Knowledge Base Development for Second Language Learning in the 3D Virtual Space

István Károly Boda, and Erzsébet Tóth, *Member, IEEE*

***Abstract*—In our study first we provide a short overview of the 3D virtual library project which started about ten years ago as part of the Cognitive Infocommunications (CogInfoCom) research. The current implementation of the virtual library model exploits the 3D features of the MaxWhere Seminar System. In our study we would like to summarize the classroom experiences that we brought together in teaching English as a second language for students of computer science majors at the Faculty of Informatics, University of Debrecen in the academic year of 2020 and 2021. Our main purpose was to improve the students' linguistic knowledge and to collect their opinions and suggestions about the content of the learning material of our virtual library and where they think is necessary to modify it. Summarizing their views, we decided to add further vocabulary items, contexts and explanations to the material. In addition, their achievements proved that successful language learning needs carefully prepared tests and exercises which support self-assessment and increase motivation. In general, the more tests are available the more efficient the learning process is. But preparing good and varied tests manually is a relatively slow and exhausting work. Therefore, we intended to use JavaScript technology to develop an algorithm which can generate tests and exercises automatically, based on the knowledge base of the virtual library.**

***Index Terms*—automatic test and exercise generation, CogInfoCom, MaxWhere Seminar System, second language learning, three-dimensional (3D) virtual library project, virtual learning environment**

I. Introduction

ABOUT ten years ago we launched a *virtual library project* as part of the cognitive infocommunications (CogInfoCom) research [1-2], focusing on the presentation of selected library content in virtual 3D environment the features of which have been thoroughly investigated by several CogInfoCom studies since then (e.g. presentation of virtual rooms and buildings in the 3D space [3-4], developing and using effective 3D learning environment [5-6], cognitive and psychological aspects of the 3D environment [7-8] etc.). The main objective of the virtual library project was to collect, organize and present relevant verbal and multimedia content in the 3D space about the ancient Library of Alexandria and classical Greek literature translated into English (e.g. texts about the life and work of Callimachus, translations of selected literary texts from the works of prominent ancient authors etc.) [9-11]. Although the 3D virtual library is designed for various purposes, *language learning* has proved to be the most promising application of the collected virtual library content [12-13] because of its potential usefulness and the importance of advanced language skills (in our case, in English) to convey the ancient cultural heritage reflected in the collected material to today's cultural environment.

The literary and cultural heritage of antiquity and especially the collection of the ancient Library of Alexandria are usually referred to as a symbol of universal human knowledge and wisdom. Its catalogue covered all Greek literary works which were once available in its holdings, and the catalogue was used as a bibliography of ancient Greek literature for centuries. The 3D virtual library model of our project (3DVLM) is based on the classification system of the ancient Library of Alexandria, especially on the famous Pinakes invented by the great scholar-poet Callimachus in the 3rd century BC [12].

The current implementation of the 3D virtual library model makes use of the innovative 3D environment of the MaxWhere Seminar System [14] the features of which, and especially the embedded web browsers called smartboards, fully support the implementation of the basic concepts of the 3DVLM [15-17]. The core content of the knowledge base of the virtual library focuses on the classical heritage the European culture is based on. Our main idea is to revive and convey the message of ancient times to the present-day culture, which seems to be really crucial with respect to the young social generations. We firmly think that this mission can be best accomplished by language learning which, along with its obvious importance in our internet-based society, can serve as a bridge between the ancient times and the "modern" culture of the XXI. century. The basic idea of the virtual library project is that with a carefully elaborated way and methodology the ancient thoughts and values can effectively be translated (both literally and figuratively) to the young members of the generation CE (i.e. the generation cognitive entities whose members are growing up in "entangled co-evolution with ICT" [2])

The index page of the current implementation of the virtual library project begins with a 2D map of the virtual library which shows and briefly explains the main entry points to the content of the library (Fig. 1).

István Károly Boda is with the Department of Mathematics and Informatics, Debrecen Reformed Theological University, Debrecen, Hungary (e-mail: boda.istvan@drhe.hu).

Erzsébet Tóth is with the Faculty of Informatics, University of Debrecen, Debrecen, Hungary (e-mail: toth.erzsebet@inf.unideb.hu).

