Ondřej Šimik
ondrej.simik@osu.cz
Uniwersytet w Ostrawie

Syntactic and semantic difficulty of text in the history part of textbooks for the educational field of the „Human and his world"

Introduction

Even though we live in the 21ˢᵗ century, which is characterized by major development in information and communication technology, with the increase in electronic media and the expansion of the Internet, traditionally printed textbooks have (and most likely will continue to have) a firm place not only in the Czech school (see Maňák, 2008). Dvořák (2005) quotes some significant researchers in the field of textbook research. For example, the Australian professor, Mike Horsley says that textbooks will continue to be the typical characteristic of all classes and they will not be replaced by any electronic products. Průcha speaks about the renaissance textbook. Mikk (2007) claims that textbooks are the nation's future. Dvořák (2005) points out that these experts base their claims on researched facts and even though individual countries work with different textbooks, various studies claim that textbooks play an important role in education regardless of the specific country. The textbook is a curriculum project (Maňák, 2007). Průcha (2008), states that textbooks represent the most specific form of the projected curriculum. Therefore, it is a means by which the student comes into direct contact with the material being taught, which the student is to claim as his own and a means by which educational targets and expected improvements are to be achieved. Therefore, it is obvious that it is important to research the quality of the textbooks, which

are used in schools, especially in today's day and age, when the textbook market is oversaturated. Currently, there are tens of publishers issuing hundreds of textbooks (for example, Greger [2004 or 2005] discovered up to 63 publishers issuing textbooks). Although a textbook has to obtain an approval clause in the Czech Republic, as stated by Jakubcová (2012), individual textbooks differ quite a bit. To a certain extent, the process of approving textbooks depends on the reviewer's subjective viewpoint. This subjectivity could at least be partially limited if the textbook's quality was measured objectively. The contributing factor is educational research of textbooks, which is gradually being developed in the Czech Republic and which should give (is giving) solutions related to the quality of textbooks. As stated by Beneš, Janoušek, Novotný (2009), requirements for the quality of textbooks were formulated and they can be summarized into the following criteria: expertise (compliance with contemporary scientific findings in the field), content (arranging the educational content in compliance with curricular documents), didactic processing (the content of the textbook is appropriate and comprehensive for the student), linguistic aspect (level of language and style), esthetic (creation and graphics) and ergonomic (quality of paper, print, weight, textbook format). Textbooks are subject to research from various points of view (explanatory text, homework material, visual components, using the textbook by the teacher and students, finding out the opinions about the textbook by the students, etc.). By rule, it most frequently deals with finding out how the textbook is equipped in terms of didactics, where several methods were written up (for example, Průcha [1998] or Sikorová [2007]). From the didactics point of view, it deems appropriate to also measure the **verbal characteristics of the text in the textbooks,** since the text especially carries the behavioral – educational content. If the student is to work with the textbook, which is (should be) its purpose, it is necessary for the student to be able to understand the text. This is especially true for subjects of a teaching character, which encompass an integrated educational field of the Human and his World (science, history, etc.). Understanding the text is one of the assumptions of understanding the curriculum and achieving educational targets. In educational research, we will also find methods of evaluating the text of textbooks according to the **accessibility of the curriculum for the student** criterion (see Průcha, 2006). This criterion is referred to as the overall **extent of difficulty of the text** (it is labeled with the letter T). Research devoted to measuring the difficulty of the text in a textbook is clearly illustrated by Červenková (2011). For Czech conditions, the most sophisticated method

of measuring the difficulty of text in textbooks is the method created by Nestler (1974), which was later elaborated by Průcha (1989) and innovated by Pluskal (1996), who especially expanded the range of examined samples in one textbook (10 samples and 200 words) as well as the number of parameters of semantic components (types of concepts) from three to five. The method of measuring the difficulty of text is also defined as the Nestler-Průcha-Pluskal method (for example, in Průcha [1997]).

Methodology of Research

The research used the quantitative methodology, based on analysis of text. The following objectives were defined within the research: **1) to find out and compare the overall extent of the difficulty of text (T) of the history part of Czech textbooks for the subject of history (the Human and his World); 2) to find out and compare the extent of syntactic difficulty of text (Ts) of the history part of Czech textbooks for the subject of history (the Human and his World);** 2a) to find out and compare the average length of sentences (V average) of the history part of Czech textbooks for the subject of history (the Human and his World); 2b) to find out and compare the average length of the sentence section (U) of the history part of Czech textbooks for the subject of history (the Human and his World); **3) to find out and compare the extent of semantic difficulty of text (Tp) of the history part of Czech textbooks for the subject of history (the Human and his World);** 3a) to find out and compare the number of common concepts (P1) of the history part of Czech textbooks for the subject of history (the Human and his World); 3b) to find out and compare the number of expert concepts (P2) of the history part of Czech textbooks for the subject of history (the Human and his World); 3c) to find out and compare the number of facto-graphic concepts (P3) of the history part of Czech textbooks for the subject of history (the Human and his World); 3d) to find out and compare the number of numerical concepts (P4) of the history part of Czech textbooks for the subject of history (the Human and his World); 3e) to find out and compare the number of repetitive concepts (P5) of the history part of Czech textbooks for the subject of history (the Human and his World); **4) to find out and compare the density coefficient of expert information in the total number of concepts (h) of the history part of Czech textbooks for the subject of history (the Human and his World); 5) to find out and compare the information coefficient in**

the total number of words (i) of the history part of Czech textbooks for the subject of history (the Human and his World).

