost vědeckých termínů v určitém oboru. Ze vzorce je zřejmé, že tyto pojmy mají největší podíl na sémantické obtížnosti textu. Mezi faktografické pojmy se řadí vlastní jména osobní, názvy lidských výtvorů a zařízení, států, národů, institucí, uměleckých výtvorů, zeměpisná jména a názvy přírodních jevů a zkratky [Hodis, 2003].

Komplexní míra obtížnosti hodnotí také hustotu odborné informace v textu dvěma samostatnými koeficienty:

$$i = \frac{P_O + P_F + P_C}{N} \cdot 100 \%$$

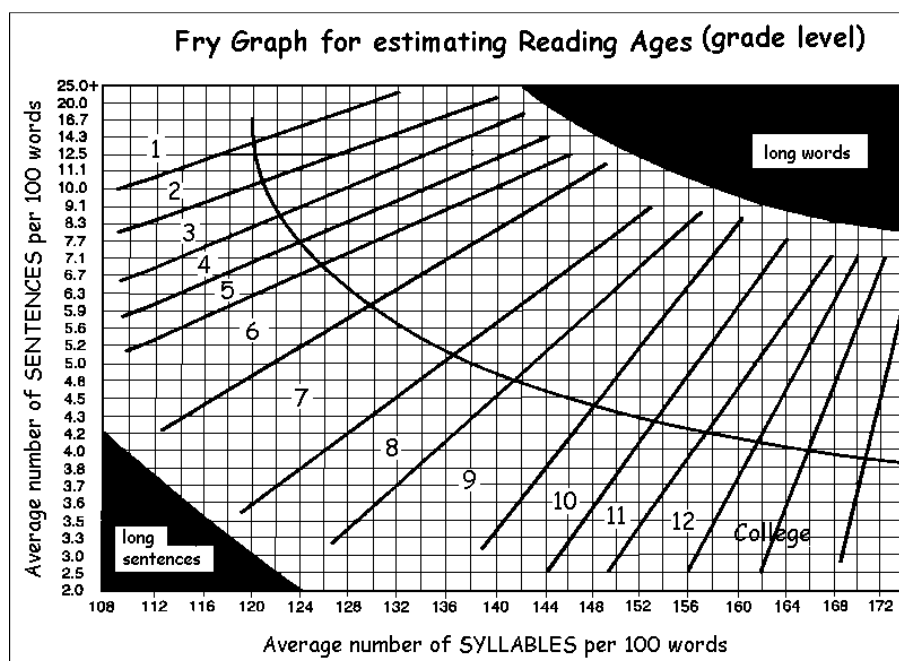
$$h = \frac{P_O + P_F + P_C}{P} \cdot 100 \%$$

Čím větší jsou hodnoty, koeficienty i výsledné body, tím složitější je text. Míra obtížnosti didaktických textů je tak nejkomplicovanějším vzorcem pro výpočet čtivosti textu pro český jazyk. Aplikace vzorců vyžaduje odborné znalosti z oboru, a proto není vhodná ani pro běžné čtenáře ani pro strojové zpracování.

2.2 Vzorce pro anglický jazyk

2.2.1 Fryův graf

Nejrozšířenější graf pro měření obtížnosti textu vytvořil Edward Fry, bývalý člen Rutgers University Reading Center. Tento graf se velmi snadno používá, pracuje pouze se dvěma veličinami – počet vět a počet slabik v úryvcích o přesně sto slovech, obvykle jsou zpracovávány tři vzorky. Tyto hodnoty se zanesou do grafu, kde jejich průsečík určí počet let vzdělání zamýšleného čtenáře.



Obrázek 2 Stupnice Fryova grafu

Graf je dostupný z <http://school.discoveryeducation.com/schrockguide/fry/fry2.html> (se svolením autora).

Křivka grafu znázorňuje průměrné texty, body pod křivkou značí vyšší než průměrnou délku vět, body nad křivkou naopak znamenají těžší slovní zásobu, jako například ve středoškolských učebnicích přírodních věd. Ve výjimečných případech může testovaný vzorek padnout do jedné z černých oblastí v rozích grafu, v tomto případě je měření neplatné, nebo nastala chyba ve výpočtu.

V případě, že je zjištěna větší variabilita ve výsledcích, je doporučeno pro vyšší přesnost měření vybrat větší počet náhodných vzorků textu.

Tento graf je často využíván k regulačním účelům, například ve zdravotnictví, aby byla publikacím zajištěna taková úroveň obtížnosti, která bude srozumitelná a přístupná širokému okruhu čtenářů.

2.2.2 FOG Index

Jak již bylo zmíněno výše, *FOG Index* pochází z poloviny dvacátého století od Roberta Gunninga, amerického vydavatele učebnic. Tento absolvent Státní univerzity v Ohiu vypořádal, že většině středoškoláků dělá potíže čtení, což souviselo s obtížností textů. Podle Gunningova názoru byly noviny a obchodní dokumenty příliš složité a zamlžovaly význam obsahu. Proto v roce 1944 zahájil rozsáhlý výzkum více než šedesáti deníků a časopisů, aby pomohl autorům a editorům správně psát pro jejich čtenáře. Zanedlouho se *FOG Index* stal nejpoužívanějším vzorcem v oboru [Bogert, 1985].

Obecný vzorec pro *FOG Index* tedy vypadá následovně:

$$FOG Index = 0,4 \cdot \left(\frac{W}{S} + \frac{CW}{W} \cdot 100 \% \right)$$

kde

W je počet slov v úryvku,

S je počet vět v úryvku,

a CW je počet slov o třech a více slabikách („těžká“ slova).

Gunning uvádí, že z těžkých slov se mají vynechat vlastní jména a složeniny a běžné anglické koncovky (*-ed*, *-es*, *-ing*), které z dvouslabičných sloves dělají tříslabičná, nemají být počítány jako slabiky. Výsledný index představuje počet let vzdělání, která musí čtenář absolvovat, aby dokázal snadno přečíst daný text.

V systému amerického vzdělávání:

Snadné čtení	
05	5. ročník
06	6. ročník
07	7. ročník
08	8. ročník
09	1. ročník střední školy
10	2. ročník střední školy
Varovná hranice	
11	3. ročník střední školy
12	4. ročník střední školy
13	1. ročník vysoké školy
14	2. ročník vysoké školy
15	3. ročník vysoké školy
16	4. ročník vysoké školy
17	postgraduální studium

Tabulka 3 Gunningova stupnice FOG Indexu

Texty, u nichž je požadováno všeobecné porozumění, by měly mít *FOG Index* zpravidla nižší než 8 [My Byline Media, nedatováno].

Výzkumy spolehlivosti vzorců čtivosti potvrdily silnou vazbu mezi *FOG Indexem* a skutečnou obtížností textu. Přestože *FOG Index* není ze vzorců nejpřesnější, stal se díky své jednoduchosti populárnější než ostatní vzorce. Obecně se jeho přesnost pohybuje kolem ± 1 vzdělávacího stupně až do třináctého. Pak už přesnost nehraje příliš roli, protože cokoliv nad tuto úroveň je neúměrně obtížné [Bogert, 1985].

2.2.3 SMOG

SMOG Readability Formula, s jistou nadsázkou občas označován také jako *Simple Measurement of Gobbledygook*⁵ vymyslel G. Harry McLaughlin v roce 1969 a publikoval ho ve svém článku *SMOG Grading – A New Readability Formula* v **JOURNAL OF READING** [McLaughlin, 1969]. V tomto článku tvrdí, že přestože je tento vzorec až směšně jednoduchý k použití, je lepší než předchozí algoritmy, což se snaží ve zbytku článku také dokázat.

⁵ Gobbledygook a podobně znějící výrazy označují v angličtině nesmyslnou řeč. Pojem používá ve své práci Robert Gunning. Odkaz je ještě umocněn významy slov *fog* (mlha) a *smog*. McLaughlin v článku uvádí, že svou míru obtížnosti tak nazval na jeho počest (a také dle atmosférických podmínek ve svém rodném městě Londýně).

Vzorec pracuje se třemi vzorky textu o deseti větách, jeden se vybere ze začátku, druhý z prostřední části a třetí ke konci textu. Ve všech větách se sečte počet slov s třemi a více slabikami. Najde se nejbližší číslo tak, aby druhá odmocnina vyšla celočíselně. Odmocní se a k výsledku se přičtou 3.

Pokud by se počet slov nacházel přibližně mezi dvěma takovými čísly, použije se to nižší z nich. Výsledkem je počet let vzdělání, které musí předpokládaný čtenář absolvovat, aby bez potíží porozuměl danému textu, s tolerancí $\pm 1,5$ roku. Vzorec lze zapsat jako

$$g = 3 + \sqrt{p}$$

kde

p je počet tří a více slabičných slov ve 30 větách.

Vzhledem k tomu, že McLaughlin tento vzorec vymyslel, jak sám poznamenává, *1969 BC (Before Computers)*, tedy před rozšířením počítačů, byl jeho vzorec lehce zjednodušen pro ruční počítání. Ve svém článku uvádí i zpřesněnou verzi určenou ke strojovému zpracování, určenou pro libovolný počet vět. Finální rovnice vypadá takto:

$$g = 3,1291 + 1,0430 \cdot \sqrt{p \cdot \frac{30}{V}}$$

kde

V je celkový počet vět v měřeném textu.

Určování srozumitelnosti dle tohoto vzorce se ujalo zejména ve zdravotnictví, kde je *SMOG* považován za nejvěrohodnější ukazatel [Fitzsimmons, 2010]. V instrukcích publikovaných různými zdravotnickými zařízeními se objevují, bez odkazů na zdroje, další pravidla pro výpočet:

- Za větu se považuje posloupnost slov zakončená tečkou, otazníkem nebo vykřičníkem.
- V dlouhých větách se středníkem a v seznamech se každá položka počítá jako samostatná věta.
- Slova se spojovníkem se počítají jako jedno slovo.
- Vlastní jména se započítávají.
- Čísla se započítají, jako by byla čtena.

Rovněž se v těchto materiálech objevily převodní tabulky pro usnadnění výpočtu a pro práci s texty kratšími než 30 vět. Původně byly sestaveny Haroldem C. McGrawem z výzkumného vzdělávacího centra v Marylandu a dnes jsou velmi rozšířené po internetu.

Bohužel, McGraw počítal hodnoty podle zjednodušené verze vzorce. Správné hodnoty (viz Tabulka 4) vycházející přímo z McLaughlinovy rovnice, nebyly dosud publikovány.

počet těžkých slov	počet let vzdělání
2 – 5	5
6 – 10	6
11 – 17	7
18 – 26	8
27 – 37	9
38 – 49	10
50 – 64	11
65 – 80	12
81 – 98	13
99 – 118	14
119 – 140	15
141 – 164	16
165 – 189	17
190 – 217	18
218 – 246	19

Tabulka 4 Tabelované hodnoty stupně SMOG

Co se postupu a převodní tabulky pro kratší texty týče, ze způsobu, jakým byl vzorec odvozen, je evidentní, že jsou statisticky nesmyslné, jak ostatně komentuje i autor na svých stránkách [McLaughlin, 2008].

2.2.4 Flesch Reading Ease Readability Formula

Tento vzorec je považován za jeden z nejstarších algoritmů pro měření čtivosti textů, převážně učebnic. Byl vyvinut v roce 1948 Rudolphem Fleschem, zastáncem *Hnutí pro srozumitelnou angličtinu*, který ho v tomto roce publikoval ve svém článku *A New Readability Yardstick*, otištěném v **JOURNAL OF APPLIED PSYCHOLOGY** [Flesch, 1948]. Od té doby se Fleschův vzorec stal standardním způsobem měření čtivosti textu v mnoha vládních organizacích Spojených států amerických, včetně například ministerstva obrany.

Samotný vzorec vypadá následovně:

$$re = 206,835 - 1,015 \cdot sl - 84,6 \cdot wl$$

kde

sl je průměrná délka věty ve slovech,

wl je průměrná délka slova ve slabikách.

Koeficienty byly upraveny tak, aby se výsledek nacházel na stupnici 0 až 100, přičemž 0 má nejvyšší obtížnost a hodnoty 100 dosahují nejlépe čitelné texty, což je pro uživatele snadno pochopitelný princip [Flesch, 1948].

100 – 90	Velmi snadný text
90 – 80	Snadný text
80 – 70	Poměrně snadný text
70 – 60	Standardní text
60 – 50	Poměrně složitý text
50 – 30	Složitý text
20 – 0	Velmi složitý text

Tabulka 5 Stupnice Fleschova vzorce pro čtivost

Již ne tak často se uvádí, že autor ve svém článku publikoval i vzorec pro odhadování přitažlivosti textu (*human interest*):

$$hi = 3,635 \cdot pw + 0,314 \cdot ps$$

kde

pw je procento „osobních slov“, tj.

- (a) všechna zájmena v první, druhé i třetí osobě kromě zájmen středního rodu a zájmen odkazujících na věci místo lidí,
- (b) všechna podstatná jména v jednotném i množném čísle, která mají ženský nebo mužský rod, ale ne ta, která mohou nabývat obou významů (např. *partner*),
- (c) skupinová slova *people* a *folks* (lidé);

ps je procento „osobních vět“, tj.