DOI: 10.36244/ICJ.2024.5.8

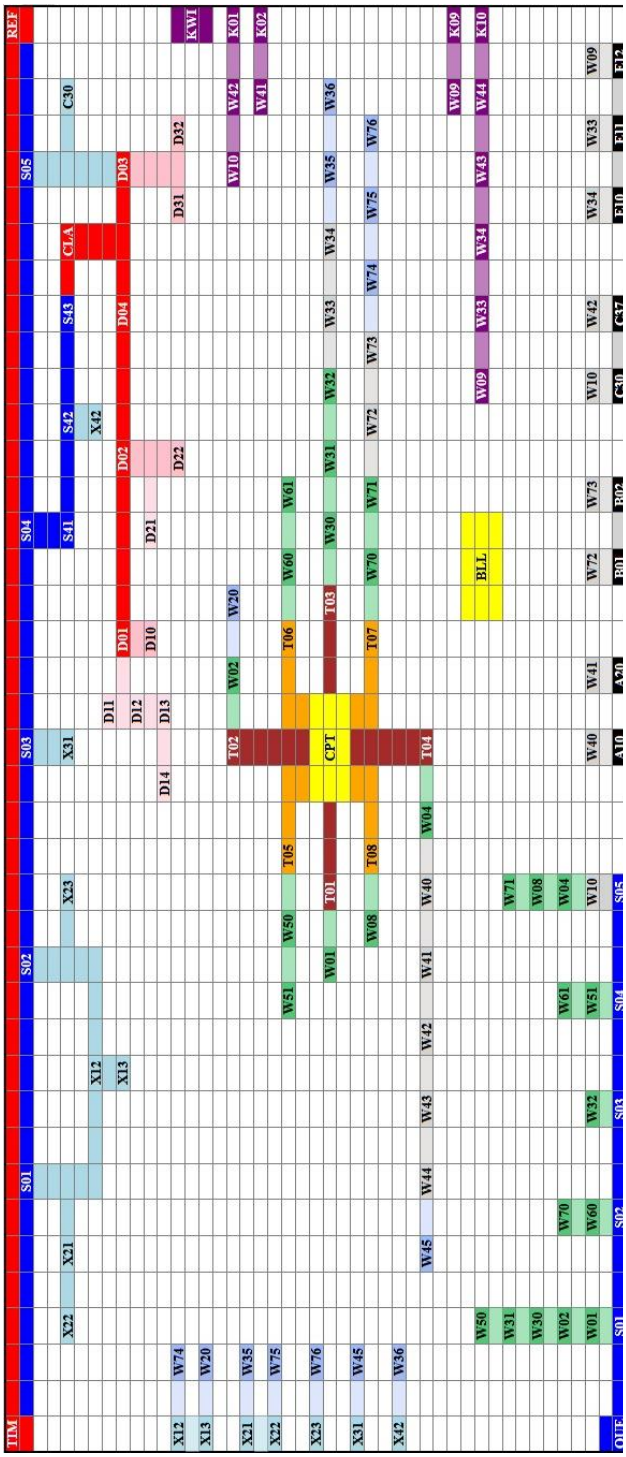

Fig. 1. The 2D map of the current implementation of the virtual library [xix]. Each abbreviation in the map refers (and is linked) to a corresponding part of the virtual library; e.g. clicking on [T06] (about in the centre) we can go to the thesaurus page described in detail in Section III.

In the following section a brief account will be provided how we were teaching English as a second language for the students of the Faculty of Informatics, University of Debrecen on the basis of the preprocessed texts and supporting material delivered by the knowledge base of our virtual library.

II. A SUMMARY OF CLASSROOM EXPERIENCES

In the academic year of 2020 and 2021 we could teach the compiled learning material of our virtual library for students of the Faculty of Informatics, University of Debrecen at Bachelor level. Our main objective was to develop the students' linguistic competence and to collect their opinions and suggestions about the content of the learning material (e.g. where they think is necessary to correct or modify it, where to add further comments, vocabulary items, explanations etc. to it). In the autumn semester an optional university course was offered for the students in a blended learning form. The only admission criterion for students to this course was to have a successful placement test which measured their skills of using grammatical rules properly in different contexts at B2 level. Using this test, we could quickly select those students from the 120 candidates for the course whose language skills were satisfactory (i.e. at upper-intermediate or advanced level). Finally, we had 60 students altogether who were divided into three different groups for teaching. Although in the beginning we could deliver our courses in a traditional way where a face to face contact was possible between teachers and students at this time of the semester, later, because of the world-wide pandemic, we had to change the way of teaching from traditional to online starting from 9th November, 2020.