Description of the Research Method

For calculation, we used the method of measuring the difficulty of text-book text, which was developed by Nestler-Průcha-Pluskal. With the aid of this method, a total of 10 different sections from every textbook at a length of 200 words (see length in the description of the research sample) are researched from one textbook. According to this method, the overall difficulty of text (T) is equal to the sum of the syntactic difficulty (Ts) value and concept or semantic difficulty (Tp). The value of syntactic difficulty depends on the number of words (N) and verbs (U) in the non-infinitive form and the number of sentences, therefore, the average length of a sentence (V) in the selected section of the textbook. We will first choose a sample of 200 words (by rule, this is several words more, since we incorporate complete sentences into the sample). We count the number of sentences and according to the following formula: $V = \sum N / \sum V$, calculate the **average length of the sentence** (V). Then, we find out the number of verbs in a specific form (not in the infinitive) and according to the $U = \sum N / \sum U$ formula, we calculate the **average length of sentence sections** (U). We then calculate the syntactic difficulty according to the following formula: **Ts = 0,1 x V x U.** The measuring unit is points.

For the **semantic (concept) difficulty (Tp)** it is necessary to identify a total of five different types of concepts in the selected text. It deals with common concepts (P1) – definitions, which the students come into regular contact with in life and it is assumed that they know them. Further, it deals with expert concepts (P2) – definitions of the given field (for example, occupant, enlightenment, defenestration, and so on), facto-graphic (P3) – concepts identifying specific phenomena, things (own personal names, the names of human creations and organizations, the names of countries, nations, institutions, geographic and scientific names, abbreviations and symbols of the above stated), numerical (P4) – all numerical data, except for page numbers, assignments, pictures (for example, years, distance, calculations) and finally, repeated concepts (P5), into which we include words, which have already occurred in the text (if they appeared more than once, we count them as repeated concepts only once). The resulting concept of difficulty (Tp) is then calculated according to the following formula: **Tp = 100 x $\sum$P / $\sum$N x ($\sum$P1 + 3$\sum$P2 + 2$\sum$P3 + 2$\sum$P4 + $\sum$P5) / $\sum$N**, where P is the sum of

all concepts in the given formula, $\sum P$ is the sum of all concepts in all 10 samples and $\sum N$ is the sum of words in all 10 samples. Finally, the **density coefficient of expert information in the total number of concepts (h)** is calculated according to the h = 100 x ($\sum P2 + \sum P3$) / $\sum P$ formula and the **density coefficient of expert information in the total number of words (i)** according to i = 100 x ($\sum P2 + \sum P3$) / $\sum N$. For statistical comparison of the results, we used the extent of variability (standard deviation, coefficient of variation), Pearson's correlation coefficient, student's T-test (whether statistics differ in arithmetic average for textbooks of individual publishers) as well as the independence chi-square test.

Research Sample

The research sample consisted of 13 textbooks issued by 7 publishers (Alter, State Educational Publishing, Prodos, Nová škola, Didaktis, Fraus and SPL Práce). The selected textbooks of the specific publisher had to fulfill the following two criteria as a whole: the content of history curriculum from the beginning of history to the end of the 20th century and a valid approval clause by the Ministry of Education. Apart from textbooks published by Prodos, defined for the 4th and 5th grade combined, every publisher was represented by two textbooks (one for the 4th grade and one for the 5th grade of primary school). It is necessary to mention that we only measured the text describing the history of the Czech nation and the individual publishers as a whole, so 10 samples at 200 words were summarized for (possibly) both textbooks of the given publisher. Specific sections of text were selected so that the topics of the textbooks would be in compliance for every publisher. The following sections were selected: The Přemyslid princes, the Luxembourg Dynasty, the Hussite Era, George of Poděbrady and Jagiellonian dynasty, White Mountain and Habsburg Monarchy, Enlightenment (Marie Theresa and Joseph II), the Revolution of 1848, the First Republic (1918–38), World War II and Communism. These topics appeared in the textbooks of all publishers researched. The total number of researched text was 70 (10 from the textbooks of every publisher), 1 234 sentences and 14 467 words, to be specific.

Research Results

Objective 1: **to find out and compare the overall extent of the difficulty of text (T) of the history part of Czech textbooks for the subject of history (the Human and his World)**.

