

## METÓDY ANALÝZY TEXTOV

Text sa spravidla chápe ako napísaný, vytlačený alebo aj ináč zaznamenaný (resp. dokumentovaný) sled súvislých výpovedí alebo iných komunikátov v ich štruktúrnom kontexte (Švec, 1998). Textové dokumenty, ako zhmotnené prostriedky sociálnej komunikácie, tvoria bohatý zdroj údajov pre vedecké bádanie. V oblasti školstva a vzdelávania sa používa pestrá paleta textových dokumentov, ako napríklad:

### a) Textové dokumenty prevažne oficiálnej povahy:

- *výskumné a vývojové*: napr. výskumné správy, dizertácie, projekty, expertízy ap.; záznamy spomienkového rozprávania;
- *politické*: napr. projekty rozvoja národnej vzdelanosti, akčné programy, politické prehlásenia a výzvy ap.;
- *legislatívne*: napr. zákony, vyhlášky, nariadenia; morálny kódex učiteľskej profesie, listina detských práv ap.;
- *štatistické*: štatistické ročenky, výkazy školských aktivít;
- *administratívne*: napr. akreditačné posudky, inšpekčné správy, zápisnice z rokovaní, štatút školy, študijný poriadok ap.;
- *kurikulárne* (dokumenty programových projektov výučby): napr. profil absolventa, učebné osnovy, učebnice, výzdoba;
- *evalvačné*: napr. vysvedčenia, testy schopností, systém hodnotenia žiakov, žiacke knižky, zdravotné záznamy o žiakoch.

### b) K dokumentom prevažne osobnej povahy patria napr.:

- učiteľove heslovité zápisy štruktúry vysvetľovaného učiva na tabuli, hodnotiace poznámky v klasifikačnom notese, pamäti;
- žiacke denníky, poznámky a "čarbanice" v zošite alebo slohové práce, výkresy, náčrty; žiacka korešpondencia ap. (Švec, 1998).

Obsahová analýza týchto edukačných (školských, osvetových a iných) dokumentov je významným metodickým prostriedkom v empirickom poznávaní a praktickom pretváraní reality (Švec, 1998).

Podľa charakteru dokumentu a tiež podľa výskumného problému sa volia rôzne stratégie analýzy. V tejto časti uvádzame dve metódy: obsahovú analýzu textových dokumentov a analýzu obťažnosti textu.

## Obsahová analýza

### Obsahová analýza textov

Obsahová analýza textov je metóda, ktorá sa uplatňuje pri všetkých výskumných postupoch, v ktorých sa pracuje so slovom (analýza dotazníkov, prepisu z interview, protokolu z pozorovania a pod.). Predstavuje však aj osobitnú výskumnú metódu na analýzu textových dokumentov. Analogicky s obsahovou analýzou verbálnej stránky textu existuje i analýza neverbálnych produktov – obrázkov, schém, grafov a pod. (Gavora, 1999).

Podľa súčasného vnímania tejto metódy medzi odborníkmi môžeme konštatovať, že má potenciál poskytovať výsledky nielen na úrovni opisu a kvantifikácii zjavných prvkov, ale aj interpretačné a vysvetľujúce závery na základe korelácie premených.

Priekopníkom pri rozvoji kvantitatívneho poňatia obsahovej analýzy textu bol Bernard Berelson, podľa ktorého ide o procedúru na objektívny, systematický a kvantitatívny opis zjavného obsahu komunikácie (Švec, 1998).

Pri kvantitatívnom postupe sa výskumník opiera o vopred explicitne vyčlenené kategórie. Tieto kategórie môžu byť na úrovni gramatických kategórií, termínov, pojmov, myšlienok a tém (termín chápeme v užšom slova zmysle ako konkrétnu interpretáciu pojmu, práve tak, ako veta je konkrétnou podobou myšlienky).

Kvantifikačné postupy, ktoré sa aplikujú v obsahovej analýze textu, sú:

- Sledovanie frekvencie výskytu analytickej kategórie – môže ísť o absolútnu alebo relatívnu početnosť.
- Analýza kontingencie, čiže spoločného výskytu dvoch analytických kategórií, v dôsledku kauzálnej asociácie.
- Zisťovanie dichotómie – prítomnosti, resp. neprítomnosti nejakého ukazovateľa.
- Analýza intenzity – poradia alebo stupňa – určenie kvality na posudzovacej škále.
- Meranie plochy venovanej určitému problému v tlačenom alebo písanom texte. (V súvislosti s rozvojom elektronických médií sa rozšírila pôsobnosť obsahovej analýzy aj na iné, netextové dokumenty. V kvantifikácii sa napr. sleduje čas venovaný určitej téme v televízii a podobne.) (Švec, 1998)

Týmto spôsobom je kvalita, určená významom (obsahom) zvolených analytických kategórií, prevedená na kvantitatívnu mieru (Gavora, 1999).

V intenciách tradičného výkladu sa prístup k obsahovej analýze člení na kvantitatívny a kvalitatívny, pričom kvalitatívny spočíva v rozbere obsahovej stránky textu, a to na rôznych úrovniach – vysvetlenie zmyslu textu až po hlboké interpretácie. Gavora (1999) uprednostňuje pre rozbor obsahovej stránky textu pojem nekvantitatívna analýza. Kvalitatívny výskum vychádza z inej paradigmy ako kvantitatívny výskum; neakceptuje a priori definované predpoklady a z nich plynúce kategórie analýzy.

## **Príprava a postup pri obsahovej analýze textov**

Obsahová analýza textu má široké uplatnenie v rôznych odboroch – lingvistike, žurnalistike, sociológii, psychológii. V oblasti pedagogického výskumu je aktuálna obsahová analýza širokého spektra textov:

- učebné osnovy, učebné texty,
- písomné úlohy žiakov,
- písomné prípravy učiteľov na vyučovanie,
- zápisnice, správy a záznamy z činnosti,
- legislatívne materiály, štatistické výstupy, mediálne správy...

Pri analýze textov sa začína *stanovením výskumnej vzorky*. Uskutočňuje sa analogicky ako pri výbere výskumnej vzorky u osôb. Základný súbor textov totiž môže byť veľmi rozsiahly. Sú to všetky texty, na ktoré sa vzťahuje hypotéza. Analýza sa však uskutočňuje s výberovým súborom textov, ktoré predstavujú reprezentatívnu vzorku zo základného súboru.