- (a) přímá řeč v uvozovkách nebo jinak uvozená, včetně spojení „řekl“ apod.,
- (b) otázky, příkazy, žádosti a jiné věty přímo adresované čtenáři,
- (c) rozkazovací věty,
- (d) gramatické neúplné věty, jejichž význam je třeba odvodit z kontextu.

Stupnice je rozložena obdobně:

100 – 60	Dramatický text
60 – 40	Velmi zajímavý text
40 – 20	Zajímavý text
20 – 10	Nepříliš zajímavý text
10 – 0	Nudný text

Tabulka 6 Stupnice Fleschova vzorce pro přitažlivost

Instrukce popsal Flesch (na rozdíl od ostatních) velmi detailně [Flesch, 1948].

1. Pokud nechcete testovat celý text, vyberte vzorky. Vyberte dostatečné množství, aby byl test vypovídající (např. 3 – 5 vzorků pro článek, 25 – 30 pro knihu). Nesnažte se vybírat dobré nebo typické úseky, postupujte čistě systematicky — například vezměte každý třetí odstavec na každé druhé stránce. Každý úryvek by měl začínat na začátku odstavce.
2. Spočítejte slova v textu. V případě vzorků vezměte z každého maximálně 100 slov. Stažené tvary i slova se spojovníky se počítají jako jedno slovo. Za slova se počítají i číslice nebo samostatná písmena.
3. Spočítejte slabiky ve stoslovných vzorcích, nebo v případě celého textu spočítejte počet slabik ve 100 slovech. V případě pochybností o slabikách použijte jakýkoliv lepší slovník. U čísel a symbolů počítejte slabiky tak, jako by se četly nahlas.
4. Určete průměrnou délku věty ve slovech tak, že počet slov vydělíte počtem vět. Vzorky zaokrouhlete na nejbližší celé věty. Při počítání vět se řiďte spíše myšlenkovými celky než interpunkcí, ale nerozdělujte věty spojené spojkami *a* a *ale*. U vzorků určete celkový průměr.
5. Spočtete „osobní slova“ a „osobní věty“ (viz výše) a dosadte do vzorců.

V roce 1976 americké námořnictvo vzorec upravilo pro určení ročníku, ve kterém bude čtenář schopen daný text přečíst bez obtíží. Tento vzorec se nazývá podle jeho tvůrců, Rudolpha Flesche a Johna P. Kincaida – *Flesch-Kincaid Readability Grade Level Score*. Rovnice vypadá takto:

$$re_2 = 0,39 \cdot sl + 11,8 \cdot wl - 15,59$$

Pracuje tedy se stejnými veličinami jako původní vzorec, pouze konstanty jsou odlišné. Hodnocení výsledku je velice snadné — výsledek se rovná ročníku předpokládaného čtenáře.

2.3 Vzorce pro jiné jazyky

Přehled nejen o vzorcích, ale i uskutečněných výzkumech určování čtivosti v dalších jazycích a jejich výsledcích podává vyčerpávajícím způsobem publikace *Readability: its past, present and future* editorů Zakaluka a Samuelse. Bylo by zbytečné zde všechny vzorce znovu uvádět, je ale zajímavé srovnat způsoby, jakými k problematice přistupují.

Nejjednodušší způsob je vzít již existující vzorec pro jiný jazyk a snažit se jej přizpůsobit pro nový jazyk. Tyto pokusy proběhly např. v holandštině, němčině či francouzštině, a to s *Fleschovým vzorcem*, případně s *Fryovým grafem*. Jelikož se jedná o jazyky s obecně delšími slovy než v angličtině, bylo třeba koeficienty resp. osy upravit.

Většina uvedených výzkumů nepřinášela žádné nové nápady — mezi nejpoužívanější kritéria patří průměrná délka věty (v počtech slov nebo slabik) a průměrná délka slov (ve slabikách nebo písmenech). *Läsbarhetsindex* (LIX), který vzešel z výzkumu Carl-Huga Björnssena ve Švédsku, penalizuje dlouhá slova (tj. delší než šest znaků). Na základě tohoto výzkumu byla zřízena oficiální dánská LIX komise, která vzorec přijala i pro dánštinu.

V případě holandštiny z dlouhých slov (delší než tři slabiky) autoři (Zondervan, van Steen a Gunneweg) vynechávají 35 nejpoužívanějších dlouhých slov. Použití seznamu slov je rovněž oblíbeným ukazatelem (francouzština De Landsheere, ruština Rock a další). Jedná se většinou o uznávané slovníky národních autorů, ale např. v případě ruštiny byl seznam sestaven ze slovíček, která se objevila alespoň v polovině učebnic ruštiny používaných v USA. Během výzkumu se však ukázalo, že získávání takových slovíček je těžší, než se čekalo. Ani ve zmíněném případě francouzského jazyka nedopadl výzkum příliš dobře, neboť syntaktická část vzorce kladla hlavní váhu na způsob, jak interpunkce rozděluje text, v čemž ovšem autoři vykazovali vysokou rozdílnost. A konečně, jeden z výzkumů pro německý jazyk (Bamberger, Vaněček) pracoval se seznamem 1 000 nejčastěji psaných slov 10letými dětmi.

Co se syntaktických kritérií týče, v jazycích, kde hrají slovesa specifickou roli, vzniklo několik vzorců beroucích v úvahu pomocná slovesa, nebo které podle sloves dělí věty (holandština, němčina). Za zmínku ještě stojí výzkum Richaudeaua pro francouzštinu, jehož vzorec penalizuje věty, pokud se mezi slovesem a předmětem nachází více než 10 slov, a naopak větu zvýhodňuje, pokud se jedná o větu potvrzovací (např. „*To je důvod, proč...*“). Mezi sémantické prediktory pak patří i procento neznámých slov, které rozhoduje, zda čtenář pochopí nebo nepochopí význam textu.

Jiné potíže nastávají v jazycích, které jsou od našeho více vzdáleny. Hindština sice ještě patří mezi indoevropské jazyky, používá však slabičné písmo, takže dělení na víceslabičná slova nemá význam. Proto ve výzkumu (Bhagoliwal), který se snažil přizpůsobit některý z existujících vzorců pro hindštinu, patřily vzorce pracující se slabikami (např. Flesch, Gunning) mezi ty méně přesné.

Vietnamština, která silně využívá diakritických, tonálních a jiných znamének, tyto společně s pomlčkou počítá ve vzorci za samostatná písmena, alespoň v případě uvedeného výzkumu (Ngueyn, Henkin). Není překvapením, že u logografických písem, jakými jsou např. čínština nebo korejština, se za hlavní jednotku měření čtivosti považují slova a jejich obtížnost dle frekvenčních seznamů (Yang, Park), případně počet tahů v jednotlivých logogramech. V Koreji je určování obtížnosti jazykových učebnic a učebnic společenských věd vyžadováno tamním ministerstvem školství.

Zmíněný výzkum v korejském jazyce je jediný, který k výpočtu obtížnosti používá pokročilých statistických metod. Všechny ostatní, pokud nepřevzaly *Fryův graf*, jsou zredukovány na lineární vzorec. Výjimku tvoří ovšem vzorec Rodrigueze Diégueze pro španělštinu z roku 1983, který nemůže v této práci chybět:

$$\begin{aligned} index = & 59,929 - 0,098x_1 - 0,321x_2 + 4,428 \log x_4 + 0,108x_{11} + 0,200x_{12} \\ & - 7,079 \log x_{16} - 25,816 \log x_{21} - 0,007x_{22}^2 - 0,012x_{25} \\ & - 0,126x_{27} - 20,420x_{28} + 5,502x_{31} \end{aligned}$$

kde

x_1 je počet čárek,

x_2 je počet středníků,

x_4 je počet čárek a nových odstavců,

x_{11} je počet vlastních jmen,

x_{12} je počet číslovek,

x_{16} je průměrný počet slov ve větě,

x_{21} je míra redundance,

x_{22} využívá frekvenci slov ze slovníku Garcíi Hoze,

x_{25} je počet osobních zájmen,

x_{27} je celkový počet teček,

x_{28} je odchylka rozložení počtu písmen ve slovech,

x_{31} je střední hodnota počtu písmen ve slovech + odchylka

Nutno podotknout, že všechny výzkumy uvedené v publikaci jsou datovány mezi lety 1930 a 1985, tedy z doby, kdy podobné výpočty probíhaly ručně. [Zakaluk, 1988]

3. Použití v praxi

3.1 Ukázka výpočtu vybraných vzorců

V této části je na úryvcích textu vysvětlen výpočet *FOG Indexu* a *Mistikova vzorce* pro český jazyk a *SMOGu* a vzorce *Flesch Reading Ease* pro anglický jazyk a znázorněna práce s *Fryovým grafem*.

3.1.1 Ruční výpočet pro český jazyk

Následující úryvek pochází z pohádky *Kterak slavný Sidney Hall kouzelníka chytil* ze sbírky **DEVATERO POHÁDEK** Karla Čapka a má přesně sto slov.

„Pak přišla depeše ze San Franciska v Americe: “Mám rýmu. Jinak vše v pořádku. Uchystejte hrušky. Sidney Hall.” Třicátého devátého dne po odjezdu došel konečně telegram z Amsterodamu v Holandsku: “Přijedu zítra večer v 7 hodin 15 minut. Uchystejte hrušky. Nejraději máslovky. Sidney Hall.” Čtyřicátého dne v 7 hodin 15 minut večer přirachotil vlak na nádraží. Z vlaku vyskočil pan Sidney Hall a za ním sestoupil kouzelník, vážný, bledý a s očima sklopenýma. Všichni detektivové čekali na nádraží a divili se tuze, že kouzelník není ani spoután. Než Sidney Hall jim jen pokynul rukou a řekl: “Čekejte mne, hoši, dnes [...]“

FOG Index

Vzorec pro výpočet *FOG Indexu*:

$$FOG Index_{CDV} = 5 + 0,4 \cdot (A + L)$$

kde

L je počet tří a víceslabičných slov, v ukázce 31 (vyznačena podtržením);

A je průměrný počet slov ve větě. Na sto slov připadlo 15 vět, $\frac{100}{15} \doteq 6,667$.

Tedy

$$FOG Index_{CDV} \doteq 5 + 0,4 \cdot (6,7 + 31) = 20,08.$$

Hodnota se nachází na samém počátku stupnice uvedené v kapitole 2.1.1 (nejjednodušší texty).

Mistríkův vzorec

Doplněním do *Mistríkova vzorce*

$$R = 50 - S \cdot V \cdot \frac{L}{N}$$

kde

N počet všech slov ve vzorku, tedy 100;

L počet různých slov ve vzorku, v ukázce 75 (vynechaná slova vyznačena řezem);

V je průměrný počet slov ve větě = $A = \frac{100}{15}$;

S je průměrný počet slabik ve slově, v ukázce je celkem 200 slabik⁶, tedy $S = \frac{200}{100} = 2$;

lze spočítat obtížnost

$$R = 50 - 2 \cdot \frac{100}{15} \cdot \frac{75}{100} = 50 - 2 \cdot 5 = 40.$$

I dle Mistríkovy stupnice se jedná o lehce srozumitelný text, i když nejjednodušší texty podle této stupnice dosahují ještě trochu vyšších hodnot.

3.1.2 Ruční výpočet pro anglický jazyk

Pro demonstraci výpočtu vzorců pro anglický jazyk byla vybrána ukázka z díla Agathy Christie **THE MYSTERIOUS AFFAIR AT STYLES**, která má 175 slov:

I am afraid I showed my surprise rather plainly. Mrs. Cavendish, who had married John's father when he was a widower with two sons, had been a handsome woman of middle-age as I remembered her. She certainly could not be a day less than seventy now. I recalled her as an energetic, autocratic personality, somewhat inclined to charitable and social notoriety, with a fondness for opening bazaars and playing the Lady Bountiful. She was a most generous woman, and possessed a considerable fortune of her own.