The average T value was 22,57 points. As a whole, textbooks from Prodos statistically show the highest above average value (from the arithmetic value of all textbooks of individual publishers – this is illustrated in graph no. 1 with the aid of standard deviation – vertical line). On the contrary, below average values of the overall difficulty of text were reached by history text of textbooks from SPL Práce. As stated by Průcha (2006), the values typical for textbooks of the 4th and 5th grades should be within the range of 22–24 points, which the majority of the textbooks, except for textbooks published by Prodos, fulfill. However, the value of the overall difficulty does not demonstrate extreme values even here. If we compare (at least in scope) the results of measuring of the difficulty of text in older history textbooks from the 1990s, where a history curriculum is incorporated, then there is a considerable decrease of the overall difficulty of text, on average by approximately 11,5% (the average value is 25,2, the highest value is 29,5 and the lowest value is 19,4 – see Průcha, 2006). Beneš, Janoušek, Novotný (2009) claim that an overall extent of difficulty lower than 20 is considered as easy. On the contrary, above 60, it is considered as inadequately difficult.
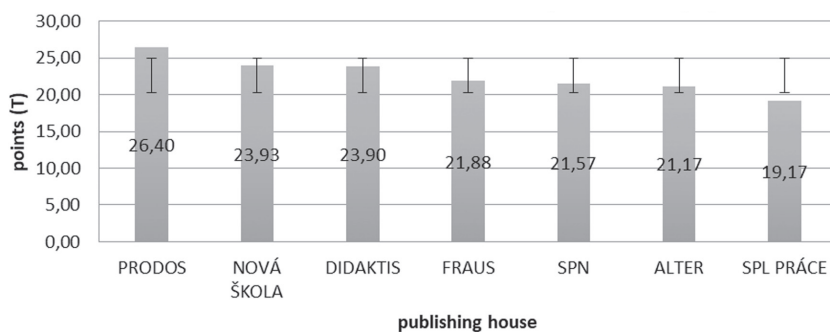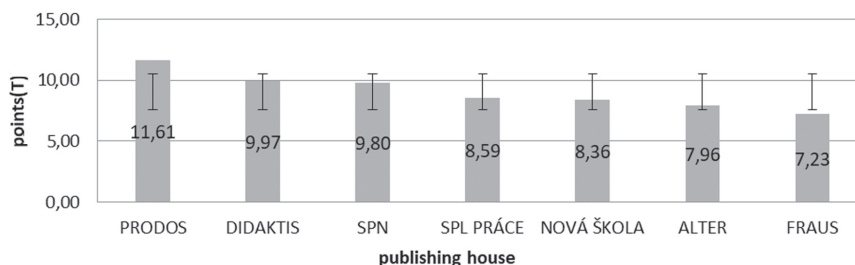

Graph 1.
Overall Extent of Difficulty of Text.
Source: own source.

Objective 2: **to find out and compare the extent of syntactic difficulty of text (Ts) of the history part of Czech textbooks for the subject of history (the Human and his World)**.

The summarized results are illustrated in the following graph.

Graph 2.
The Extent of Syntactic Difficulty of Text (Ts).
Source: own source.

The average value of syntactic difficulty reached a value of 9,07 points (SD = 1,37), which can be considered as a relatively low value, which is adequate to students at the first level in the 4th and 5th grades, for whom these textbooks are intended. Statistically above average values (within the researched file and a median value of 9,07) were reached by texts published by Prodos. On the contrary, statistically below-average values were reached by texts in textbooks published by Fraus. The highest variability inside the textbooks of the given publisher was measured for SPL Práce textbooks (SD = 2,14) and the lowest for textbooks published by Didaktis (SD = 1,17). Overall, it can be concluded that individual texts have a similar, not too different syntactic difficulty. This is also obvious from the point of view of the average length of the sentence and the average length of the sentence section (see objectives 2a and 2b below). However, the comparison of the syntactic difficulty according to the individual topics is interesting. The results are summarized in the following table. It is clear that there are significant differences between the processing of individual topics in the textbooks of various publishers. The biggest difference is seen in the topic of Jiří z Poděbrad/Jagellon. The most similar syntactic difficulty is then visible in the whole topic of the Lucemburk's Dynasty (see values of standard deviation – SD).

Table 1

Comparing the Extent of Syntactic Difficulty of Text among Individual Topics and Publishers