Ďalším krokom je *voľba významovej jednotky* obsahu.

Môže ísť o jednotlivé slová, myšlienky (tvrdenia) alebo témy.

Táto voľba vyplýva zo zámerov štúdie.

**Príklad:**

Ako príklad uvádzame štúdiu Urbanová, Orolínová, Held (2004), ktorej cieľom bolo prehodnotiť obsah a formu prírodovedného vzdelávania na Slovensku najmä v súvislosti s realizáciou niektorých medzinárodných projektov (meraní) zameraných na hodnotenie úrovne vzdelanosti žiakov v jednotlivých krajinách (napr. TIMSS, PISA), v ktorých slovenskí žiaci dosahujú len priemerné výsledky. Predpokladalo sa, že tento stav súvisí práve so spôsobom a obsahom prírodovedného vzdelávania na našich školách.

Prostredníctvom obsahovej analýzy kurikulárnych dokumentov (u nás vzdelávacích štandardov), bolo snahou získať relevantné podklady pre zhodnotenie súčasnej situácie týkajúcej sa vyučovania prírodovedných predmetov u nás aj v zahraničí.

Za základnú vzorku sa v zhode so zámermi štúdie pokladali texty kurikulárnych dokumentov zo všetkých krajín, ktoré sa zúčastnili štúdií TIMSS a PISA. Výberový súbor, s ktorým sa výskum realizoval, však predstavoval reprezentatívnu vzorku krajín s rôznou úspešnosťou (vysokou, nízkou i strednou) v hodnotiacich štúdiách TIMSS a PISA.

**Stanovenie analytických kategórií**

Kategórie klasifikujú významové jednotky. Vychádzajú z príslušného výskumného problému a zo stanovenej hypotézy.

Väčšina metodologickej literatúry zhodne uvádza, že správne stanovenie analytických kategórií je kľúčovým momentom obsahovej analýzy. Na kategórie sa kladú tieto požiadavky:

- a) musia byť primerané skúmanému problému,
- b) musia byť vyčerpávajúce, t.j. musia zahrňovať každý možný prvok obsahu, ktorý súvisí so skúmaným problémom,
- c) musia sa vzájomne vylučovať, t.j. nesmú sa prekrývať. Významová jednotka, ktorá vstupuje do jednej kategórie, nesmie vstupovať zároveň do druhej kategórie (Gavora, 1999).

**Príklad:**

V spomenutom modelovom príklade obsahovej analýzy sa za analytické kategórie vzali kategórie PISA, ktoré sa mali sledovať a hodnotiť pri testovaní v roku 2006 (tieto požiadavky pre prírodovednú gramotnosť predstavujú medzinárodný konsenzus). Boli rozdelené na 4 oblasti - oblasť A bola zameraná na hodnotenie prírodovedných poznatkov, oblasť B na poznatky o vede, oblasť C na vedecké postupy a oblasť D na prírodovednú orientáciu, v ktorej bol dôraz na afektívnu zložku.

**Tabuľka 4 OBLASŤ A – Základné oblasti pre meranie prírodovedných poznatkov**

<b>Fyzikálne systémy</b>
štruktúra a vlastnosti hmoty (napr. tepelná a elektrická vodivosť)
fyzikálne a chemické zmeny (napr. skupenstvá hmoty, reakčná rýchlosť)
pohyb a sily (napr. rýchlosť, zrýchlenie, hybnosť)
energia a jej transformácie (napr. uchovávanie a uvoľňovanie)
interakcie energie a hmoty (napr. elektromagnetické vlny, zvukové a seizmické vlny)
<b>Živé sústavy</b>
bunky (napr. štruktúry a funkcie, DNA, rastlinné a živočíšne)
ľudia (napr. zdravie, výživa, sústavy orgánov, trávenie, dýchanie, obeh, ochorenia)
populácie (napr. druhy, evolúcia, biodiverzita, genetická premenlivosť)
ekosystémy (napr. potravné reťazce, trvalá udržateľnosť)
hmota, energia, štruktúra (napr. zložitosť a štruktúra organizmov, hierarchia)
<b>Zem a vesmír</b>
(krajinné) sféry Zeme (napr. litosféra, atmosféra, hydrosféra)
energia v (krajinných) sférach Zeme (napr. zdroje, globálna klíma)
zmeny v krajinných sférach Zeme (napr. tektonika pohybu Zeme, geochemické cykly, konštruktívne a deštruktívne sily)
história Zeme (napr. fosílie, pôvod a evolúcia)
Zem vo vesmíre (napr. gravitácia, slnečná sústava)

**Tabuľka 5 OBLASŤ B – Základné oblasti pre hodnotenie poznatkov o vede**

<b>Veda a jej aplikácie</b>
veda (napr. chápanie sveta prírody, zodpovedané otázky, pokrok ktorý ovplyvňuje spoločnosť)
technológie (napr. snahy riešiť problémy ľudí, nové postupy a výrobné procesy, analýza rizík, produkty, ktoré ovplyvňujú spoločnosť, pomáhajú ľuďom prispôbiť sa, môžu spôsobovať nové problémy, poznatky často nie sú verejné, úžitok verzus náklady, nechcené dôsledky)
veda a technika (napr. veda často napreduje vďaka novým technológiám, riešené technické problémy môžu podnietiť vedecký výskum a naopak)
<b>Vedecký výskum</b>
účel (napr. získať dôkaz, ktorý pomôže odpovedať na vedecké otázky, súčasné aktuálne názory, modely, teórie, metodiky výskumu)
pozorovanie a pokus (napr. rôzne otázky si vyžadujú rozdielny vedecký výskum, súčasné vedecké poznanie)
výsledky výskumu (napr. nový poznatok, nové metódy, nový výskum)
<b>Vedecké dokazovanie</b>
údaje (napr. kvantitatívne, kvalitatívne)
meranie (napr. zahrnutá neurčitnosť, opakovanosť, odchýlky, presnosť)
<b>Vedecká argumentácia</b>
druhy (napr. hypotézy, zákony, teória, model)
formulácia (vedomosť a nový dôkaz, tvorivosť a predstavivosť, logika)
pravidlá (napr. logická konzistentnosť, podloženosť dôkazmi, podklad v historických alebo súčasných vedomostiach)
výsledky (napr. nový poznatok, nová metóda, nový výskum)
<b>Veda a jej aplikácie v spoločnosti</b>
vplyv (napr. veda a technika ovplyvňuje spoločnosť prostredníctvom poznania, postupov, produktov a pohľadov na svet)
výzvy (napr. otázky a potreby spoločnosti často podnecujú otázky pre vedecký výskum a problémy pre inováciu technológií)
obmedzenia (napr. veda nevie zodpovedať všetky otázky a technika vyriešiť všetky problémy spoločnosti alebo naplniť všetky ľudské očakávania)