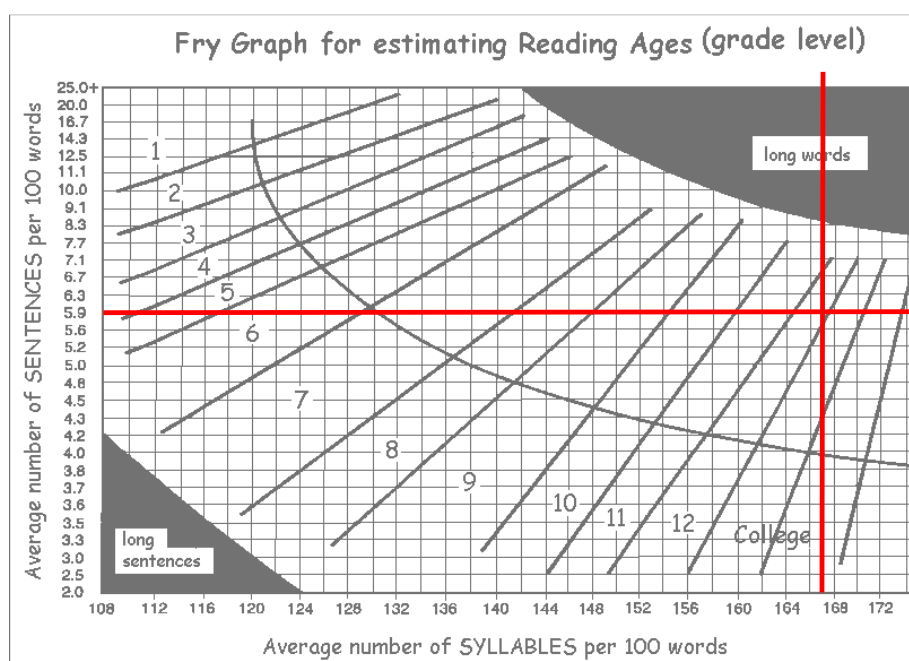
Their country-place, Styles Court, had been purchased by Mr. Cavendish early in their married life. He had been completely under his wife's ascendancy, so much so that, on dying, he left the place to her for her lifetime,

⁶ Mistrík uvádí, že skutečnost, zda se neslabičné předložky počítají či nikoliv, má zanedbatelný vliv na výsledek, u kterého je potřeba připustit odchylku ± 10 %. Sám je počítá jako slova jednoslabičná. [Mistrík, 1968]

as well as the larger part of his income; an arrangement that was distinctly unfair to his two sons. Their step-mother, however, had always been most generous to them; indeed, they were so young at the time of their father's remarriage that they always thought of her as their own mother.

Fryův graf

Pro práci s *Fryovým grafem* je potřeba pouze 100 slov (v ukázce vyznačena kurzívou), ve kterých se spočítají věty (5,9) a slabiky (167). Zjištěná čísla se vynesou do grafu a výsledný průsečík určuje počet let vzdělání předpokládaného čtenáře nutný pro bezproblémové čtení daného textu (viz Obrázek 3).



Obrázek 3 Ukázka použití Fryova grafu

Z grafu je patrné, že pro pohodlné přečtení vybraného úryvku je potřeba minimálně 12 let vzdělání, zejména kvůli vysokému počtu dlouhých slov ze začátku úryvku.

SMOG

K výpočtu *SMOGu* stačí spočítat tři- a víceslabičná slova v deseti větách textu (celý úryvek, slova vyznačena podtržením), odmocnit je a přičíst 3. Výsledek určuje počet let vzdělání předpokládaného čtenáře:

$$g = 3 + \sqrt{p} = 3 + \sqrt{25} = 8$$

Flesch Reading Ease Readability Formula

Fleschův vzorec pracuje se dvěma veličinami – průměrnou délkou vět ve slovech (wl) a průměrnou délkou slov ve slabikách (sl). Úryvek sestává z 10 větných celků, 175 slov a 270 slabik, tedy $sl = \frac{270}{175} = 1,54$ a $wl = \frac{175}{10} = 17,5$.

Tím pádem

$$\begin{aligned} re &= 206,835 - 1,015 \cdot sl - 84,6 \cdot wl \\ &= 206,835 - 1,015 \cdot 17,5 - 84,6 \cdot 1,54 \\ &= 206,835 - 17,763 - 130,284 \\ &= 58,788. \end{aligned}$$

Výsledek odpovídá obtížnosti na hranici Standard (standardní) a Fairly Difficult (poměrně složité), což zhruba odpovídá vypočtenému *SMOGu*. Oba tyto výpočty se provádějí z celého úryvku, kdežto *Fryův graf* pracuje jen se začátkem textu, který je obtížnější.

3.1.3 Počítačové programy pro anglický jazyk

Většina vzorců pro měření čtivosti byla určena pro ruční zpracování ať už knihovníky, autory nebo nakladateli. Z toho také pochází umělá omezení na délky úryvků a případná zjednodušení vzorců.

V současnosti lze měření obtížnosti textu provádět i pomocí počítačových programů, které si poradí s celými texty. Na mnoha webových stránkách lze počítat obtížnost podle různých vzorců i bez nutnosti cokoli instalovat do počítače. Nejobsáhlejší je v tomto směru server <http://www.readabilityformulas.com/>, který zdarma nabízí měření obtížnosti textu podle 7 vzorců, Fryova a Raygorova grafu, Dale-Challových metody využívající slovník těžkých slov a dalších dvou výpočtů specializovaných pro určité cílové skupiny. Autoři nabízejí i program s podobným repertoárem ke stažení a instalaci, za \$60 pro jednotlivce a za \$420 pro celé univerzity či firmy.

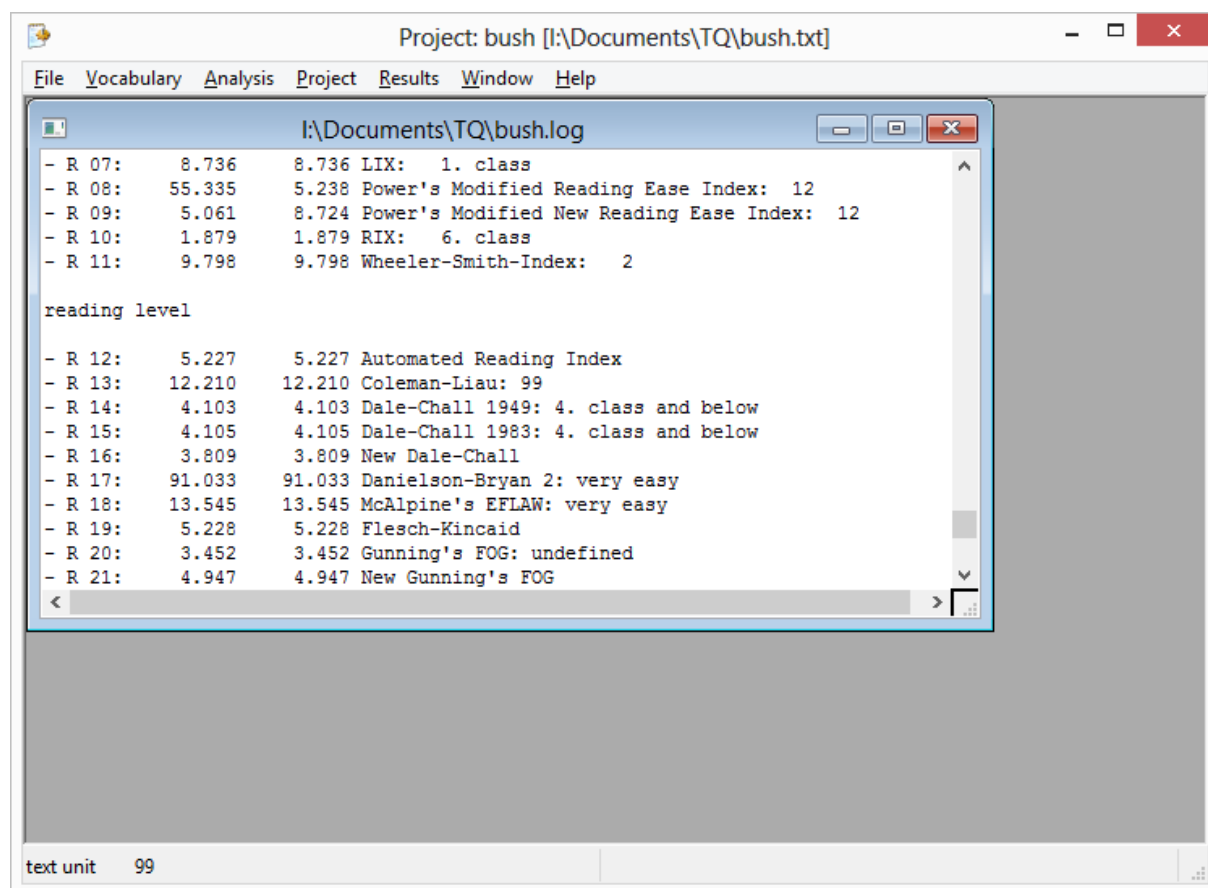
TextQuest

Každý český čtenář zajímající se o problematiku čtivosti a obtížnosti textu narazí na práce PhDr. Zdeňka Jonáka z Výzkumného ústavu pedagogického MŠMT. V jednom ze svých článků představuje program *TextQuest* německé firmy Social Science Consulting [Jonák, 2000].

Program *TextQuest*, původně INTEXT, byl vyvinut v osmdesátých letech na univerzitě v Münsteru. Jedná se o jednu z mála aplikací na obsahovou analýzu textu, bohužel její

uživatelské rozhraní ani způsob ovládání se od první verze nijak nezměnily a působí nemalé potíže i s použitím přiloženého návodu. Na pracnost postupu poukazuje i autor článku.

Co se měření obtížnosti textu týče, *TextQuest* nabízí výpočet dle 71 vzorců pro měření čtivosti, z čehož 36 je pro angličtinu, 21 pro němčinu, 7 pro španělštinu, 3 pro nizozemštinu, 2 pro francouzštinu a po jednom pro dánštinu a švédštinu. Stává se tak softwarem s nejvyšším počtem vzorců pro výpočet čtivosti textu.



Obrázek 4 Ukázka výstupu měření čtivosti v aplikaci *TextQuest*

Software ovšem stojí 800 € pro jednotlivce a 6 000 € pro univerzity. Na stránkách <http://www.textquest.de/> je k dispozici zkušební verze zdarma, která má omezení na délku zpracovávaného textu.

K získání výsledků měření čtivosti jsou zapotřebí následující kroky, které nejsou nikde takto uvedeny:

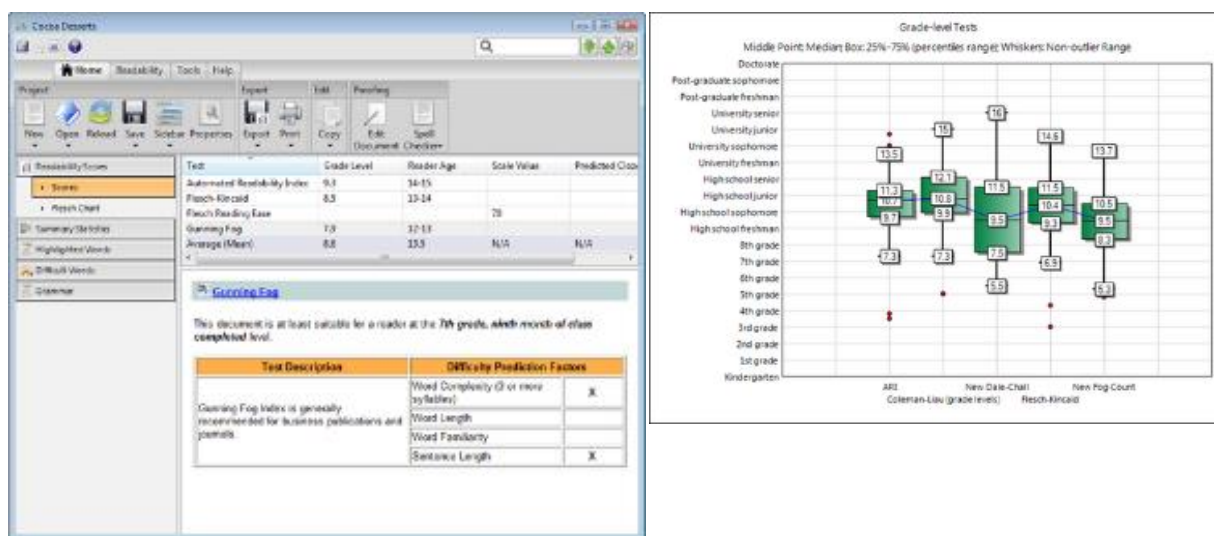
1. Text k analýze je potřeba mít připraven v textovém souboru bez formátování.
2. Po spuštění *TextQuestu* je potřeba založit projekt. V menu Project zvolte Project name a otevřete textový soubor.

3. Dalším krokem je sestavení systémového souboru. Při založení projektu bude tato možnost nabídnuta automaticky nebo ji najdete v menu File | System file. V okně Generate system file změňte položku Format of the input file na hodnotu sentence format. Potvrďte a případná varování během sestavování ignorujte.
4. Posledním krokem je samotná analýza. V menu Analysis zvolte Readability analysis. Zaškrtněte Process all text units, potvrďte parametry a případná varování během analýzy můžete opět ignorovat.
5. Výsledek najdete v okně s výpisem na jeho konci.

Výsledky jsou uvedeny z vzorců pro celé texty a také jednotně na stupnici od 0 (nejtěžší texty) do 100 (nejlehčí texty), takže je prakticky nelze srovnat s ručními výpočty. Uživatelé, které by zajímalo, jak program jednotlivé výpočty provádí, si mohou pořídit vysvětlující dokumentaci za 1 500 €.

Readability Studio

Prívětivějším produktem je pak *Readability Studio* od americké firmy Oleander Software,⁷ které se specializuje čistě na měření obtížnosti textu.