Ts According to Topic	DIDAKTIS	ALTER	Nová škola	FRAUS	SPN	PRÁCE	PRODOS	Average	Comparison	SD
1848 and the Revolution	9,85	*8,77*	11,05	9,31	**15,4**	12,82	13,79	11,57	↑	2,3
World War 2	10,84	*6,37*	10,3	9,18	**14,3**	11,43	9,89	10,33	↑	2,2
First Republic	11,9	*5,91*	11,01	7,29	11,8	10,88	**12,31**	10,15	↑	2,3
Communism	9,44	6,57	8,08	*7,85*	10,1	13,08	**13,29**	9,77	↑	2,4
Hussite Era	7,82	9,71	*5,84*	12,29	7,86	6,86	**13,83**	9,17	↓	2,7
Enlightenment – Marie Theresa and Joseph II	**12,13**	9,35	10,51	7,08	8,75	*6,64*	9,29	9,11	↓	1,8
George of Poděbrady / Jagiellons	7,74	7,94	6,18	9,25	9,18	*4,91*	**16,59**	8,83	↓	3,5
White Mountain and the Habsburg Monarchy	9,56	10,71	7,21	*4,67*	8,55	7,69	**12,12**	8,64	↓	2,3
Luxembourgs Dynasty	**10,1**	9,35	**10,1**	*5,05*	7,64	8,28	9,79	8,62	↓	1,7
Přemyslid Princes	**12,48**	7,23	7,33	*5,95*	9,2	9,15	8,76	8,59	↓	1,9

Source: own source.

Objective 2a–b: **to find out and compare the average length of sentences (V average) and the average length of the sentence section (U) of the history part of Czech textbooks for the subject of history (the Human and his World)**.

The average sentence length is 11,85 words. The results for textbooks of individual publishers are illustrated in the following table[1].

---

[1] ↑ – above average value, ↓ – below average value.

Table 2

Comparing the Average Length of Sentences and Sentence Section among Textbooks of Individual Publishers

	Average Sentence Length	Difference of Arithmetic Average (t-test, p value)	SD	Average Length of Sentence Section	Comparison with Arithmetic Average (t-test)	SD
PRODOS	14,49	↑ (0,003)*	2,12	8,05	↑ (0,072)	0,62
DIDAKTIS	12,63	↑ (0,052)	1,17	7,98	↑ (0,245)	0,85
SPN	12,02	↑ (0,519)	1,71	8,29	↑ (0,089)	1,04
SPL PRÁCE	12,01	↑ (0,504)	2,14	7,29	↓ (0,353)	1,03
NOVÁ ŠKOLA	10,89	↓ (0,163)	1,47	7,82	↑ (0,580)	0,97
ALTER	10,61	↓ (0,051)	1,42	7,56	↓ (0,775)	0,71
FRAUS	*10,32*	↓ (0,016)**	1,35	7,30	↓ (0,476)	1,32
Average	11,85			7,76		

*α = 0,01; ** α = 0,05.

Source: own source.

Only textbooks published by Prodos demonstrated significantly higher average lengths of sentences (t-test: p = 0,003)[2]. On the contrary, textbook published by Fraus on average contain significantly less words in one sentence than textbooks published by the remaining publishers (t-test: p = 0,016)[3]. Regarding the difference inside the individual samples of textbooks (publishers), the biggest difference in the average number of sentences is demonstrated by textbooks published by SPL Práce and Prodos. The lowest difference is then demonstrated by textbooks published by Didaktis and Prodos (see standard deviation). The researched textbooks do not differ in the average lengths of sentence sections (on average 7,76 words). The biggest difference in texts inside individual textbooks was in textbooks published by Fraus. The smallest difference was in textbooks published by SPN and SPL Práce. On the contrary, the most similar length of sentence section can then be found in textbooks published by Prodos and Alter.

---

[2] α = 0,01.

[3] α = 0,05.

Objective 3: **to find out and compare the extent of semantic difficulty of text (Tp) of the history part of Czech textbooks for the subject of history (the Human and his World)**.

Semantic difficulty of the researched texts ranges from approximately 10,5 to 15,5 points. Therefore, it forms the majority part of the overall difficulty of the text. The average value of all texts is 13,50 points. A statistically significant difference from the average of all researched texts are demonstrated by texts of textbooks published by Nová škola (Tp = 15,57 points). Statistically lower are then texts of SPN and SPL Práce textbooks (11,77, or 10,58 points) – these represent the lowest difficulty of text among the researched textbooks (see the following table – the line above the column means standard deviation). The total values of semantic difficulty are adequate to the textbooks for the second period of the first level of primary school. The differences in the textbooks published by Prodos, Fraus, Didaktis, Alter and SPN are not statistically significant from the total average.
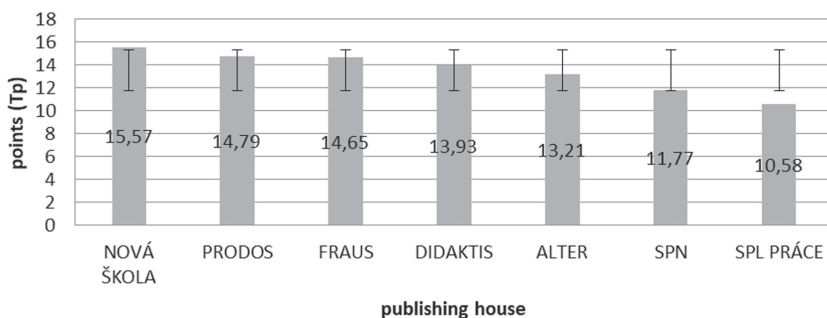
Graph 3.
The Extent of Semantic (Concept) Difficulty of Text.
Source: own source.