**Tabuľka 6 OBLASŤ C – Vedecké postupy**

<i>určiť (identifikovať) alebo formulovať vedecké otázky</i>
<i>určiť a aplikovať vhodný poznatok z prírodných vied alebo o vede (ako takej)</i>
<i>systematicky hľadať ďalšie informácie k vedeckej otázke alebo problému</i>
<i>interpretovať dôkaz a odvodené vedecké závery</i>
<i>komunikovať vedecké názory (myšlienky)</i>

**Tabuľka 7 OBLASŤ D – Prírodovedná orientácia**

<b>Uznanie prínosu prírodných vied</b>
<i>oceňovanie úlohy a príspevku vedy a techniky k pochopeniu javov a kultúrnemu svetu</i>
<i>uvedomenie si, že aplikácie poznatkov prírodných vied a technológií môžu priniesť etické problémy</i>
<i>hodnotenie prínosu mnohých ľudí z rôznych krajín a kultúr k rozvoju vedy a techniky</i>
<b>Prírodovedná zvedavosť</b>
<i>záujem o otázky z oblasti prírodných vied (prípadne so vzťahom k nim)</i>
<i>záujem o ďalšie prírodovedné poznatky a zručnosti, používanie množstva zdrojov a metód, vrátane uskutočňovania výskumu</i>
<i>pretrvávajúci záujem o prírodné vedy, vrátane zvažovania zamestnania so vzťahom k prírodným vedám</i>
<b>Prírodovedné uvažovanie</b>
<i>hodnotenie dôležitosti zväženia iných (alternatívnych) pohľadov, myšlienok a vysvetlení</i>
<i>hodnotenie využitia reálnych informácií a racionálnych interpretácií pri analýze a hodnotení</i>
<i>hodnotenie vplyvu sociálnych vzťahov na tvorbu racionálnych, dôkazmi podložených záverov</i>
<i>uznávanie opatrných logických procesov pri vyvodzovaní záverov</i>
<b>Environmentálna zaangažovanosť</b>
<i>preukázanie zmyslu pre osobnú i zdieľanú zodpovednosť za trvalo udržateľné prostredie</i>
<i>docenenie osobných, spoločenských a environmentálnych dôsledkov navrhovaných aktivít</i>
<i>preukázanie ochoty prispieť k zachovaniu trvalo udržateľného rozvoja</i>
<b>Zásady bezpečnosti</b>
<i>preukázanie záujmu o otázky bezpečnosti a pochopenie potreby pravidiel a obmedzení</i>
<i>preukázanie uvedomenia si priamych a nepriamych následkov svojho konania</i>
<i>preukázanie schopnosti identifikovať situácie zahŕňajúce osobnú bezpečnosť a dodržiavať príslušné bezpečnostné pravidlá a usmernenia</i>

Za významovú jednotku sa pokladala jedna položka (výrok) v štandarde. Za analytické kategórie sa považovali kategórie PISA pre hodnotenie prírodovednej gramotnosti (sú uvedené v tabuľkách 4 - 7).

Kvantifikačný postup, ktorý bol uplatnený spočíval len (vzhľadom na rôznorodosť použitých kurikulárnych dokumentov jednotlivých krajín) v zisťovaní prítomnosti vs. neprítomnosti zvolených analytických kategórií v kurikulárnych dokumentoch jednotlivých krajín.

**Kvantifikácia analytických kategórií**

Zisťuje sa ich frekvencia, t.j. absolútny počet, relatívny počet (percento), priemer, smerodajná odchýlka a pod. Obvyklé sú aj korelácie medzi kategóriami alebo štatistické testy rozdielov (napr. CHI-kvadrát).

Pri obsahovej analýze je výhodné prebrať už existujúce analytické kategórie z iných výskumov (pokiaľ sú vhodné na dané účely) a nekonštruovať za každú cenu

vlastné kategórie. Vyvíjanie nových analytických kategórií je totiž prácné. Podobné pravidlo však platí pre akékoľvek výskumné nástroje.

Vo výskume možno použiť akýkoľvek počet analytických kategórií, ktorý vyhovuje zámeru výskumu. Každú kategóriu okrem toho možno rozdeliť do niekoľkých subkategórií.

Od zámeru skúmateľa závisí i šírka danej kategórie. Úzke kategórie vedú k jemnejšej analýze, širšie naopak k hrubšej analýze. Šírku kategórií si stanoví skúmateľ na základe svojich výskumných zámerov. Frekvencie sa zoradia do prehľadných tabuliek, prípadne sa vyjadria graficky (Gavora, 1999).