Obrázek 5 Ukázka výstupu měření čtivosti v aplikaci Readability Studio

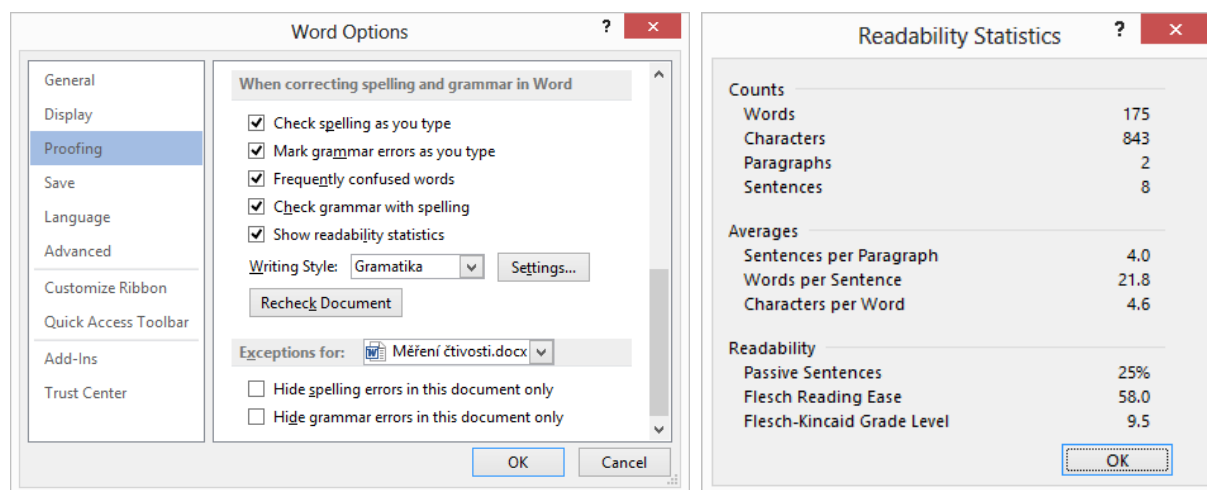
Cena je také příznivější, \$70 pro jednotlivce a \$600 pro celé univerzity a firmy, ale autoři bohužel nenabízejí žádnou možnost software si před zakoupením vyzkoušet.

Microsoft Word

A konečně každý uživatel aplikace Microsoft Word by měl vědět, že i ten disponuje funkcí na měření obtížnosti textů v angličtině již od verze 97, a to vzorcem Rudolpha Flesche zmíněným

⁷ <http://www.oleandersolutions.com/>

v kapitole 2.2.4. V nastavení kontroly pravopisu a gramatiky je nutné zaškrtnout Zobrazit statistické údaje čitelnosti a pak provést v dokumentu kontrolu pravopisu. Následně se zobrazí statistika dokumentu obsahující počet slov, písmen, odstavců a vět, průměrné počty vět v odstavci, slov ve větě a písmen ve slově, počet pasivních vět v procentech a výpočet vzorců *Flesch Reading Ease* a *Flesch-Kincaid Grade Level*.⁸



Obrázek 6 Ukázka výstupu měření čtivosti v aplikaci Microsoft Word

3.1.4 Počítačové programy pro český jazyk

Pro češtinu žádný podobný, veřejně dostupný program, který by nabízel výpočet obtížnosti textu, není autorce znám.

Centrum distančního vzdělávání

Centrum distančního vzdělávání Univerzity Palackého v Olomouci bylo v roce 2011 zrušeno stejně jako jeho webové stránky. Program na výpočet *FOG Indexu*, určený pro účastníky vzdělávacích kurzů, se již nepodařilo dohledat, ale pracovníci z Oddělení pro podporu dalšího vzdělávání zpřístupnili pro účely této práce webovou stránku, kde je možné nechat si spočítat *FOG Index* vlastního českého textu.

Maceška

Vzhledem k neexistenci dostupného programu na měření čtivosti českých textů a potřebě realizovat nové nápady představené v této práci byla ve spolupráci s Matematicko-fyzikální fakultou navržena jednoduchá aplikace — *Maceška*. Aplikace je k dispozici na kompaktním disku přiloženém k této práci a na Windows 8 ji lze spustit bez instalace přímo z disku. Na starších verzích Windows je potřeba zajistit .NET Framework 4.5 (rovněž přiložen na CD).

⁸ Více podrobností o této funkci na <http://office.microsoft.com/HP010354286.aspx>.

Maceška počítá obtížnost libovolně dlouhého textu dle *FOG Indexu* a *Mistričkova* vzorce tak, jak byly tyto vzorce popsány v kapitole 2. Ovládání je velmi snadné, čtivost se totiž počítá automaticky z vloženého textu.

Maceška — měření čtivosti textu

Text k měření: Divadelní hra (řádky začínají jménem postavy a dvojtečkou) Slova v textu:

Česká národní rada se usnesla na tomto ústavním zákoně

Preambule

Mý, občané České republiky v Čechách, na Moravě a ve Slezsku, v čase obnovy samostatného českého státu, věrni všem dobrým tradicím dávné státnosti zemí Koruny české i státnosti československé, odhodlání budovat, chránit a rozvíjet Českou republiku v duchu nedotknutelných hodnot lidské důstojnosti a svobody jako vlast rovnoprávných, svobodných občanů, kteří jsou si vědomi svých povinností vůči druhým a zodpovědnosti vůči celku, jako svobodný a demokratický stát, založený na účtě k lidským právům a na zásadách občanské společnosti, jako součást rodiny evropských a světových demokracií, odhodlání společně strážít a rozvíjet zděděné přírodní a kulturní, hmotné a duchovní bohatství, odhodlání řídit se všemi osvědčenými principy právního státu, prostřednictvím svých svobodně zvolených zástupců přijímáme tuto Ústavu České republiky

Hlava první

Dynamika textu:

Pozice	Slovo	Počet	Slabik
120	a	177	1
89	republiky	118	4
972	Článek	115	2
19	se	87	1
1002	je	86	1
2519	nebo	69	2
2165	o	67	1
30	na	59	1
99	v	56	1
4597	do	51	1
83	České	50	2
8806	vlády	44	2
1329	zákon	44	2
4997	Senátu	41	3

Vět celkem: 320
Slov celkem: 5421
Slov jedinečných: 1509
Slov tří- a více slabičných: 2312
Slabik celkem: 12968
Průměrná délka věty ve slovech: 16,94
Průměrná délka slova v slabikách: 2,39

Mistrík: 38.72
FOG Index: 28.84

Obrázek 7 Ukázka měření v aplikaci Maceška

3.2 Vzorec jako prostředek

3.2.1 Hodnocení knih

Použití vzorců k hodnocení a třídění vydaných publikací a textů se samo nabízí.

Byly osloveny následující knihovny pro zjištění, jakým způsobem třídí dětskou literaturu, a zda, případně jak, používají ke své práci vzorce pro měření čtivosti textu:

- *Městská knihovna Hlavního města Prahy* (<http://www.mlp.cz/>)
- *Městská knihovna Benešov* (<http://www.knihovna-benesov.cz/>)
- *Knihovna Jiřího Mahena v Brně* (<http://www.kjm.cz/>)
- *Knihovna města Olomouce* (<http://www.kmol.cz/>)
- *Knihovna města Ostravy* (<http://www.kmo.cz/>)
- *Jihočeská vědecká knihovna v Českých Budějovicích* (<http://www.cbvk.cz/>)
- *Knihovna města Hradce Králové* (<http://www.knihovnahk.cz/>)
- *Knihovna Václava Čtvrťka v Jičíně* (<http://www.knihovna-jicin.cz/>)
- *Městská knihovna Jihlava* (<http://www.knihovna-ji.cz/>)
- *Krajská knihovna Karlovy Vary* (<http://www.knihovnakv.cz/>)
- *Krajská vědecká knihovna Liberec* (<http://www.kvkli.cz/>)
- *Krajská knihovna v Pardubicích* (<http://www.knihovna-pardubice.cz/>)
- *Severočeská vědecká knihovna Ústí nad Labem* (<http://www.svkul.cz/>)
- *Krajská knihovna Františka Bartoše* (<http://www.kfbcz.cz/>)
- *Místní knihovna Velké Losiny* (<http://www.losiny.cz/org/knihovna>)

Také byla obeslána následující nakladatelství:

- *Albatros* (<http://www.albatrosmedia.cz/>)
- *Argo* (<http://www.agro.cz/>)
- *Modrý slon* (<http://www.modryslon.cz/>)
- *Paseka* (<http://www.paseka.cz/>)
- *Portál* (<http://www.portal.cz/>)
- *Rotag* (<http://www.rotag.cz/>)
- *Státní pedagogické nakladatelství* (<http://www.spn.cz/>)

Z celkem 15 oslovených knihoven přišlo zpátky osm odpovědí (Ostrava, Brno, Olomouc, Pardubice, Karlovy Vary, Jičín, Velké Losiny a Praha). Ani v jedné z těchto knihoven nepoužívají vzorce pro měření obtížnosti textu a ani o jejich zavedení neuvažují.

Knihy se snaží posuzovat individuálně, na základě vlastních zkušeností, obsahu nebo třeba podle stáří hlavního hrdiny.

Kam knihu zařadíme [sic!] rozhodujeme podle toho, jaký je její obsah, jakým fontem (velikost) je psána, podle množství ilustrací, struktury vět a také podle cílové skupiny, pro kterou si myslíme, že bude dostupná (snažíme si vždy představit konkrétní čtenářskou skupinu). Do hodnocení vstupuje také hodnocení dostupné na internetu (spíše u zahraničních titulů, u českých většinou nebývá ještě k dispozici... když titul zpracováváme).

Knihovna Jiřího Mahena v Brně

Řídíme se třeba tím, zda podobnou knihu už máme, podle stáří hlavního hrdiny, podle toho, co se probírá ve škole (u encyklopedií), podle televizní předlohy (večerníčky). [...] Stává se, že se kniha přeznačuje na jinou kategorii podle ohlasu čtenářů (většinou dospělí upozorňují) nebo podle toho, kdo si pro ni chodí ji číst.

Knihovna města Olomouce

Podle knihovnice z Karlových Varů je důvodem nepoužívání vzorců pro měření obtížnosti textu přílišná časová náročnost a to, že tyto vzorce nezohledňují „např. kvalitu textu a i další atributy, které má kniha pro dětské čtenáře splňovat.“ Vedoucí dětského oddělení knihovny Jiřího Mahena v Brně vidí problém s technickým provedením v kontextu časových možností na zpracování nového fondu.

Ve třech knihovnách o vzorcích nikdy neslyšeli.

Z oslovených nakladatelství zareagovalo pouze Státní pedagogické nakladatelství, které ovšem vzorce pro měření obtížnosti textu ve své práci nepoužívá, spoléhá na zkušenost redaktorů i autorů a názor recenzentů. Na žádného z nich se ale nepodařilo získat kontakt pro dohledání bližších informací.

V zahraničních knihovnách je situace překvapivě obdobná. Ani *Library of Congress* vzorec k měření svých publikací nepoužívá, nicméně na dotaz odpověděli do druhého dne a připojili rešerši na toto téma.

Na druhou stranu však vznikají, alespoň ve Spojených státech, samostatné soukromé instituce a společnosti, které vyhledávání podle obtížnosti textu nabízejí, ať už zdarma nebo za úplatu. Za všechny je možné uvést např. systém Lexile,⁹ který zavedl hodnocení obtížnosti textu i zkušenosti čtenáře vlastní stupnicí, a čtenáři tak mohou vyhledávat knihy, které odpovídají jejich schopnostem, nikoliv podle délky svého vzdělání. K hodnocení knih používají vzorec, který, stejně jako ostatní firmy v tomto oboru, drží v tajnosti (na svých stránkách uvádějí, že pracuje s délkou vět a četností slov). Pro evaluaci čtenářů zajišťuje prostřednictvím partnerů i vládních organizací písemné testy. Mezi partnery patří například ETS se zkouškami TOEFL, nebo OCLC s databází WorldCat.

Podobným přístupem by bylo možné vyřešit nezáměr knihoven o zmiňovanou problematiku i u nás. Studie ohlasu a úspěchu tohoto systému mezi čtenáři je sice nad rámec této práce, z komerčního zájmu lze však usuzovat, že existuje dostatečné množství lidí, kteří o takový systém stojí.