However, the average **semantic** difficulty of history text among textbooks of individual publishers differs. Statistically significant differences among the overall difficulty of history texts in textbooks of individual publishers are expressed in the following table (T-test, values: p)[4]:

---

4   +++ significant difference (higher) on the surface $\alpha = 0,001$, - - significant difference (lower) on the surface $\alpha = 0,01$.

Table 3

Comparing the Extent of Semantic Difficulty among Texts in Textbooks of Individual Publishers

Publishing house (Tp)	Nová škola	SPN	Práce	Alter	Fraus	Prodos	Didaktis
Nová škola (15,57)	x	+++	+++	0	0	0	0
SPN (11,77)	0	x	0	0	--	--	--
Práce (10,58)	0	0,3	x	0	--	--	--
Alter (13,21)	0,132	0,292	0,105	x	0	0	0
Fraus (14,65)	0,349	0,006	0,003	0,371	x	0	0
Prodos (14,79)	0,879	0,001	0,001	0,132	0,357	x	0
Didaktis (13,93)	0,132	0,004	0,002	0,478	0,714	0,178	x

Source: own source.

From the table, it is clear that texts from Nová škola textbooks reach statistically significant higher Tp values than texts of textbooks published by Prodos and SPL Práce. Therefore, compared to these textbooks, they are more difficult. On the contrary, history texts of textbook published by Prodos and SPL Práce reach statistically significantly lower values than texts of textbooks published by Fraus, Prodos and Didaktis. Therefore, they are less difficult. Especially for SPL Práce, the cause can be due to text written in the "narrative" form of a story as opposed to the traditional explanatory text. Also, text in SPN textbooks demonstrate a simple sentence structure, which by rule is the common language used by children. However, in general, none of the observed textbooks demonstrated extreme values, which would put it in a position of an overly difficult textbook for the given school level. However, it is evident that differences in text among individual textbooks of the researched publishers exist and it is up to the teacher to decide which textbook to use. Similarly as for syntactic difficulty the **individual texts differ (relatively significantly) in terms of semantic difficulty among individual topics**, as is illustrated in the following table. **Bold text** identifies the **highest** and the *lowest* (also in italics) values for the given topic whole. The topics are ranked from highest to lowest semantic difficulty (the average value for the given topic whole). **Colored text** identifies the **highest**, or the *lowest* differences inside individual topics (according to the standard deviation – SD values;

the higher the value the higher the difference in semantic difficulty of text in individual topic wholes among the researched publishers).

Table 4
*Comparing the Semantic (Concept) Difficulty among Individually Themed Units*

Topic/publishing house	Nová škola	SPN	PRÁCE	ALTER	FRAUS	PRODOS	DIDAKTIS	Average	SD
Hussite Era	12,92	15,23	**8,16**	**20,7**	16,9	19,45	12,13	**15,07**	**4,06**
World War 2	17,32	*11,94*	12,7	15,08	**16,42**	14,01	14,85	**14,62**	*1,78*
First Republic	16,09	*10,49*	14,89	12,41	**17,82**	14,04	16,25	**14,57**	2,31
Communism	17,91	12,28	**10,99**	14,01	12,42	**17,99**	16,02	**14,52**	2,62
1848 and the Revolution	17,72	*10,08*	10,89	**19,17**	10,99	18,41	14,22	**14,5**	3,63
White Mountain and the Habsburg Monarchy	14,75	*11,9*	13,14	13,71	13,57	**15,78**	15,2	**14,01**	*1,23*
Přemyslid Princes	14,15	12,07	12,98	*8,86*	13,96	**18,84**	13,06	**13,42**	2,76
Enlightenment – Marie Theresa and Joseph II	**17,32**	*9,16*	9,96	9,55	14,49	14,62	16,02	**13,02**	3,13
Luxembourgs Dynasty	13,79	11,61	**7,65**	11,37	13,34	13,34	**15,15**	**12,32**	2,25
George of Poděbrady/ Jagiellons	14,76	13,98	**6,34**	9,28	**17,78**	11,82	11,39	**12,19**	3,47

Source: own source.

Objective 3a–e: **to find out and compare the individual types of concepts**.
Semantic (or also concept) difficulty is expressed by the number of different types of concepts. Simply said, the more special concepts (expert, factographic and numerical) a text contains, the higher the semantic difficulty (Tp) is. From the following graph, it is clear that approximately 50% of substantive concepts (from which Tp is calculated) consist of common concepts, which are not specific for the given field (in our case history). The most can be found in texts of textbooks published by Alter (around 59%) and the least (approx. 48%) in texts of textbooks published by Prodos. We also compared the statistic numbers of individual types of concepts (with the aid of a chi-squared test)