**Príklad:**

**Tabuľka 8** Výskyt kategórií PISA v štandardoch jednotlivých krajín

			Výskyt kategórií PISA v kurikulárnych dokumentoch jednotlivých krajín									
			OBLASŤ A		OBLASŤ B		OBLASŤ C		OBLASŤ D		Celkovo	
Krajina	Typ kurikul. dokumentu	Zahrnuté predmety	%	absolútny výskyt	%	absolútny výskyt	%	absolútny výskyt	%	absolútny výskyt	%	absolútny výskyt
Fínsko	štandard	Science (F, Ch, B, G)	80	12/15	60	9/15	100	5/5	37,5	6/16	62,75	32/51
Holandsko	štandard	F, Ch	46,47	7/15	60	9/15	100	5/5	43,75	7/16	54,9	28/51
Anglicko	štandard	Science (F, Ch, B, G)	86,67	13/15	66,67	10/15	100	5/5	50	8/16	70,59	36/51
Kalifornia	štandard	Science (F, Ch, B, G)	100	15/15	66,67	10/15	100	5/5	6,25	1/16	60,78	31/51
Kolorádo	štandard	Science (F, Ch, B, G)	100	15/15	73,33	11/15	40	2/5	25	4/16	62,75	32/51
Česko	štandard	F, Ch, B	100	15/15	53,33	8/15	20	1/5	43,75	7/16	60,78	31/51
Slovensko	štandard	F, Ch, B, G	100	15/15	66,67	10/15	60	3/5	43,75	7/16	70,59	36/51
Poľsko	štandard	F, Ch, B, G	86,67	13/15	26,67	4/15	60	3/5	43,75	7/16	52,94	27/51
Maďarsko	osnovy	F, Ch, B, G	86,67	13/15	40	6/15	60	3/5	37,5	6/16	54,9	28/51

## Interpretácia zistených frekvencií

Zistené kvantitatívne údaje sa slovne opíšu, vysvetlia a interpretujú.

Súčasným trendom vo výskume je hľadanie vzťahov medzi niekoľkými analytickými kategóriami (teda nielen stanovovanie frekvencií izolovaných kategórií), alebo hľadanie vzťahu medzi obsahovými kategóriami a inými premennými (získovanými ďalšími výskumnými nástrojmi).

Kerlinger (1972) uvádza, že je potrebné splniť isté podmienky, aby sa kvantifikácia mohla realizovať. Smú sa kvantifikovať len tie kategórie, ktorých výskyt je dostatočný. Ak ide o málo frekventovanú kategóriu, t.j. ak sa javy nevyskytujú v každej zo skúmaných kategórií v dostatočnom množstve, neumožňuje to štatistické zovšeobecňovanie. Mnohé obsahové analýzy neplnia túto podmienku, a preto v týchto prípadoch môžeme hovoriť len o exploratívnom použití obsahovej analýzy (Gavora, 1999).

**K príkladu:** Zo záverov obsahovej analýzy našich, ale aj zahraničných kurikulárnych dokumentov vyplynulo, že v nich absentujú hlavne zložky zamerané na vedecké postupy. Autori skonštatovali, že nie sú podchytené ani v učebniciach a je potrebné hľadať možnosti, akým spôsobom túto situáciu zmeniť.

## Validita a reliabilita obsahovej analýzy textu

Obsahová analýza sa obyčajne validizuje za pomoci inej metódy.

Napríklad obsahová analýza učiteľových výrokov o vlastnej práci sa porovnáva s pozorovaním jeho práce v triede. Platí to však i naopak. Obsahová analýza validizuje inú výskumnú metódu. Odpovede v dotazníku sa porovnávajú s písomným produktom skúmaných osôb (napr. voľnou písomnou prácou). Určí sa miera zhody. Ak výskum nemá zostať len na úrovni opisu reality, ale má overovať hypotézy alebo hľadať súvislosti, vždy má ísť o porovnanie výsledkov získaných pomocou obsahovej analýzy s výsledkom z iného zdroja. Platí to hlavne pre zisťovanie príčin a následkov.

Reliabilita obsahovej analýzy textu je tým vyššia, čím sú analytické kategórie presnejšie vymedzené. Na druhej strane však príliš úzke kategórie znižujú reliabilitu a naopak široké kategórie ju zvyšujú (znižuje sa však jemnosť diferenciacie).

Reliabilita obsahovej analýzy sa obyčajne zisťuje mierou zhody výsledkov analýzy dvoch alebo viacerých zaškolených ľudí. Táto miera zhody sa zisťuje pri každej analytickej kategórii a pri všetkých kategóriách súborne.

Jej charakteristickou črtou je, že ide o non-reaktívnu výskumnú metódu. Autor textu nie je obyčajne prítomný pri analýze obsahu textu, nemôže teda reagovať, nevlýva na skúmateľa (v niektorých prípadoch ani nevie, že sa jeho text podrobí obsahovej analýze) (Gavora, 1999).

## Literatúra

Gavora, P. *Úvod do pedagogického výskumu*. Bratislava : Univerzita Komenského, 1999. ISBN 80-223-1342-4

Kerlinger, F. N. *Základy výskumu chování*. Praha : Academia, 1972, pp. 547 – 562.

Švec, Š. a kol. *Metodológia vied o výchove*. Bratislava : IRIS, 1998. ISBN 80-88778-73-5

Urbanová, A.; Orolínová, M.; Held, L. A. Priebežné výsledky riešenia projektu integrovaného vyučovania prírodných vied pre základnú školu. In: *Sborník příspěvků – Mezinárodní seminář didaktiků chemie, sborník prací Pedagogické fakulty Masarykovy univerzity v Brně*, Brno – Šlapanice : Pedagogická fakulta MU, 2004, s. 102 – 111. ISBN 80 – 210 – 3616 – 5

## Obťažnosť textov

### Určovanie obťažnosti učebných textov

Jedným z variantov analýzy textu je určovanie jeho obťažnosti, čo je jeden z najdôležitejších didaktických parametrov.

Ktorýkoľvek text má určité obsahové a formálne črty, ktorými je viac alebo menej usposobený pre komunikáciu s určitými subjektami. Keď začneme čítať určité texty, intuitívne pociťujeme, či ide o texty ľahké, jednoduché, ľahko zrozumiteľné alebo o texty ťažké, zložité až nezrozumiteľné. Toto intuitívne hodnotenie textu subjektom je odrazom objektívne existujúcich komunikačných vlastností textu. Obťažnosť textu učebnice môžeme chápať ako objektívnu vlastnosť, ktorá je daná jeho špecifickými charakteristikami. To, že je to vlastnosť nezávislá od subjektu, umožňuje, aby táto vlastnosť bola zisťovaná a vyhodnocovaná objektívnymi procedúrami (Průcha, 1997).

Obťažnosť textu je súhrn takých vlastností textu, ktoré existujú objektívne v ktoromkoľvek texte a v procese učenia majú vplyv na percepciu, chápanie a spracovanie textovej informácie učiacim sa subjektom (Průcha, 1998).