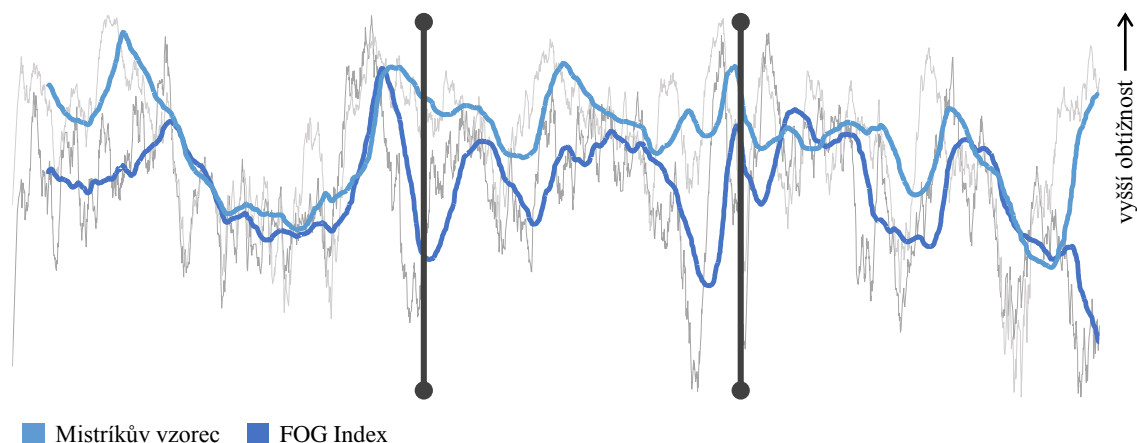
3.2.2 Dynamika textu

Automatizovaný výpočet obtížnosti textu přináší i nové možnosti analýz, které by buď vyžadovaly přečtení celé knihy, nebo dokonce nebyly vůbec možné kvůli lidskému subjektu vtaženému do děje.

Mezi takové ukazatele patří např. dynamika textu, která nám umožní sledovat vývoj obtížnosti textu v průběhu díla. Následující grafy vznikly pomocí programu Maceška výpočtem *FOG Indexu* a *Mistríkova vzorce* aplikovaného na úryvek textu začínající prvním slovem, na úryvek začínající druhým slovem, třetím slovem atd.

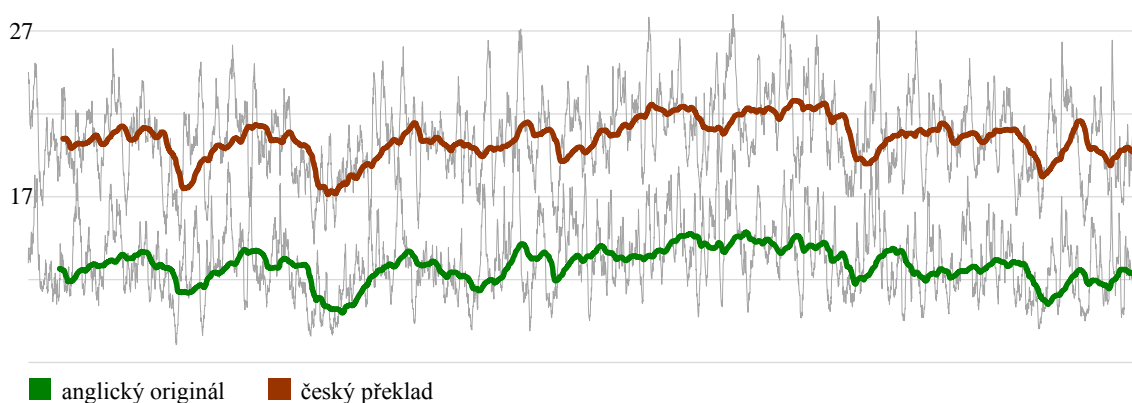
Graf 1 znázorňuje dynamiku textu divadelní hry **MATKA** od Karla Čapka. Jednotlivá dějství jsou oddělena svíslými čarami. Nejzajímavější je zvýšení obtížnosti textu v druhé části prvního dějství, shodně rozpoznané oběma vzorci, stejně tak jako pokles čtivosti na konci dějství třetího.

⁹ <http://www.lexile.com/>



Graf 1 Dynamika textu v divadelní hře Matka

Další zajímavý pohled může nabídnout srovnání této charakteristiky mezi různými jazykovými verzemi téhož díla. Následuje porovnání dynamiky textu anglické a české verze **HOBITA** Johna Ronalda Reuela Tolkiena použitím *FOG Indexu* (viz. Graf 2)



Graf 2 Dynamika textu v různých jazykových verzích jednoho díla

Tyto charakteristiky by mohly mimo jiné pomoci při porovnávání děl různých autorů.

3.3 Vzorec jako cíl

Všechny výše zmíněné vzorce byly výsledkem dlouhého výzkumu a pokusů o zjištění, které prvky textu jsou rozhodující pro určení jeho obtížnosti. Vznikaly nejen pro ověření obtížnosti textů stávajících, ale i pro určení obtížnosti textů při jejich tvorbě.

Autoři mají možnost sledovat během své práce její předpokládanou obtížnost a přizpůsobovat tak text kýženému výsledku. Americké ministerstvo obrany používá vzorce jako standardní způsob kontroly při tvorbě svých dokumentů a formulářů, armáda k zajišťování únosné

obtížnosti technických manuálů a mnohé státy pro úpravu svých legislativních dokumentů. Jako příklad lze uvést zákon státu Florida,¹⁰ který nařizuje pojišťovacímu sektoru používat ve svých podmínkách hodnocení alespoň 45 bodů dle Fleschova vzorce.¹¹

Nedávný výzkum *Readability of Wikipedia* publikovaný Teunem Lucassenem, Roaldem Dijkstrou a Janem Maartenem Schraagenem,¹² který poukazuje na vyšší obtížnost textů na anglické Wikipedii i Wikipedii psané tzv. *Simple English*, než by podle autorů odpovídalo úrovni zamýšleného čtenáře v obou případech, zahájil projekt na jejich nápravu. Autoři předpokládali pro „normální“ Wikipedii úroveň 60, tedy spodní hranici úrovně Standard podle Fleschova vzorce, který použili pro testování. Pro *Simple English* Wikipedii zvolili úroveň 80, čili spodní hranici úrovně Easy. V tomto výzkumu bylo Fleschově testu podrobena (po vyřazení článků označených jako *seznam*, *přesměrování* a *stránka pro odstranění dvojmyslů*) 88 % článků z anglické Wikipedie a 85 % textů ze *Simple English* Wikipedie. V anglické Wikipedii mělo přes tři čtvrtiny testovaných článků vyšší než požadovaný stupeň obtížnosti a zhruba 7 % článků má na Fleschově stupnici hodnotu 30 a méně (spodní hranice úrovně Difficult). Přestože autoři upozorňují na některé nedostatky, které se ve výzkumu vyskytují (*Simple English* Wikipedie není ještě plně vyvinutý projekt, což se projevuje nejenom malým počtem článků, ale i jejich délkou; pro testování byl použit pouze jeden automatizovaný vzorec pro měření obtížnosti textu; výsledky všech vzorců pro měření čtivosti by měly být interpretovány s jistou opatrností), poukazují na nutnost opravit a zjednodušit především nejčtenější články na Wikipedii tak, aby se opět zaměřily na průměrného čtenáře. Autoři vytvořili i webovou stránku,¹³ na které lze ověřit obtížnost jakéhokoliv článku na Wikipedii, případně i vlastního textu, pomocí Fleschova vzorce. Tento projekt je výsledkem spolupráce Univerzity v Twente, Babbletics a Norské organizace pro aplikovaný výzkum TNO. [Lucassen, 2012]

Od používání vzorců pro měření čtivosti textu se začalo v Americe upouštět zejména z důvodu jejich nespolehlivosti a toho, že ničí styl psaní, což tvrdili jejich kritici. Jak vysvětloval již Robert Gunning, vzorce pro měření čtivosti nebyly nikdy určeny pro to, aby se podle nich psalo. To uvádí i Judith Bogert z Pennsylvánské univerzity ve svém článku *In Defense of the FOG Index* [Bogert, 1985]. Dle jejího názoru není na vině vzorec, ale jeho zneužívání. Cílem totiž není slepě aplikovat *FOG Index* na napsaný text a pak zkrátit věty a smazat tří- a víceslabičná

¹⁰ Hlava XXXVII, kapitola 627.4145 Readability language in insurance policies.

¹¹ Způsoby použití ve Spojených státech převzaty z poznámek k výsledkům testu čtivosti na <http://www.readabilityformulas.com/freetests/six-readability-formulas.php>.

¹² Publikovaný 3. září 2012.

¹³ <http://www.readabilityofwikipedia.com/>

slova. Tím se čtivost textu sice možná zlepší ve výsledku vzorce, ale ne ve skutečnosti. Převážně studenti by se měli prací s *FOG Indexem* poučit, v čem spočívá chyba v jejich psaní tak, aby ji mohli odstranit. *FOG Index* by tedy při psaní měl být představen jako pomůcka, ne jako pravidlo.

3.4 Spolehlivost

Zajímavým problémem, který v tomto oboru zůstává opomenut, je měření spolehlivosti vzorců. Vášnivých diskusí na toto téma proběhlo mnoho, ale jak poznat, zda navrhovaný vzorec opravdu funguje? A odpovídají vůbec hodnoty skutečnosti?

3.4.1 Vzorové texty

K zodpovězení těchto otázek je třeba změřit texty, jejichž obtížnost byla stanovena odlišným, pokud možno objektivním způsobem. Existuje i jiný obor, který pracuje s obtížností textu, a tím je studium poruch čtení — dyslexie. Dětský psycholog Zdeněk Matějček¹⁴ se zabýval dyslexií a tvorbou diagnostických nástrojů v devadesátých letech. Při vyšetření čtení se sledují základní charakteristiky čtenářského výkonu (rychlost, správnost, technika čtení a porozumění) na standardizovaných, různě obtížných textech [Rittichová, 2006].

Ke standardizaci bylo vybráno sedm textů: *Alena, Kvočna a kocour, Zajíček, Kapr, Sýkorky, Krtek a Jedle a dub*, které jsou — společně s návodem na hodnocení a interpretaci — k dispozici klinickým psychologům, psychiatrům a speciálním pedagogům. [Rittichová, 2006].

Pracovnice Pražské pedagogicko-psychologické poradny byly natolik laskavé, že zapůjčily několik těchto textů pro změření jejich čtivosti vzorci. Jelikož jsou texty stále používány k diagnostice dětí, není možné je běžně publikovat, a v práci jsou uvedeny jen krátké ukázky.

Zapůjčeny byly tři texty: *Jedle a dub*, který se řadí mezi ty nejlépeší, *Kapr*, který je o něco těžší, a *Krtek*, který je ze všech textů nejtěžší. Výsledky výpočtu měření jsou uvedeny v tabulce 7.

	FOG Index	Mistríkův vzorec
Jedle a dub	18,82	33,81
Kapr	17,02	38,32
Krtek	17,7	40,12

Tabulka 7 Výsledky měření čtivosti na standardizovaných textech

¹⁴ prof. PhDr. Zdeněk Matějček, CSc.

Dle *FOG Indexu* jsou všechny tři texty snazší než nejsnazší rozsah na stupnici.

Dle *Mistríkova vzorce* spadají všechny texty do rozsahu 30—40 (druhý nejsnazší).

Ještě překvapivější je, že výsledky nesouhlasí s obtížností, na kterou byly texty standardizovány. V případě *Mistríkova vzorce* dokonce poukazují na opačné vztahy obtížnosti.

Přitom čtivost textů je rozdílná a zřejmá i z krátkých ukázek:

Jedle a dub: „[...] Tebe od mládí ošetřovali ve školkách, potom v pasekách o tebe pečovali, vsadili tě do dobré půdy, proto jsi vzrostlá. [...]“

Krtek: „[...] Pověřuji psa, aby každého, kdo si bude naříkat na krutý život, přivedl ke mně. Za nějaký čas před podzemním obydlím krtkovým zaštěkal pes. [...]“

Existují ale i jiné texty, se kterými mají vzorce potíže, a nemusí být zrovna sestavovány zkušeným psychologem.

„Žil kdys kdes chlap. Dá se říct též muž či kmán, jak kdo chce. Já ho zvu chlap. Nu, a on krad, ten chlap. Krad, co moh a kde moh. Zvlášt', když měl hlad. Byl den jak květ, pták pěl, klas zrál a nad vším jas. Jen hvozd se tměl, tam, kde spal chlap. Spal den a noc, neb pil jak Dán. Když jím třás chlad, vstal a chtěl jít dál, leč měl hlad. [...]“

Chlap, děd, vnuk, pes a hrob je jednoslabičná pohádka Jana Wericha o sto větách. Její obtížnost dle *Mistríkova vzorce* je 47,29, tedy téměř nejnižší možná, a *FOG Index* je 7,76, neboli hluboko pod nejsnazší rozsah této stupnice. Přesto lze přečtením snadno zjistit, že tato pohádka nepatří mezi nejjednodušší texty vůbec, jak se snaží použité vzorce naznačit.

3.4.2 Terénní výzkum

Čtivost textu by mohly ovlivňovat i jiné, těžko měřitelné faktory, například skutečnost, že čtenář předlohu zná. Na zjištění, zda předchozí znalost knihy či filmové adaptace ulehčí dětem čtení v porovnání s ukázkou, jejichž předlohu neznají, byl proveden výzkum na 23 žácích 4. třídy ZŠ Lupáčova v Praze a na 16 žácích 2. třídy ZŠ T. Stolzové v Kostelci nad Labem. Dětem byly předloženy dvě ukázky z textů o stejné délce (viz Příloha A). První ukázka byla z pohádky *Kterak král kočku kupoval* z knihy **DEVATERO POHÁDEK** od Karla Čapka, druhá ukázka byla z knihy **HARRY POTTER A KÁMEN MUDRCŮ** od Joanne Kathleen Rowlingové.