and we discovered that overall, there is no significant difference in the number of concepts in texts of textbooks of individual publishers (the following graph illustrates approximately the same percentage frequency – difference in several small units of percent). There is one exception in common (general) concepts, where the chi-square test value was p = 0,009[5]. The numbers in the columns capture the absolute frequency per individual types of concepts in textbooks of the given publisher.
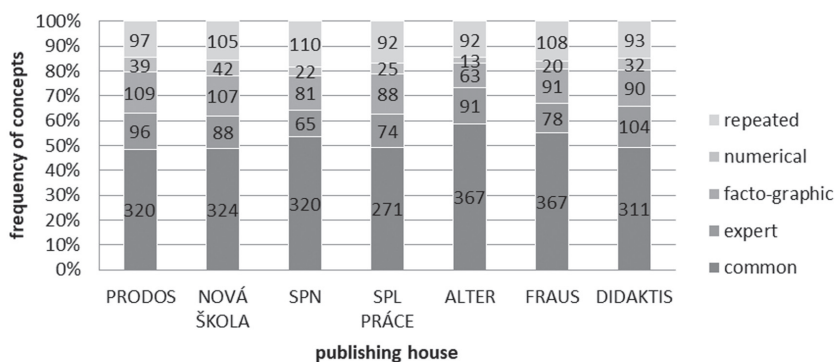
Graph 4.
The percentage distribution of individual types of concepts in texts of textbooks of the observed publishers.
Source: own source.

Comparing the number of concepts of individual types according to partial topic wholes is once again interesting. The results are presented in the following table. From the table, it is clear that regardless of the topic whole, they reach the highest number of common concepts (around 32 concepts). There are around 8–9 expert concepts as well as factual concepts appearing in the texts. On average, there are 10 repeated concepts. Numerical concepts have the lowest representation (on average, less than three in the researched text).

---

[5] $\alpha = 0.01$.

Table 5

*The Average Number of Individual Types of Concepts within the Topic Whole*

Topic Whole / typ of concept	common	expert	facto-graphic	numerical	repeated
Přemyslid Princes	30,57	6,14	11,71	2,29	12,29
Luxembourgs Dynasty	31,29	6,14	9,86	2	11,29
Hussite Era	33,57	12,14	7,29	1,14	10
George of Poděbrady/Jagiellons	30,43	9	6,71	1,71	11
White Mountain and the Habsburg Monarchy	29,29	9	9,14	5	9,86
Enlightenement – Marie Theresa and Joseph II	37,14	9	4,29	3,14	8,29
1848 and the Revolution	35,29	9,57	6,57	3,43	9,43
First Republic	32	5,57	13,86	3,14	10,71
World War 2	29,29	7,71	14,71	2,57	9,29
Communism	36,86	10,86	5,71	3,14	7,43
Avarage ($\bar{X}$)	**32,57**	**8,51**	**8,99**	**2,76**	**9,96**
Standard Deviation (SD)	**2,82**	**2,03**	**3,34**	**1,02**	**1,37**
The Coefficient of Variation (V)	**8,67%**	**23,80%**	**37,13%**	**37,04%**	**13,76%**

Source: own source.

If we compare the **variability** of data, using the variation coefficient which eliminates differing absolute values, than the greatest differences can be seen in factual and numerical concepts. The number of repeated and common concepts differs only slightly. Their numbers in the individual topic wholes are very similar (the higher the variation coefficient value, the higher the difference in the number of the types of concepts – in the texts researched for topic wholes).

Objective 4 and 5: to find out and compare the density coefficient of expert information in the total number of concepts (h) and the information coefficient in the total number of words (i) of the history part of Czech textbooks for the subject of history (the Human and his World).

The density coefficient of expert information for the observed texts of individual publishers is summarized in the following graph. The density coefficient of expert information in the total number of words reaches an average value of **8,46%**, in the total number of concepts it is **27,86%**. For comparing these coefficients, there is no official methodology. Therefore, generally, it is necessary to only emerge from comparing similar research. For science textbooks at the second level of primary school (Hrabí, 2005), the value is in the range of 23–33%, the h value is 64–76%. Smutková (2012) demonstrates somewhat lower values for their science textbooks – i = 23–27%; h = 55–66%. Pluskal (1996) claims that the i coefficient for the second level is around 20–25% and the h coefficient is between 40–50%. From this, we can conclude that the values for the first level of primary school will be lower. Our research (Šimik, 2014) of science textbooks for the first level demonstrated similar values (i around 8%, h between 18–35%). In comparing this research, the coefficients (i, h) of expert information can generally be considered as adequate for textbooks on the first level of primary school. The density coefficient of expert information in the number of concepts demonstrates a greater balance between the individual textbooks or publishers than science textbooks. A high correlation (r = 0,841) was confirmed between both coefficients. The higher the density of expert information in all concepts, the higher it also is in all words.
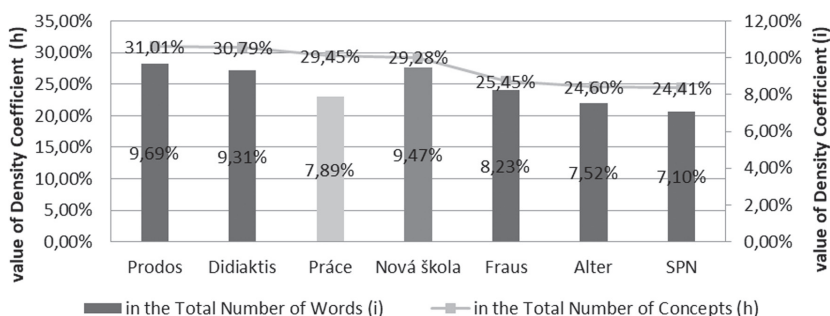
Graph 5.
Density Coefficient in Expert Information (i, h) – in the Total Number of Words, in the Total Number of Concepts.
Source: own source.