Obťažnosť textu možno určiť odhadom, ktorí robia experti alebo používatelia textu (učitelia, žiaci). Často sa pritom používajú dotazníky alebo posudzovacie škály, na ktorých respondenti hodnotia jednotlivé aspekty textu. (Průcha, 1984, 1998 podľa Gavoru, 1999).

**Príklad**

<i>jednoduchý jazyk</i>	1 2 3 4 5 N	<i>zložitý jazyk</i>
<i>ťažké pojmy</i>	1 2 3 4 5 N	<i>primerané pojmy</i>
<i>podrobný výklad</i>	1 2 3 4 5 N	<i>všeobecný výklad</i>
<i>zaujímavý text</i>	1 2 3 4 5 N	<i>nudný text</i>
<i>atď.</i>		

Pri inom spôsobe zisťovania obťažnosti sa žiakom dávajú úlohy na: selekciu informácií, usporiadanie informácií, dopĺňovanie chýbajúcich slov v texte a pod. a porovnávajú sa ich výsledky pri každom texte. Podľa výsledkov sa určí obťažnosť textu. (Gavora, 1999)

Odlíšny spôsob zisťovania obťažnosti textu je založený na objektívnych parametroch textu a je nezávislý od posudzovateľov alebo žiakov. Týmito parametrami sú obyčajne dĺžka vety, zložitost' vety, množstvo a obťažnosť pojmov, miera opakovateľnosti pojmov, atď. (Gavora, 1999). Obťažnosť určitého textu pre čitateľa je teda ovplyvňovaná lexikálnymi, ale aj syntaktickým faktormi. Tieto sú začleňované do tzv. vzorcov pre výpočet obťažnosti textu.

Jednotlivé ukazovatele sú potom včlenené do vzorca, do ktorého sa dosadia údaje z textu a pomocou ktorého sa vypočíta obťažnosť textu. Existuje niekoľko desiatok vzorcov na výpočet obťažnosti textu, ktoré sa líšia zdôrazňovaním alebo naopak potlačovaním niektorých textových parametrov. Sú to napr. vzorce Flescha, Challa, Mikka, Pisareka, Mistríka a ďalších. U nás je najrozšírenejší Nestlerovej vzorec, ktorý v modifikovanej podobe používa aj J. Průcha (Gavora, 1999).

**Vzorce obťažnosti textu**

Veľmi jednoduchým univerzálnym vzorcem obťažnosti textu je **vzorec LIX**, ktorý vyvinul Švéd C. H. Björnsson. Obťažnosť textu učebnice sa vypočítava podľa vzorca:

$$LIX = Lm + Lo$$

kde Lm – priemerná dĺžka vety vyjadrená počtom slov v súbore 200 viet,  
Lo – priemerná dĺžka slov s počtom viac ako 6 písmen v súbore 2000 slov.

Súbory viet (syntaktický faktor) sa vytvárajú z 20 vzoriek po 10 viet vybraných systematicky z rôznych častí učebnice. Súbory slov (lexikálny faktor) sa vytvárajú z 20 vzoriek po 100 slov. Miera LIX je vybavená škálou obťažnosti, kde:

- LIX = 20 – 30 bodov ..... veľmi ľahké texty
- LIX = 30 – 40 bodov ..... stredne obťažné texty
- LIX = 40 – 50 bodov ..... veľmi obťažné texty
- LIX = 50 – 60 bodov ..... extrémne obťažné texty

Björnssonova miera LIX je vybavená normami obťažnosti didaktických textov pre jednotlivé ročníky. Napr. pre švédske školy je norma takáto:

ročník školy	1.	2.	3.	4.	5.	6.	7.	8.	9.
normovaná obťažnosť (LIX)	17 – 19 bodov	21 – 23 bodov	24 – 27 bodov	27 – 31 bodov	29 – 34 bodov	31 – 36 bodov	32 – 38 bodov	34 – 40 bodov	35 – 42 bodov



Táto miera sa považuje za príliš jednoduchú, pretože pracuje len s dvomi parametrami textov a preto nie je celkom validná. Uvádzame ju ako demonštráciu možnosti stanoviť pre jednotlivé vekové stupne žiakov kritériá obťažnosti, ktoré by mali spĺňať didaktické texty v príslušnom ročníku školy.

Ďalším použiteľným vzorcom je **Mistríkov vzorec obťažnosti**, ktorý začleňuje tri parametre textu:

- V – priemerná dĺžka viet (je príznakom zložitosti vyjadrovania myšlienok),
- S – priemerná dĺžka slov vyjadrená počtom slabík (je príznakom pojmovej zaťažnosti textu),
- I – index opakovania slov (je charakteristikou lexikálnej variability textu).

Mistríkova miera používa škálu obťažnosti s hodnotami medzi 0 – 50 bodov, v ktorej najľahšie texty majú 40 – 50 bodov a najťažšie 0 – 10 bodov. Vzorec má takýto tvar:

$$R = 50 - V \cdot S / I, \quad \text{kde } I = N / L$$

pričom N – počet všetkých lexikálnych jednotiek textu,  
L – počet rôznych lexikálnych jednotiek.

Výhodou Mistríkovo vzorca je, že zahrňuje aj lexikálnu variabilitu textu. V prípade školských učebníc je to ukazovateľ využívania rozdielnych slov v texte, ktorý má priamu súvislosť s rozsahom slovnej zásoby žiakov určitého veku. Tento vzorec má však aj nedostatok: nezahrňuje meranie kvalitatívnej stránky lexiky a syntaxe textu. Pre čitateľa, ktorý sa snaží porozumieť čítanému textu, je relevantný nielen kvantitatívny rozsah lexiky textu, ale aj jeho kvalitatívne parametre – či sú napr. v texte používané odborné termíny a ako často sa opakujú. Podobne aj v syntaktickom faktore je relevantná nielen dĺžka viet, ale aj to, aké zložité sú vetné konštrukcie, napr. či ide o súvetie s väčším alebo menším počtom propozícií (vetných úsekov).