Oba texty byly podle vzorce pro měření obtížnosti textu *FOG Index* ohodnoceny stupněm 18, tedy jako materiál velmi jednoduchý ke čtení, s nízkým výskytem dlouhých slov a krátkými větami.

Žákům byly předloženy dotazy, zda

- a) již dříve četli některé z děl, ze kterých byly ukázky,
- b) viděli filmové zpracování Harryho Pottera.

Dále se měli rozhodnout

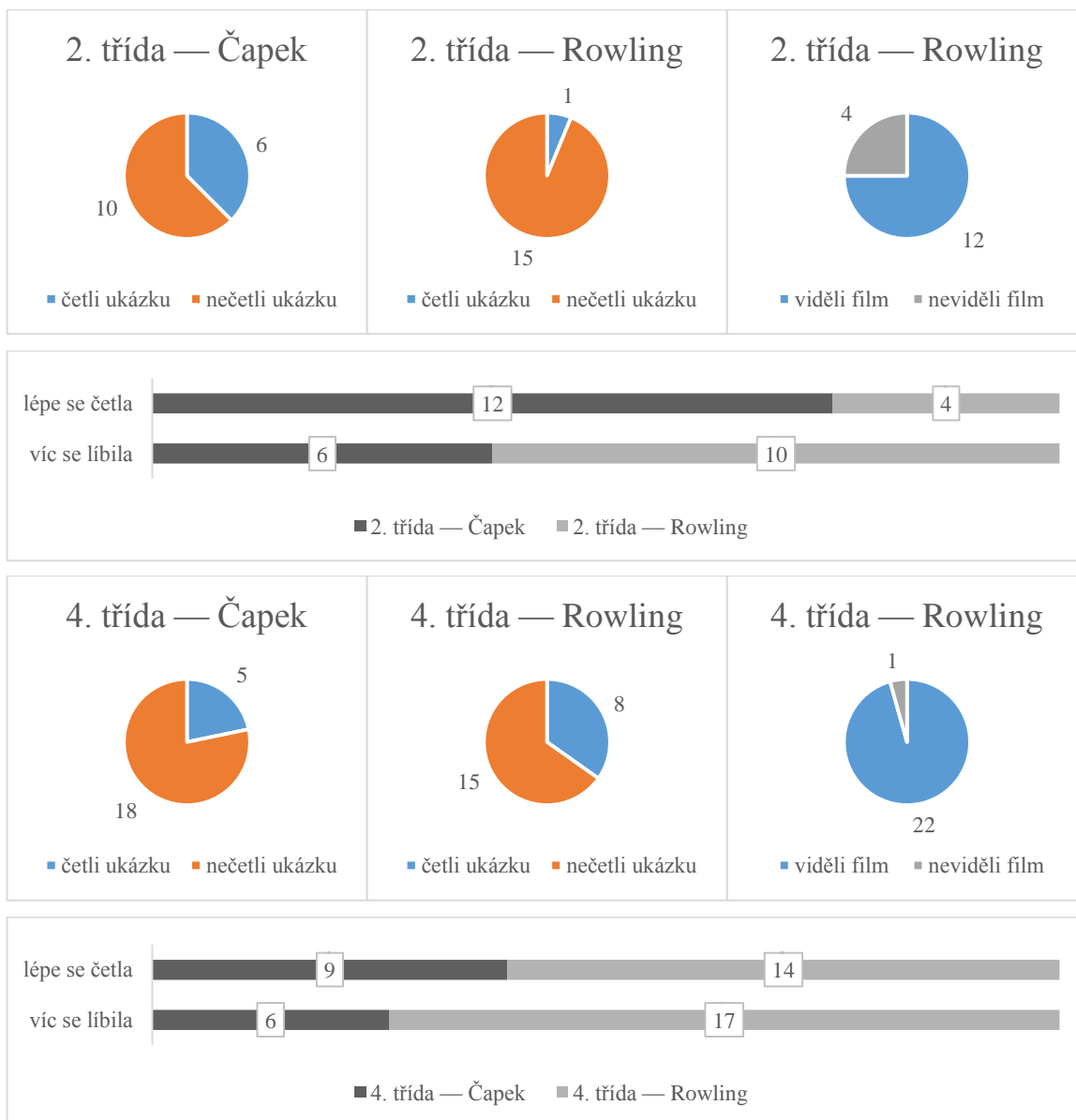
- a) která z ukázek se jim četla lépe,
- b) která z ukázek se jim líbila více,

a svá rozhodnutí zdůvodnit.

Výsledky odpovědí viz Graf 3. Ve čtvrté třídě se $\frac{2}{3}$ žáků lépe četla ukázka, od které viděli filmovou předlohu. Naopak ve druhé třídě se $\frac{3}{4}$ žáků lépe četla pohádka, přestože jich většina filmovou předlohu Harryho Pottera také viděla. Devět z dětí jako důvod uvedlo cizí slova. Tři ze čtyř dětí, kterým se četl lépe Harry Potter, svou volbu zdůvodnily tím, že ukázku už znají.

Většina žáků 4. třídy, kteří se rozhodli pro Harryho Pottera, napsala jako důvod to, že již viděli filmové zpracování.

Ze získaných odpovědí plyne, že pokud čtenář viděl filmovou předlohu, je o 6,8 % větší pravděpodobnost, že se mu bude to samé dílo v tištěné podobě číst lépe, než jiný text se stejnou naměřenou obtížností.



Graf 3 Odpovědi terénního výzkumu

4. Informační podpora tématu

4.1 Organizace zabývající se čtením

Čtením ve smyslu gramotnosti, srozumitelnosti textu apod. se ve světě zabývá mnoho národních i mezinárodních organizací. Níže je uveden stručný přehled těch nejdůležitějších:

Plain Language Action and Information (USA)

Plain Language Action and Information Network (PLAIN) je skupina federálních zaměstnanců napříč mnoha agenturami se širokým spektrem specializací, která podporuje používání jasného jazyka při psaní vládních dokumentů. Neformálně se schází od 90. let 20. století. Jejím cílem je prosadit používání jasného jazyka ve veškeré vládní komunikaci. Nabízí také školení pro vládní agentury či pomoc při editování dokumentů.

Center for Plain Language (USA)

Center for Plain Language je organizace, jejímž cílem je, aby vládní a obchodní dokumenty byly jasné a srozumitelné. Podle webových stránek centrum podporuje všechny, kteří jasný jazyk používají, učí technikám jasného jazyka ty, kteří by ho používat měli a vybízí veřejnost, aby požadovala, aby všechny dokumenty, které dostává, byly napsané v jasném jazyce. Kromě toho také například provádí a sdílí výzkumy, které určují pokyny, jak používat jasný jazyk a nejlepší techniky jeho učení. Centrum také každoročně uděluje ceny ClearMark, tedy ocenění za nejlepší text v jasném jazyce a WonderMark za nejnesrozumitelnější text.

Plain Language Association International (Kanada)

Plain Language Association International (PLAIN) je nezisková organizace založená v roce 1993, sestává z organizací, profesionálů a zastánců jasného jazyka. Jejím cílem je zvyšovat povědomost veřejnosti o jasném jazyce, pomáhat lidem pochopit a používat při psaní zásady jasného jazyka, podporovat používání jasného jazyka v Kanadě i v zahraničí či vyvíjet standardy a pokyny pro oblast jasného jazyka.

Plain English Campaign (Velká Británie)

Plain English Campaign byla založena již v roce 1979. Tato nezávislá organizace je financována z komerčně poskytovaných služeb, jako školení a editování dokumentů. Dokumentům uděluje Crystal Mark (za poplatek £200-500 spojený s posouzením dokumentu), která je zárukou, že daný text je napsaný jasným jazykem. Tuto značku už získalo přes 21 000 dokumentů z Velké Británie, Spojených států amerických, Austrálie, Dánska, Nového Zélandu a Jižní Afriky. Posuzování textů je individuální, některé podmínky jsou však stejné: například

bezchybné použití interpunkce, využívání aktivních sloves, průměrná délka věty od 15 do 20 slov atd.

International Reading Association

IRA (International Reading Association) je mezinárodní organizace institucí i jednotlivců zaměřená na gramotnost. V současnosti má více než 60 000 členů, kteří podporují vysokou úroveň gramotnosti například zlepšováním kvality pokynů pro vyučování čtení či šířením výzkumů a informací o čtení. Za tímto účelem vydává *IRA* také knihy a časopisy, mezi nimi například i *Reading Teacher* či *Reading Research Quarterly*, které se opakovaně zabývaly čtivostí a obtížností textu (viz rešerše provedená v databázi ERIC).

International Federation of Library Associations and Institutions

IFLA (International Federation of Library Associations and Institutions) je vedoucí světová organizace zastupující zájmy knihoven, informačních služeb a jejich uživatelů. Byla založena v roce 1927 v Edinburghu a v současné době má přes 1600 členů ve více než 150 státech, mezi nimiž je i Národní knihovna České republiky. Jednou z aktivit, které *IFLA* dlouhodobě provozuje, je i podpora čtenářství a gramotnosti v knihovnách. K tomuto účelu byla zřízena Sekce pro gramotnost a čtení, která se zabývá například zvyšováním povědomí o důležitosti čtení, vývojem čtenáře či čtenářskými schopnostmi a zvyky. Tato sekce spolupracuje kromě některých dalších sekcí *IFLA* také s významnými organizacemi pro podporu čtení, jako například *IRA* (viz výše) či *IBBY* (International Board on Books for Young People).

4.2 Odborné databáze

4.2.1 ERIC

ERIC (Education Resources Information Center), tedy Informační centrum pro vzdělávací zdroje, je digitální knihovna výzkumů a informací týkajících se vzdělávání. *ERIC* byl založen roku 1966 a je sponzorován Institutem pro vzdělávání a vědu pod americkým ministerstvem školství. Tato databáze obsahuje více než 1,4 milionů bibliografických záznamů odborných časopisů, s více než 300 000 plnými texty. Kromě jednoduchého a rozšířeného vyhledávání v databázi obsahuje *ERIC* také thesaurus, tedy kontrolovaný slovník, který

umožňuje uživateli zpřesnit jeho dotazy a tím získat lepší výsledky vyhledávání. Vzhledem k tomu, že vzorce pro měření čtivosti byly vytvořeny primárně pro účely školství, byla v databázi *ERIC* provedena rešerše s deskriptory "Readability" a "Readability Formulas" s omezením zdroje na články z časopisů (Journal Articles).

Z tohoto vyhledávání vychází, že nejvíce se v oblasti vzdělávání této problematice věnují seriály uvedené v následující tabulce.

Název periodika	Záznamů	ISSN	Vydavatel	Dostupnost v ČR
Journal of Reading	66	0022-4103	International Reading Association	nedostupný
Reading Teacher	33	0034-0561	Wiley-Blackwell	dostupný na PedF Masarykovy univerzity (od roku 2006)
Reading Improvement	20	0034-0510	Project Innovation, Inc.	nedostupný
Journal of Business Communication	20	0021-9436	SAGE Publications Ltd	nedostupný
Reading Research Quarterly	19	0034-0553	Wiley-Blackwell	nedostupný
Reading Horizons	18	0034-0502	Western Michigan University	nedostupný

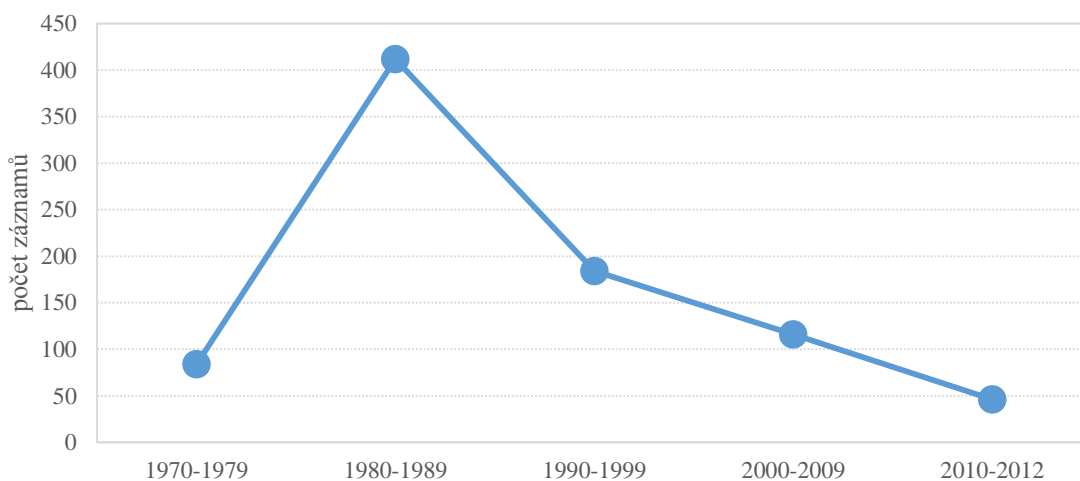
Advanced Search

Obrázek 8 Rozšířené vyhledávání databáze ERIC

Journal of Educational Psychology	13	0022-0663	American Psychological Association	dostupný v 8 knihovnách
Journal of Reading Behavior	12	0022-4111	National Reading Conference (U.S.)	nedostupný
Journal of Technical Writing and Communication	11	0047-2816	Baywood Publishing Company, Inc.	nedostupný
Journalism Quarterly	11	0196-3031	Association for Education in Journalism	dostupný ve 2 knihovnách
Reading Psychology	10	0270-2711	Taylor & Francis, Ltd	nedostupný
Visible Language	10	0022-2224	Sharon H. Poggenpohl	dostupný v Moravské galerii Brno
Reading World¹⁵	10		Taylor & Francis, Ltd	nedostupný
Social Studies	10	0037-7996	Taylor & Francis, Ltd	nedostupný

Tabulka 8 Výsledky vyhledávání v databázi ERIC

Až do 70. let 20. století se v databázi nevyskytují žádné články na toto téma, v letech 1970-1979 ERIC uchovává 84 článků, pak následuje masivní skok, kdy v letech 1980-1989 přibylo 412 článků. Poté se počty záznamů opět snižují, v letech 1990-1999 pouze 184 a 116 v letech 2000-2009. V posledních třech letech ERIC indexoval 46 článků s deskriptory "Readability" nebo "Readability formulas". Tato data znázorňuje Graf 4.