We also monitored how the overall difficulty of the text influences its individual components (and also whether individual components of difficulty of text correlate among themselves). For the calculation, we used the Pearson's correlation coefficient. From the following table, it is clear that there is a high correlation between the overall difficulty of text and the density coefficient of expert information in the total number of words. This proves that the more concepts a text contains the more difficult is it. Concepts (semantic) difficulty demonstrates a higher influence on the value of the overall difficulty than syntactic difficulty (however, even this is higher than 0,600). There is a weaker correlation (r = 0,536) between the overall difficulty of text and the density coefficient of expert information in the total number of concepts. It is possible to read other correlations among the individual components from the table.

Table 6

Comparing the Correlation Coefficients for Partial Components of the Overall Difficulty of Text

Pearson's Correlation Test	T-i	T-tp	T-Ts	T-h	h-i	Tp-i	Ts-h	Ts-i	Tp-h	Tp-Ts
r	**0,824**	**0,779**	**0,654**	0,536	**0,841**	**0,732**	0,554	0,429	0,248	0,035

T – total difficulty of text, Tp – semantic difficulty, Ts – syntactic difficulty, h – density coefficient of expert information in the total number of concepts, i – density coefficient of expert information in the total number of words.
Source: own source.

Conclusion

Measuring the difficulty of history text in history textbooks (history part of the educational section of the Human and his World), we discovered that the overall difficulty ranges from 19–26 points, which can be considered as low difficulty. According to these results, all textbooks are adequate to the age of the student which they are intended for, in terms of the difficulty. The highest level of difficulty was demonstrated by textbooks published by Prodos. The lowest level of difficulty was demonstrated by textbooks published by SPL Práce. The textbooks of other publishers statistically do not significantly differ from the overall average (approx. 22,5 points). Difference – and relatively high – can be identified when comparing partial topics presented by textbooks of individual publishers, both in semantic and syntactic fields of

difficulty. The differences, even more than double, are relatively frequent. The reason will most likely be the diversity of the author's approach, since the author (even though it was dealing with the same topic), by rule, viewed it from a different perspective (see tales 1 and 4). The average length of a sentence was less than 12 words and the average length of a sentence section was less than 8 words. Above average values were reached by texts from Prodos textbooks. On the contrary, significantly below-average values were reached by texts from Fraus textbooks (shorter and less complicated sentences). A higher difficulty of text was demonstrated by topics from modern history (19th and 20th century) and topics of the Hussite Era. Common (general) definitions were predominant among the concepts – almost around 50% of all substantive concepts. On average, we identified around 10–15% of expert concepts from all substantive concepts. Relatively many (up to 1/5) were repeated concepts, which appear to be a typical feature for primary school textbooks (repeated concepts). The density coefficient of expert information in the total number of concepts as well as in the total number of words was low (approx. 28%, or 8,5%). Overall, the difficulty of the text corresponds with the values for textbooks at the first level. Semantics has a higher influence on the overall difficulty of text than syntactic difficulty. However, the correlation of both is relatively high (around 0,700). The stated results indicate that all researched textbooks are similar as a whole. Partial differences appeared among the individual topics. In the matter of selecting the specific textbook, other factors, than the difficulty of text (for example, its graphics, the issue of how well it is equipped from the didactics point of view, activation component – types of textbook assignments or the price) will most likely be the deciding factors.

References

Beneš, P., Janoušek, R., Novotný, M. (2009). Hodnocení obtížnosti textu středoškolských učebnic. *Pedagogika*, *3*.

Červenková, I. (2011). *Užívání učebnic v činnostech žáků na 2. stupni ZŠ*. Olomouc: Palacky University.

Dvořák, D. (2005). Učebnice: nepřátelé, nebo spojenci? *Učitelské noviny*, *41*.

Greger, D. (2004). Učebnice jako realizační scénář kurikula. In: E. Walterová et al. (eds.), *Úloha školy v rozvoji vzdělanosti*. Brno: Paido.

Greger, D. (2005). Proces schvalování učebnic v historicko-srovnávací perspektivě. *Pedagogická orientace*, *3*.

Hrabí, L. (2005). Učebnice přírodopisu a jejich obtížnost. *Pedagogická orientace*, *3*.