Komplexnejšou mierou obťažnosti je **Nestlerovej vzorec** (autorkou je nemecká psychologička Kathrin Nestlerová), ktorý bol modifikovaný aj pre české a slovenské texty. Táto miera spĺňa tri základné požiadavky:

1. je dostatočne komplexná (zahŕňa aj sémantické aspekty textu),
2. je dostatočne operatívna (dá sa pomerne jednoducho aplikovať),
3. je dostatočne validná (t.j. s overenou koreláciou ku kognitívnym schopnostiam žiakov školského veku).

V súčasnosti sa v praxi používa inovovaná miera Nestlerovej – Průchu – Pluskala, ktorá sa skrátene nazýva **miera T** – symbol T označuje stupeň obťažnosti textu. Obťažnosť textu je meraná na základe vzorca:

$$T = T_s + T_p$$

kde  $T_s$  (syntaktická obťažnosť) a  $T_p$  (pojmová, sémanticko-lexikálna obťažnosť) sa vypočítajú podľa zvláštnych vzorcov:

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

V – priemerná dĺžka viet

U – priemerná dĺžka vetných úsekov (predikatívnych propozícií)

$$T_P = 100 * \sum P / \sum N * (\sum P_1 + 3\sum P_2 + 2\sum P_3 + 2\sum P_4 + \sum P_5) / \sum N, \text{ kde}$$

N – počet všetkých slov

P – počet všetkých pojmov

P<sub>1</sub> – bežné pojmy

P<sub>2</sub> – odborné pojmy

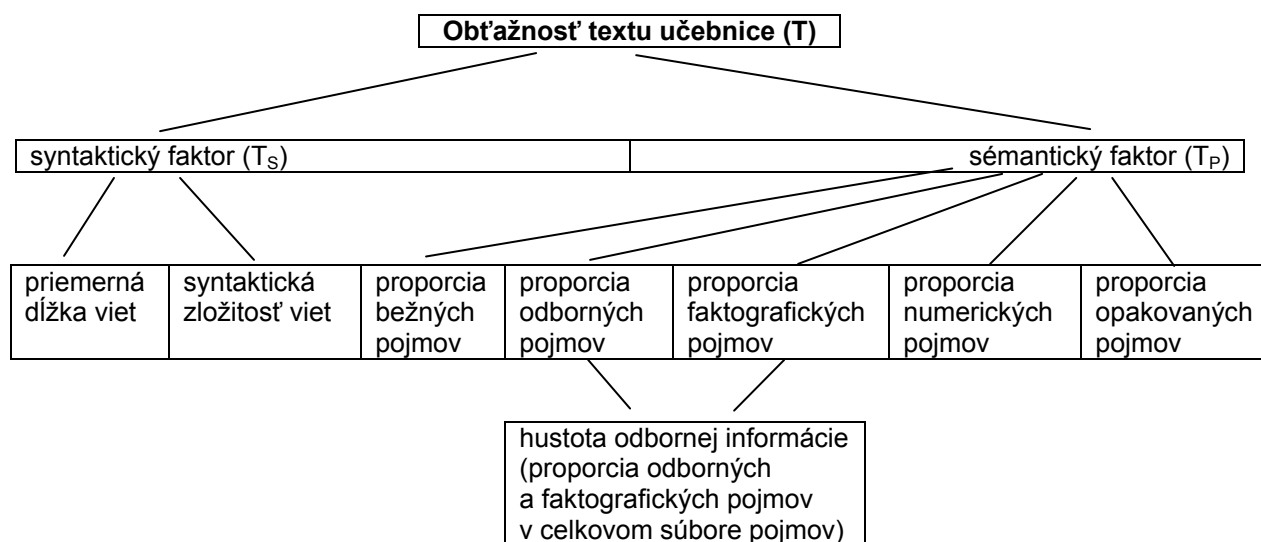
P<sub>3</sub> – faktografické pojmy

P<sub>4</sub> – číselné údaje

P<sub>5</sub> – opakované údaje

Σ - suma – súhrnný počet daných jednotiek

Miera T zahrňuje tieto parametre:



**Inštrukcia pre analýzu obťažnosti textu učebníc (Průcha, 1998):**

### Zložky analýzy

Aplikácia miery T pozostáva z niekoľkých krokov:

1. Výber vzoriek textu z učebnice.
2. Výpočet stupňa syntaktickej obťažnosti (T<sub>S</sub>).
3. Výpočet stupňa sémantickej obťažnosti (T<sub>P</sub>).
4. Výpočet stupňa celkovej obťažnosti (T).
5. Výpočet koeficientov hustoty odbornej informácie (i, h).
6. Interpretácia výsledkov analýzy.

Aby bola dosiahnutá maximálna presnosť analýzy, je nutné venovať veľkú pozornosť výberu vzoriek a zachovaniu predpísanej kategorizácie skúmaných jednotiek textu.

### Opis jednotlivých krokov analýzy

#### 1. Výber textových vzoriek

- a) Zo skúmaného textu sa vyberie 5 vzoriek, každá s rozsahom najmenej 200 slov. Podľa modifikácie M. Pluskala je vhodnejšie vybrať 10 vzoriek s rozsahom 200 slov. Za „slovo“ sa považuje akýkoľvek slovný, číselný alebo symbolický výraz (vrátane skratiek), ktorý je v texte oddelený grafickými medzerami alebo grafickými rozdeľovacími znakmi (interpunkčnými ai.). Ako slová sa počítajú aj ustálené značky (napr. mier, hmotností, mien, fyzikálnych veličín, matematických a hudobných pojmov ai.) – napr. km, Sk, log, %, atď., SR, USA a pod.

Vzorky by mali rovnomerne pokrývať rôzne časti (témy) učebnice a nemali by byť vyberané z niekoľkých prvých stránok, kde zvyčajne býva netypický text. Nadpisy a marginálie sa do vzoriek nevyberajú. Každá zo vzoriek musí tvoriť súvislý text.

- b) V každej vzorke sa odpočíta 200 slov. Keďže dvojsté slovo sa spravidla nekryje s koncom vety, dopočíta sa vzorka až k najbližšiemu koncu vety. Tým vzniknú vzorky, ktoré o niečo presahujú počet 200 slov.
- c) Stanoví sa počet viet v každej vzorke. „Veta“ sa pre daný účel definuje ako akákoľvek postupnosť „slov“, ktorá začína veľkým písmenom a končí bodkou alebo iným grafickým znakom (otáznik, dvojbodka, rámeček ai.).