Graf 4 Vývoj počtu záznamů v databázi ERIC v čase

¹⁵ Nyní Literacy Research and Instruction – ISSN 1938-8071.

4.2.2 ProQuest

Při použití deskriptoru "Readability" bylo v ProQuestu vyhledáno 305 výsledků ve vědeckých časopisech. Nejvíce záznamů se vyskytuje v následujících seriálech:

Název periodika	Záznamů	ISSN	Vydavatel	Dostupnost v ČR
IEEE Transactions on Professional Communication	25	0361-1434	Institute of Electrical and Electronics Engineers	dostupný ve 2 knihovnách
The Journal of Business Communication	17	0021-9436	SAGE Publications Ltd	dostupný ve 2 knihovnách
Journal of Information Science	7	0165-5515	SAGE Publications Ltd	dostupný v 6 knihovnách
Accounting, Auditing & Accountability Journal	6	0951-3574	Emerald Group Publishing Limited	dostupný v CIKS VŠE Praha
Technical Communication	6	0049-3155	Society for Technical Communication	nedostupný

Tabulka 9 Výsledky vyhledávání v databázi ProQuest

4.2.3 EbscoHost

Při použití deskriptoru "Readability" v knihovnických a informačních databázích (*Computers & Applied Sciences Complete, LISTA with FT, e-Book Collection*) databázového centra EbscoHost bylo vyhledáno 66 výsledků bez duplicitních záznamů. Z toho 56 bylo v akademických periodikách.

Název periodika	Záznamů	ISSN	Vydavatel	Dostupnost v ČR
Journal of Consumer Health on the Internet	4	1539-8285	Taylor & Francis, Ltd	nedostupný
Journal of Scholarly Publishing	4	1198-9742	University of Toronto Press	nedostupný
Journal of Technology in Human Services	4	1522-8835	Haworth Press	nedostupný
Communications of the ACM	2	0001-0782	Association for Computing Machinery	dostupný ve 21 knihovnách
Computational Linguistics	2	0891-2017	The Association for Computational Linguistics	dostupný ve 2 knihovnách

Health Informatics Journal	2	1460-4582	SAGE Publications Ltd	nedostupný
-----------------------------------	---	-----------	-----------------------	------------

Tabulka 10 Výsledky vyhledávání v databázi EbscoHost

4.3 Kvalifikační práce

V rámci informační podpory tématu bakalářské práce byla provedena i rešerše v databázi vysokoškolských kvalifikačních prací <http://www.theses.cz>, která sdružuje kvalifikační práce 27 vysokých škol. Tato databáze neumožňuje vyhledávat klíčová slova, ale při kliknutí na slovo uvedené v seznamu seřazeném podle abecedy nebo podle četnosti jsou zobrazeny práce, které mají tento pojem uvedený v seznamu klíčových slov. Kromě základních slov, tedy „čtivost textu“ a „obtížnost textu“ byla na doporučení vedoucího práce pro znázornění širších souvislostí vybrána následující slova: „čitelnost textu“, „čtenářství“, „čtenářská gramotnost“ a „čtení“, z nichž poslední tři zmíněná se vyskytují v seznamu klíčových slov.

Termín	Fulltextové vyhledávání	Klíčové slovo
Čtivost textu	78 záznamů	—
Obtížnost textu	94 záznamů	—
Čitelnost textu	213 záznamů	—
Čtenářství	709 záznamů	29 záznamů
Čtenářská gramotnost	304 záznamů	22 záznamů
Čtení	884 záznamů	24 záznamů

Tabulka 11 Výsledky vyhledávání v databázi theses.cz

Bohužel tato databáze vysokoškolských prací má velmi omezené možnosti vyhledávání, což nedovoluje provedení podrobnější statistiky.

V digitálních repozitářích Univerzity Karlovy bylo provedeno vyhledávání stejných termínů jako v Theses.cz, s následujícími výsledky:

Termín	Před rokem 2006	Po roce 2006
Čtivost textu	1	24
Obtížnost textu	57	35
Čitelnost textu	15	84
Čtenářství	7	261
Čtenářská gramotnost	6	117
Čtení	162	8438

Tabulka 12 Výsledky vyhledávání v repozitářích Univerzity Karlovy

5. Závěr

V této práci zaměřené na metody měření obtížnosti textu byly prezentovány aspekty textu, jejichž vliv na obtížnost textu byl v minulosti do hloubky zkoumán. Z tohoto výzkumu vznikly představené vzorce pro měření čtivosti pro český i anglický jazyk a bylo vysvětleno a předvedeno, jak fungují.

Několik knihoven a nakladatelství bylo dotázáno ohledně používání těchto vzorců a bylo zjištěno, že využívány nejsou. Pro knihovny je to pochopitelné, z důvodu obtížnosti práce s tištěným dokumentem a časové náročnosti ručního výpočtu je počítání čtivosti každé knihy velmi obtížně realizovatelné. Mohly by ale například ve svých internetových katalozích zjišťovat od čtenářů kromě oblíbenosti titulu také hodnocení jeho čtivosti a získávat tím cennou zpětnou vazbu.

V případě nakladatelství, kde mají k dispozici elektronické verze publikací, je ovšem situace zcela jiná. Pokud by se výpočty prováděly automatizovaně, mohly by vzorce např. nakladatelstvím dětské literatury ušetřit čas strávený hodnocením zaslaných rukopisů, kdy vzorec může být prvním vodítkem, že text je pro dětské čtenáře příliš obtížný. Kdyby navíc nakladatelství uváděla na všech vydávaných knihách standardizovaný údaj o věku předpokládaných čtenářů, knihovny i knihkupectví by měly jistotu, že všechny knihy jsou hodnoceny jednotně a objektivně, namísto slepé důvěry v úsudek nebo odhad anonymního člověka v pozadí.

Samozřejmě, že stejně tak jako existují čtenáři, kteří dají přednost nezávislému hodnocení nebo se obejdou bez osobního kontaktu s pracovníky knihovny, vždy budou i čtenáři, kteří se o pomoc raději obrátí na zkušeného knihovníka.

Výzkumy měření čtivosti textu byly utlumeny před čtyřiceti lety a závratný rozvoj výpočetní techniky tento obor překvapivě příliš nepoznamenal. V práci byly zmíněny možnosti moderní výpočetní lingvistiky a uvedeny příklady, které ukazatele by se daly k automatizovanému měření využít, a rovněž byly nastíněny nové metody použití takových měření. Vzniká tak mnoho zajímavých možností k dalším výzkumům v tomto oboru.

Významným přínosem této práce je mimo jiné návrh a vytvoření programu Maceška pro měření obtížnosti českých textů pomocí *FOG Indexu* a *Mistričkova vzorce*. Tento program může značně napomoci dalšímu rozvoji zkoumání na poli automatizovaného měření obtížnosti textu.

Čtivost této práce je 15,7 podle *FOG Indexu* a 45,1 podle *Mistričkova vzorce*.

Seznam obrázků

Obrázek 1 Schematický popis zdrojů Komplexní míry obtížnosti [PRŮCHA, 2013]	15
Obrázek 2 Stupnice Fryova grafu	17
Obrázek 3 Ukázka použití Fryova grafu	29
Obrázek 4 Ukázka výstupu měření čtivosti v aplikaci TextQuest.....	31
Obrázek 5 Ukázka výstupu měření čtivosti v aplikaci Readability Studio.....	32
Obrázek 6 Ukázka výstupu měření čtivosti v aplikaci Microsoft Word.....	33
Obrázek 7 Ukázka měření v aplikaci Maceška.....	34
Obrázek 8 Rozšířené vyhledávání databáze ERIC	46

Seznam tabulek

Tabulka 1 Stupnice FOG Indexu dle CDV ÚPOL	13
Tabulka 2 Stupnice Mistríkova vzorce	14
Tabulka 3 Gunningova stupnice FOG Indexu	19
Tabulka 4 Tabelované hodnoty stupně SMOG.....	21
Tabulka 5 Stupnice Fleschova vzorce pro čtivost	22
Tabulka 6 Stupnice Fleschova vzorce pro přitažlivost	23
Tabulka 7 Výsledky měření čtivosti na standardizovaných textech.....	40
Tabulka 8 Výsledky vyhledávání v databázi ERIC	47
Tabulka 9 Výsledky vyhledávání v databázi ProQuest	48
Tabulka 10 Výsledky vyhledávání v databázi EbscoHost.....	49
Tabulka 11 Výsledky vyhledávání v databázi theses.cz.....	49
Tabulka 12 Výsledky vyhledávání v repozitářích Univerzity Karlovy	49

Seznam grafů

Graf 1 Dynamika textu v divadelní hře Matka	38
Graf 2 Dynamika textu v různých jazykových verzích jednoho díla	38
Graf 3 Odpovědi terénního výzkumu	43
Graf 4 Vývoj počtu záznamů v databázi ERIC v čase.....	47

Seznam použité literatury

About IFLA. *IFLA* [online]. Poslední změna 17. 1. 2013 [cit. 22. 1. 2013]. Dostupný z: <http://www.ifla.org/about>.

About Literacy and Reading Section. *IFLA* [online]. Poslední změna 5. 10. 2012 [cit. 22. 1. 2013]. Dostupný z: <http://www.ifla.org/about-the-literacy-and-reading-section>.

About PLAIN. *Plain Language Association International* [online]. Plain Language Association International, ©2011. [cit. 21. 1. 2013] Dostupný z: <http://www.plainlanguagenetwork.org/networkindex.html>.

About us. *Plain Language: Improving the Communication from the Federal Government to Public* [online]. [cit. 21. 1. 2013]. Dostupný z: <http://www.plainlanguage.gov/site/about.cfm>.

BOGERT, J. In Defense of the Fog Index. *Business Communication Quarterly* [online]. 1985, **48**(2), 9-12 [cit. 10. 1. 2013]. ISSN 1552-4191. Komerčně dostupný z: <http://bcq.sagepub.com/content/48/2.toc>.

Center for Plain Language [online]. Falls Church, VA [cit. 21. 1. 2013]. Dostupný z: <http://centerforplainlanguage.org/>.

ČAPEK, Karel. *Menší prózy*. 1. souborné vyd. Praha: Československý spisovatel, 1992 [cit. 16. 2. 2013]. Spisy, sv. 10.

Digitální univerzitní repozitář [online]. Praha: Univerzita Karlova v Praze, ©2007 [cit. 16. 2. 2013]. Dostupný z: <http://digitool.cuni.cz/>.

Filozofická fakulta Univerzity Karlovy v Praze. *Český národní korpus* [online]. [cit. 16. 2. 2013]. Dostupný z: <http://www.korpus.cz/>.

FITZSIMMONS, PR, MICHAEL, BD, HULLEY, JL, SCOTT, GO. A readability assessment of online Parkinson's disease information. *The Journal of the Royal College of Physicians of Edinburgh*. **40**(4), 292-296 [cit. 10. 1. 2013]. ISSN 2042-8189. Dostupný z: http://www.rcpe.ac.uk/journal/issue/journal_40_4/fitzsimmons.pdf.

FLESCHE, Rudolph. A New Readability Yardstick. In: *The Classic Readability Studies* [online]. Costa Mesa (CA): Impact Information, 2006, s. 99-111. [cit. 10. 1. 2013]. Dostupný z: www.impact-information.com/impactinfo/research/classics.pdf.