Jakubcová, P. (2012). *Analýza didaktické vybavenosti učebnic občanské výchovy pro 6. a 7. ročník základní školy*. Diplomová práce. Brno.

Maňák, J. (2007). Učebnice jako kurikulární projekt. In: J. Maňák, P. Knecht (eds.), *Hodnocení učebnic*. Brno: Paido.

Maňák, J. (2008). Funkce učebnice v moderní škole. In: P. Knecht et al. (eds.), *Učebnice z pohledu pedagogického výzkumu*. Brno: Paido.

Mikk, J. (2007). Učebnice: budoucnost národa. In: J. Maňák, P. Knecht (eds.), *Hodnocení učebnic*. Brno: Paido.

Nestler, K. (1974). Zur Wirkung lexikalischer und syntaktisher Textmerkmale auf das Lernergebnis. *Informationen zu Schulbuchfragen*, *17*.

Pluskal, M. (1996). Zdokonalení metody pro měření obtížnosti didaktických textů. *Pedagogika*, *1*.

Průcha, J. (1989). *Teorie, tvorba a hodnocení učebnic*. Praha: ÚÚVPP.

Průcha, J. (1997). *Moderní pedagogika*. Praha: Portál.

Průcha, J. (1998). *Učebnice: teorie a analýzy edukačního média: příručka pro studenty, učitele, autory učebnic a výzkumné pracovníky*. Brno: Paido.

Průcha, J. (2006). Učebnice: teorie, výzkum a potřeby praxe. In: J. Maňák, D. Klapko (eds.), *Učebnice pod lupou*. Brno: Paido.

Průcha, J. (2008). Možnosti výzkumu učebnice ve vztahu k učení. In: P. Knecht et al. (eds.), *Učebnice z pohledu pedagogického výzkumu*. Brno: Paido.

Sikorová, Z. (2007). Návrh seznamu hodnotících kritérií pro učebnice základních a středních škol. In: J. Maňák, P. Knecht (eds.), *Hodnocení učebnic*. Brno: Paido.

Šimik, O. (2014). *Učebnice přírodovědy pohledem pedagogického výzkumu*. Ostrava: Ostravská univerzita.

Smutková, T. (2012). *Srovnání učebnic přírodopisu v České a Slovenské republice*. Diplomová práce. Brno.

Syntaktyczna i semantyczna trudność tekstu w historycznej części podręczników w obszarze edukacyjnym „Człowiek i jego świat"

Streszczenie: W artykule przedstawiono wyniki badań dotyczących składniowej i semantycznej trudności tekstu w czeskich podręcznikach historii na pierwszym poziomie szkoły podstawowej. Zbadano łącznie 13 podręczników z 7 wydawnictw. Dla analizy użyliśmy metody Nestler-Průcha-Pluskal. Wyniki pokazują, że całkowita trudność tekstu mieści się w skali od 19 do 27 punktów i zawiera ocenę odpowiednich trudności tekstowych dla danego roku. Składniowe trudności wahają się w przedziale od 7 do 12 punktów, a średnia długość zdania to około

12 słów. Semantyczna trudność tekstu jest nieco wyższa (między 10 a 16 punktów) w porównaniu ze składniową (około 9 punktów). Zidentyfikowaliśmy większe różnice w semantycznej trudności tekstu. Do istotnych statystycznie różnic należą te między poszczególnymi jednostkami tematycznymi. Tematy związane z historią nowożytną (XIX i XX wieku) wydają się trudniejsze. Współczynnik gęstości informacji zawodowej w ogólnej liczbie pojęć wynosi około 27%, co jest widoczne w stosunkowo niskiej trudności w porównaniu z innymi badaniami. Jest to jednak wystarczające dla danego poziomu podręczników (dla 4 i 5 klasy).

Słowa kluczowe: składniowa i semantyczna trudność tekstu, podręcznik do nauki historii, szkoła podstawowa

Syntactic and Semantic Difficulty of Text in the History Part of Textbooks for the Educational Field of the „Human and His World"

Summary: The article presents the results of the research regarding the syntactic and semantic difficulty in Czech history textbooks for the first level of primary school. We examined a total of 13 textbooks from 7 publishers. The Nestler-Průcha-Pluskal method was used for analysis. The results show that the overall difficulty of text ranges between 19–27 points, which can be evaluated as adequate text difficulty for the given year. Syntactic difficulty oscillates between 7–12 points and the average length of a sentence is around 12 words. The semantic difficulty of text is somewhat higher (between 10–16 points) compared to the syntactic (around 9 points). We identified greater differences in the semantic difficulty of text. The statistically significant differences are among the individual themed units. Topics related to modern history (19th and 20th century) appear to be more difficult. The density coefficient of expert information in the total number of concepts ranges around 27%, which is seen as relatively low difficulty in comparison to other research. It is however adequate to the given (4th and 5th) grade.

Keywords: syntactic and semantic difficulty, history textbook, primary school