**Príklad:** V učebnici Zeměpis 7 bolo vybraných 5 vzoriek:

- na s. 18	203 slov, 18 viet
- na s. 47	214 slov, 14 viet
- na s. 73	208 slov, 16, viet
- na s. 102	203 slov, 20 viet
- na s. 128	204 slov, 16 viet

Celkový súbor, ktorého analýza sa uskutočňuje, teda predstavuje:

$$\sum N = 1032 \text{ slov}$$

$$\sum V = 84 \text{ viet}$$

## 2. Výpočet syntaktickej obťažnosti

- a) Vypočíta sa priemerná dĺžka vety (V) podľa vzorca:  $V = \sum N / \sum V$

**V predchádzajúcom príklade je to teda**  $V = 1032 / 84 = 12,28$ .

- b) V každej vzorke sa zistia slovesá v určitom tvare (nie v infinitíve – neurčitku). Pre urýchlenie práce Průcha doporučuje slovesá farebne podčiarkovať. Pozor! Zložené tvary sloviess skladajúce sa z dvoch alebo viacerých slov sa považujú za 1 „sloveso“.

**Príklad:** Očekávali by sme = 3 slová, ale 1 sloveso

Uskutočnili ste = 2 slová, ale 1 sloveso

- c) Zistí sa celkový počet sloviess v danom súbore a potom sa vypočíta priemerná dĺžka vetných úsekov (U) podľa vzorca:  $U = \sum N / \sum U$

**Príklad:** Vo vyššie uvedenom súbore bolo zistených celkom 96 sloviess, takže

$$U = 1032 / 96 = 10,75$$

- d) Syntaktická obťažnosť textu sa vypočíta podľa vzorca:  $T_s = 0,1 * V * U$

**Príklad:** V učebnici Zeměpis 7 je  $T_s = 0,1 * 12,28 * 10,75 = 13,201$

## 3. Výpočet sémantickej obťažnosti

Výpočet sa uskutočňuje z hodnôt zistených pre 5 kategórií pojmov ako nositeľov sémantickej informácie:

- P<sub>1</sub> – bežné pojmy,
- P<sub>2</sub> – odborné pojmy,
- P<sub>3</sub> – faktografické pojmy,
- P<sub>4</sub> – číselné údaje,
- P<sub>5</sub> – opakované pojmy.

- a) V každej vzorke sa zistia (a môžu sa podčiarknuť obyčajnou ceruzkou) všetky podstatné mená („pojmy“), vrátane:

- abstraktných a dejových podstatných mien, napr. spoločenstvo, väzby, funkcia, zahájenie a pod.,
- spodstatnených prídavných mien, napr. pracujúci, poddaní,
- osobných mien a priezvisk, napr. Karol IV, Charles Darwin,
- skratiek označujúcich rôzne pojmy, napr. OSN, Rh-faktor.

- b) V každej vzorke sa zistia (a podčiarknu určitou farbou) všetky odborné pojmy ( $P_2$ ), t.j. tie podstatné mená, ktoré majú platnosť odborného termínu v nejakom obore. Dvojslovné a viacslovné termíny sa počítajú ako jeden pojem.

**Príklad:** V texte *Na každé teleso pôsobí gravitačná sila v smere...* je celkom 8 slov, z toho jeden bežný pojem (*smere*) a dva odborné pojmy (*teleso, gravitačná sila*).

Niekedy býva problémom odlišenie odborných pojmov od bežných a faktografických pojmov. V sporných prípadoch je nutné sa opierať o terminologické slovníky alebo konzultovať s príslušnými odborníkmi.

- c) V každej vzorke sa zistia (a opäť podčiarknu inou farbou) faktografické pojmy ( $P_3$ ). Pre daný účel (v príklade) sú to:
- vlastné osobné mená (rodné mená, priezviská, názvy bytostí), napr. Archimedes, Přemyslovci, Rusalka,
  - názvy ľudských výtvorov a zariadení, štátov, národov, inštitúcií, umeleckých výtvorov ai., napr. Národné divadlo, Kelti, Európska únia,
  - zemepisné názvy a názvy prírodných javov, napr. Venuša, Podunajská nížina,
  - všetky skratky a značky pre výrazy z uvedených skupín, napr. SR, SRN.

Dvojslovné a viacslovné faktografické pojmy počítame opäť ako jeden pojem  $P_3$ .

**Napr.** Zmluva o hospodárskej a menovej únii je výraz zložený zo 6 slov, ale počíta sa ako jeden faktografický pojem (pretože ide o názov jedného ľudského výtvoru).

- V každej vzorke sa zistí počet kvantitatívnych (číselných) údajov ( $P_4$ ) vyjadrujúcich napr. letopočty, vzdialenosti, hmotnosť a pod. Do tejto kategórie nepočítame odkazy na čísla strán, úloh, obrázkov ai. Tieto pojmy podčiarkujeme ďalšou farbou.
- V každej vzorke sa zistí počet opakovaných pojmov ( $P_5$ ). Sú to pojmy kategórií  $P_1$  až  $P_4$ , ktoré sa v texte vzorky už vyskytli.

**Príklad:** *Dopady meteoritov na začiatku vývoja Zeme boli také obrovské... Na povrchu Zeme tak vznikol prvý stály oceán obklopený atmosférou...*

*V tomto texte je slovo Zem jednak faktografickým pojmom ( $P_3$ ), jednak ho počítame do kategórie opakovaných pojmov ( $P_5$ ).*

- Všetky pojmy, ktoré nepatria do kategórie  $P_2$  až  $P_4$ , t.j. tie, ktoré zostali podčiarknuté len obyčajnou ceruzkou, tvoria množinu bežných pojmov ( $P_1$ ). Sú to výrazy, s ktorými sa žiaci stretávajú vo svojom každodennom živote, poznajú ich význam. V niektorých prípadoch môže byť určitý výraz buď bežným pojmom alebo odborným pojmom. To je potrebné rozlíšiť podľa kontextu, v ktorom sa výraz vyskytuje.