- GUNNING, Robert. *The technique of clear writing* [online]. New York: McGraw-Hill, 1968 [cit. 10. 1. 2013]. Dostupné z: <http://archive.org/details/techniqueofclear00gunn>.
- HODIS, Vladislav. Hodnocení učebnice Praktické činnosti pro 6. - 9. ročník základních škol: Provoz a údržba domácnosti. *e-Pedagogium* [online]. 2003, III [cit. 10. 1. 2013]. ISSN 1213-7499. Dostupný z: <http://epedagog.upol.cz/eped3.2003/clanek07.pdf>.
- International Reading Association* [online]. International Reading Association, ©1996-2013 [cit. 21. 1. 2013]. Dostupný z: <http://www.reading.org/>.
- JONÁK, Zdeněk. Co vše by měl (mohl) o knize vědět školní knihovník. *Daidalos: informační server pro knihovníky* [online]. [cit. 10. 1. 2013]. ISSN 1213-1482. Dostupný z: <http://daidalos.ff.cuni.cz/prilohy/skolnik/ssic07.php>.
- JONÁK, Zdeněk. TEXTQUEST: software pro obsahovou analýzu. *Ikaros* [online]. 2000, 4(5) [cit. 10. 1. 2013]. ISSN 1212-5075. Dostupný z: <http://www.ikaros.cz/textquest-software-pro-obsahovou-analyzu>.
- JONÁK, Zdeněk. Stanovení jazykové obtížnosti učebnic vzhledem k věku žáka. In: *Metodický portál RVP* [online]. 24. 5. 2005 [cit. 10. 1. 2013]. Dostupný z: <http://clanky.rvp.cz/clanek/c/ZUC/239/STANOVENI-JAZYKOVE-OBTIZNOSTI-UCEBNIC-VZHLEDEM-K-VEKU-ZAKA.html/>.
- LUCASSEN, Teun, DIJKSTRA Roald, SCHRAAGEN, Jan Maarten. Readability of Wikipedia. *First Monday* [online]. 2012, 17(9) [cit. 10. 1. 2013]. ISSN 1396-0466. Dostupný z: <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3916/3297>.
- MCLAUGHLIN, Harry G. SMOG Grading – A New Readability Formula. *Journal of Reading*. 1969, 12(8), 639-646 [cit. 10. 1. 2013]. Dostupný z: [http://www.harrymclaughlin.com/SMOG_Readability_Formula_G._Harry_McLaughlin_\(1969\).pdf](http://www.harrymclaughlin.com/SMOG_Readability_Formula_G._Harry_McLaughlin_(1969).pdf).
- MCLAUGHLIN, Harry G. *SMOG: Simple Measure of Gobbledygook* [online]. [cit. 10. 1. 2013]. Dostupný z: <http://www.harrymclaughlin.com/SMOG.htm>.
- MISTRÍK, Jozef. Meranie zrozumiteľnosti prehovoru. *Slovenská reč* [online]. 1968, 33(3), 171-178. [cit. 10. 1. 2013]. Dostupný z: <http://www.juls.savba.sk/ediela/sr/1968/index.html>.
- My Byline Media. The Gunning's Fog Index (or FOG) Readability Formula. *Readability Formulas* [online]. [cit. 10. 1. 2013]. Dostupný z: <http://www.readabilityformulas.com/gunning-fog-readability-formula.php>.

Plain English Campaign [online]. Plain English Campaign, ©2013 [cit. 21. 1. 2013].

Dostupný z: <http://www.plainenglish.co.uk/>.

PRŮCHA, Jan. *Moderní pedagogika*. 5. aktualiz. a dopl. vyd. Praha: Portál, 2013. ISBN 978-80-262-0456-5.

Repozitář závěrečných prací [online]. Praha: Univerzita Karlova v Praze. 14. 2. 2013 [cit. 16. 2. 2013]. Dostupný z: <https://is.cuni.cz/webapps/zzp/>.

RITTICHOVÁ, Jana. *Reedukace specifických poruch učení u žáků ve speciálních třídách při běžných základních školách v okrese Jihlava*. Brno, 2006. Diplomová práce, Masarykova univerzita v Brně. [cit. 10. 1. 2013]. Dostupný z:

http://is.muni.cz/th/166556/pedf_m/?lang=en.

ROWLINGOVÁ, J. K. *Harry Potter a Kámen mudrců*. 2. vydání. Praha: Albatros, 2001. ISBN 80-00-00995-1.

Univerzita Palackého v Olomouci. Výpočet obtížnosti. *Centrum distančního vzdělávání* [online]. Olomouc, ©2005 [cit. 12. 6. 2012]. Dostupný z:

http://www.cdiv.upol.cz/www/autori_obtiznost.htm.

Vysokoškolské kvalifikační práce [online]. Brno: Masarykova univerzita [cit. 26. 1. 2013].

Dostupný z: <http://www.theses.cz/>.

ZAKALUK, Beverley L., SAMUELS, Jay S. *Readability: its past, present and future*. Newark: International Reading Association, 1988. ISBN 0-87207-795-0.

Příloha A. Výzkumný dotazník

J. K. Rowlingová: Harry Potter a kámen mudrců

Četl/a jsem knihu? ANO – NE

Viděl/a jsem film? ANO – NE

„Tatínek se zbláznil, vid’?“ zeptal se Dudley sklesle tety Petunie navečer. Strýc Vernon zastavil na nábřeží, všechny je zamkl v autě a zmizel.

Začalo pršet. Po střeše vozu bubnovaly veliké kapky. Dudley se rozfňukal.

„Je pondělí,“ řekl matce. „Dnes večer dávají Velkého Humberta. Já chci někam, kde bude televize.“

Pondělí! To Harrymu něco připomnělo. Pokud bylo pondělí – a díky televizi se na Dudleyho obvykle dalo spolehnout, že ví, co je za den – pak zítra, v úterý, měl on sám mít jedenácté narozeniny. Harryho narozeniny samozřejmě nikdy nebyly žádná sláva – loni mu Dursleyovi darovali ramínko na šaty a staré ponožky po strýci Vernonovi. Ale přesto, jedenáct vám není každý den.

To už byl strýc Vernon zpátky a usmíval se. Navíc si nesl i jakýsi dlouhý, úzký balík a neodpověděl, když se ho teta Petunie zeptala, co to koupil.

„Našel jsem vynikající místočko!“ řekl. „Pojďte! Všichni ven!“

Vystoupili a zjistili, že venku je velice chladno. Strýc Vernon ukazoval na něco, co daleko v moři vypadalo jako veliký útes. Nahoře na něm byla přilepená ta nejubožejší chatrč, jakou si dokážete představit. Jedna věc byla jasná: televizi tam určitě nemají.

„Na dnešní noc předpovídají bouři!“ oznámil strýc Vernon radostně a tleskl rukama. „A tento dobrý muž bude tak laskav a půjčí nám svůj člun!“

To už se k nim přiloudal bezzubý stařík a s dost zlomyslným úšklebkem jim ukazoval starou veslici, která se kolébala v ocelově šedé vodě pod nimi.

„Nějaké zásoby jsem už opatřil,“ řekl strýc Vernon, „takže všichni na palubu!“

V loďce byla strašlivá zima. Za krk jim zatékala ledová vodní tříšť a do tváří je bodal mrazivý vítr. Připadalo jim, že to trvalo celé hodiny, než dorazili k útesu, kde je strýc Vernon, i když se smekal a klouzal, dovedl k polorozpadlému stavení.

Uvnitř to vypadalo hrozně: páchlo to tam chaluhami, vítr profukoval škvírami v dřevěných stěnách a ohniště bylo vlhké a prázdné. V chatrči byly jen dvě místnosti.

Ukázalo se, že jejich zásoby obnášejí po jednom balíčku smažených bramborových lupínků pro každého a čtyři banány. Strýc Vernon se pokusil rozdělat oheň, ale prázdné obaly od brambůrků jen čadily a zkroutily se.

„Teď by se nám pár těch dopisů docela hodilo, co říkáte?“ nadhodil rozmarně.

Byl ve velice dobré náladě. Očividně si myslel, že nikdo nemá sebemenší možnost dostat se v bouři až sem a doručit jim poštu. Harry s ním v duchu souhlasil, i když sám z toho žádnou radost neměl.

S příchodem noci se kolem nich rozběhla ohlášená bouře. Stěny chatrče zalévala tříšť ze vzdutých vln a špinavými okny lomcoval divoký vítr. Teta Petunie našla ve druhé místnosti několik zplesnivělých příkrývek a ustlala Dudleymu na pohovce prožrané od molů. Sama se uložila se strýcem Vernonem na hrbatou postel v sousední místnosti a Harrymu nezbylo nic jiného, než si najít to nejměkčí místo na podlaze a schoulit se pod tou nejtenčí a nejpotrhanější příkrývkou.

Jak noc ubíhala, bouře řádila stále zuřivěji. Harry nemohl usnout. Třásl se zimou a mlel sebou, aby našel nějakou pohodlnější polohu, a v žaludku mu kručelo hladem. Dudleyho chrápání teď přehlušovalo vzdálené burácení hromů, které začalo před půlnocí. Svítící ciferník hodinek, které visely spáči na tučném zápěstí přes okraj pohovky, Harrymu prozradil, že mu za deset minut bude jedenáct let.

Karel Čapek: Devatero pohádek

Četl/a jsem knihu? ANO – NE

„Inu, co platno,“ vzdychl král, „to už tu kočku musím koupit. Ale vy jste, babičko, jářku, taškářka.“

Tak tedy nezbylo králi, než aby vysázel na stůl stříbrné tolarý; pak vzal své mamince z hlavy ten malý, ze všech nejmenší čepeček a položil jej na tolarý. Čepeček byl tak malý, že se pod něj vešlo zrovna jen pět tolarů.

„Tumáte, babičko, svých pět tolarů a jděte spánembohem,“ řekl král a byl rád, že to pořídil tak lacino.

Ale babička vrtěla hlavou a povídá: „To jsme si, pane králi, neujednali. Vy mně máte dát tolik tolarů, co se vejde nejlepšího stříbra pod čepec vaší maminky.“

„Vždyť ale vidíte,“ bránil se král, „že se pod čepec vejde zrovna jen pět tolarů z nejlepšího stříbra.“

Babička vzala čepeček do ruky, hladila jej, točila na dlani a povídá pomalu: „Myslím, pane králi, že nejlepší stříbro na světě jsou stříbrné vlasy vaší maminky.“

Král se podíval na babičku, podíval se na maminku a řekl tiše: „Máte pravdu, babičko.“

Tu babička posadila čepeček pěkně na hlavu královy maminky, pohladila jí bílé vlasy a řekla: „A teď mi, pane králi, dáte tolik tolarů, co je stříbrných vlásků pod čepcem vaší maminky.“

Král se podivil, král se zamračil, král se konečně usmál a pravil: „Vy jste ale, babičko, tuze veliká taškářka.“

Nu ale, děti, přísaha je přísaha, a tak musel král zaplatit babičce, oč žádala. Poprosil tedy svou maminku, aby si sedla, a nařídil svému nejvyššímu účetnímu, aby jí spočítal stříbrné vlasy, které se jí vejdu pod ten čepeček. Účetní počítá, počítá, a králova maminka pěkně drží, ani se nepohne, a vtom, to víte, stáří má rádo spánek a lehko zadřímne – zkrátka králova maminka usnula.

Zatímco tak spala, počítal účetní vlasek po vlásku; a když se právě dopočítal jednoho tisíce, snad že silněji zatahal za jeden stříbrný vlas, králova maminka se probudila.

„Au,“ vykřikla, „proč mne budíte? Měla jsem živý sen; zdálo se mi, že právě překročil příští král hranice naší země.“

Babička sebou trhla. „To je divné,“ vyhrkla, „zrovna dnes se má ke mně nastěhovat můj vnouček z přespolní říše.“

Ale král ani neposlouchal a zvolal: „Odkudpak, maminko, odkudpak přijde příští král naší země? Ze kterého královského dvora?“

„To nevím,“ řekla králova maminka, „protože jste mne zrovna probudili.“

Zatím nejvyšší účetní počítal dál, a králova maminka zase usnula. Účetní počítá, počítá až do dvou tisíc, a tu se mu zase utrhla ruka a on zatahal silně za stříbrný vlas.

„Ale chaso,“ vykřikla králova maminka, „proč jste mne probudili? Právě se mi zdálo, že příštího krále sem nepřivede nikdo jiný než tahle černá kočka.“

„Jděte, maminko,“ divil se král, „kdopak jakživ slyšel, že by kočka někoho do domu přivedla?“

„Už to tak bude,“ řekla králova maminka, „ale teď mne nechte spát.“

Znovu usnula králova maminka, a znovu počítal účetní. Když se dopočítal vlasu třítisícího a posledního, zatřásla se mu ruka a on nechtě zase silně zatahal.

„Ach vy daremní,“ vykřikla králova maminka, „nenecháte starou bábu chvilku spát. Zrovna se mi zdálo, že sem příští král přijede se svým celým domem.“

„No to mně už, maminko, odpusťte,“ řekl na to král, „ale tohle asi nebude pravda. Kdopak by mohl s sebou přivést celý královský zámek?“

„Nemluv, hochu, nadarmo,“ kárala ho maminka. „Člověk nikdy neví, co se může stát.“

**Která ukázka se Ti lépe četla?
Proč?**

Harry Potter – Devatero pohádek

**Která ukázka se Ti více líbila?
Proč?**

Harry Potter – Devatero pohádek