**Príklad:** *V texte Zem a planéty jej podobné vznikli chladným nabaľovaním je výraz chladné nabaľovanie vo výklade vzniku slnečnej sústavy určite odborným pojmom ( $P_2$ ), ale v inom kontexte môže byť výraz nabaľovanie považovaný za bežný pojem ( $P_1$ ), žiakom známy.*

- Sémantická obťažnosť analyzovaného textu sa vypočíta podľa vzorca:

$$T_P = 100 * \sum P / \sum N * (\sum P_1 + 3\sum P_2 + 2\sum P_3 + 2\sum P_4 + \sum P_5) / \sum N$$

Vo vzorci majú jednotlivé kategórie pojmov rozdielnu váhu podľa toho, aká je predpokladaná obťažnosť určitých pojmov pre žiakov. Z tohto dôvodu majú odborné pojmy najväčšiu váhu, a to 3, faktografické a kvantitatívne pojmy majú váhu 2 (viď. Pluskal, 1996).

#### 4. Výpočet celkovej obťažnosti textu

Uskutoční sa podľa vzorca  $T = T_S + T_P$

**Príklad:** Pri meraní stupňa obťažnosti výkladového textu na tému *Atmosféra* vypočítal Pluskal (1996) tieto hodnoty:

učebnica ZŠ:  $T_S = 11,35$ ;  $T_P = 30,74$ ; takže  $T = 11,35 + 30,74 = 42,09$

učebnica gymnázia:  $T_S = 16,96$ ;  $T_P = 34,75$ ; takže  $T = 16,96 + 34,75 = 51,71$

#### 5. Výpočet koeficientov hustoty odbornej informácie (i, h)

Tieto koeficienty sa počítajú podľa vzorcov:

$$i = 100 * (\sum P_2 + \sum P_3 + \sum P_4) / \sum N \quad h = 100 * (\sum P_2 + \sum P_3 + \sum P_4) / \sum P$$

- koeficient (i) udáva proporciu pojmov nesúcich odbornú informáciu v celkovom počte slov,
  - koeficient (h) udáva túto proporciu v celkovom počte pojmov.
- Hodnoty oboch koeficientov sú vyjadrené v percentách.

**Príklad:** Vo vyššie uvedených textoch prezentujúcich výklad témy *Atmosféra* boli zistené takéto hodnoty koeficientov i a h (Pluskal, 1996):

učebnica ZŠ:  $i = 20,73 \%$   $h = 52,68 \%$

učebnica gymnázia:  $i = 22,87 \%$   $h = 54,50 \%$

### Interpretácia výsledkov analýzy

Zistené hodnoty obťažnosti textu (T) a koeficientov hustoty odbornej informácie je možné interpretovať k rôznym účelom evalvácie a komparácie, a to:

#### a) Vzájomné porovnávanie učebníc:

- porovnávanie učebníc rôznych predmetov v tom istom ročníku,
- porovnávanie učebníc rovnakého predmetu v jednotlivých ročníkoch školy,
- porovnanie učebníc určitého predmetu v rôznych druhoch škôl,
- porovnanie učebníc rovnakých predmetov publikovaných rôznymi nakladateľstvami,
- porovnanie učebníc rovnakých predmetov v historickom vývoji (napr. obťažnosť učebníc zemepisu od minulého storočia po súčasnosť,
- porovnanie učebníc rovnakých predmetov v rôznych krajinách (napr. učebnice dejepisu v slovenských, českých, nemeckých, francúzskych ai. školách).

#### b) Detailné objasnenie vlastností konkrétnej učebnice:

Týka sa hlavne identifikácie príčin, prečo v určitej učebnici dochádza k nadmernej obťažnosti a čo je teda nutné v texte učebnice korigovať. Týmto spôsobom je možné dospieť ku konkrétnym doporučeniam na zníženie nadmernej obťažnosti textu učebnice.

**Príklad:** V bodoch 4 a 5 týchto inštrukcií sme uviedli údaje o texte českých učebníc zemepisu pre základnú školu a pre gymnázium (podľa Pluskala, 1996). Ukazuje sa pritom, že pre danú tému výkladu (ale podobne to môže platiť aj pre celú učebnicu) sa texty týchto učebníc prakticky neodlišujú čo do hustoty odbornej informácie: Aj žiaci ZŠ, aj žiaci gymnázia sa stretávajú vo svojich učebniciach s textami s vysokou hustotou odbornej informácie (koeficient h), t.j. viac ako 50 % všetkých pojmov sú také, ktoré nesú odbornú informáciu. To je určite neprímeraná proporcia hlavne v učebniciach pre základnú školu. Je takisto nutné hľadať príčinu tohto stavu:

*Prečo je v téme Atmosféra týchto učebníc taká vysoká hustota odbornej informácie? Detailné zistenia ukazujú, že je to predovšetkým v dôsledku vysokej proporcie odborných pojmov, keď v oboch učebniciach tvoria pojmy  $P_2$  18 % všetkých slov. Naopak faktografické pojmy  $P_3$  a číselné údaje  $P_4$  tvoria veľmi nízke proporcie medzi všetkými slovami daného textu.*

Ak sa má teda zvýšiť zrozumiteľnosť výkladu oboch učebníc, je nutné urobiť tieto korekcie:

- V texte učebnice pre ZŠ znížiť počet odborných pojmov  $P_2$ .
- V texte učebnice pre gymnázium je nutné uskutočniť to isté, ale navyše je potrebné znížiť stupeň syntaktickej obťažnosti textu ( $T_S = 16,9$ ), ktorý prispieva k tomu, že celkovo má výklad danej témy v tejto učebnici neúmerne vysokú celkovú obťažnosť textu ( $T = 51,7$ ).

## **Literatúra**

- Gavora, P.: *Úvod do pedagogického výskumu*. 2. vyd. Bratislava : Univerzita Komenského, 1999. 236 s. ISBN 80-223-1342-4
- Průcha, J.: *Moderní pedagogika*. Praha : Portál, 1997. 495 s. ISBN 80-7178-170-3
- Průcha, J. *Učebnice: Teorie a analýzy edukačního média*. Brno : Paido, 1998. 148 s. ISBN 80-85931-49-4
- Pluskal, M.: Zdokonalení metody pro měření obtížnosti didaktických textů . *Pedagogika*, 46, 1996, č. 1, s. 62 - 76.