Below we provide a short overview of the activities we planned, the good experiences and practices we had, and the main difficulties we had to deal with during the courses. We **emphasize in bold the most common problems** we faced and the suggestions or observations the students made to handle them.

A. Teaching environment and the use of information technology

The lessons for the two groups of students were placed in the same big lecture hall where we could use only a projector and the blackboards as teaching aids. But the lessons for the third group were located in a smaller classroom which was equipped with the same teaching aids. In this traditional learning environment, we decided to present the two-dimensional version of the learning material for the students which was the same as the three-dimensional one available in the 3D Castle space of the MaxWhere VR platform.

B. Texts taught during the course

According to the time schedule of the course syllabus, during 6 weeks we could focus on the learning material about Callimachus and the Library of Alexandria in two teaching hours (i.e. 2*50 min) per week. Below we provide the list of the primary and secondary texts from our virtual library that we were teaching for the students during the lessons. (The codes that identify the texts can be seen in Fig. 1).

- Callimachus [S01],

  Cyrene, Libya [X12],

  The Dorians [X13],

  The Ptolemaic dynasty [X21],

  The Ptolemaic Kingdom [X22],

- Ancient Greek Literature [CLA],
  The nine lyric poets [D21],
- Callimachus in Alexandria [S02],
  Alexandria [X23],
- The Great Library of Alexandria [S03],
  The Mouseion [X31],
- The Pinakes [S04],
  The collection of the Great Library [S41],
  Callimachus' classification system [S42],
  Binomial nomenclature [X42],
  The structure of the Pinakes [S43],
- The works of Callimachus [S05].

C. Text comprehension task

In the case of the S01, X12, X13, X21, X22 content units we could practice text comprehension by answering to various questions which were closely related to the content of the given texts. In the text comprehension tasks we perceived that the students could easily follow and understand the main topic of the complex texts in general. However, they often failed **to explain the precise meaning of unknown words or phrases, or to recognize new verb or collocation patterns** emerging in various contexts.

D. Vocabulary Teaching

We often made use of the vocabulary items integrated into the learning material. Those items are presented for the students in order to help them to understand better and memorize the **spelling, pronunciation and definition of selected words** and their related term(s) as well. While we were reading and processing the texts about Callimachus, the students were asked **to mention or provide a synonym or antonym of, to define, or paraphrase the meaning of selected words or expressions** [18]. We learned that it was the most difficult task for them. In case we explored a new word, expression or collocation in the text we were studying with the students, we suggested that they should learn and memorize it 'as is'.

E. Collecting opinions from the students

We distributed a small paper among students towards the end of the lessons to collect their opinions or proposals about the content of the learning material by posing them the following question: *"Which words, collocations or grammatical structures caused you any difficulties in the learning material?"* According to these feedbacks the students had a big **problem with the pronunciation and spelling of proper names** occurring in the texts. For example, they asked us to add new vocabulary entries to the learning material about the spelling and pronunciation of the name of the major ethnic groups in Classical Greece (not to mention historical figures, mythological characters, geographical names etc.).

F. Preparation for the lessons

In advance, we prepared a lot for these English lessons and looked for the precise *definition* of each new word and its most frequently used collocations in a monolingual dictionary [e.g. 19] which seemed to be relevant in the context. In several cases we also chose *sentence samples* or, in general, *concordances* (from various sources, e.g. from literary works, monolingual or quotation dictionaries etc.) which conveyed the various shades of meaning of the given word in different contexts. There were cases when we had to provide further explanations for the students about the meaning of *special terms*, e.g. those used for describing historical periods, art periods etc. ("Archaic Greece", "Old Comedy", "New Comedy", "Hellenistic Period" etc.) From time to time, we also checked the *grammar* [22] used in the studied texts if further grammatical explanations are required for the students.

G. Integration of concordances into teaching

Concordances can provide a simple and valuable help to teach real English in the classroom [20-21]. We selected a lot of concordances that we thought might be helpful for the students, and organized them around specific keywords in separate *thesaurus pages*. Although it is not very easy to find the best method for integrating concordances successfully into teaching, we presented the content of the compiled thesaurus pages to the students and explained the page structure for them putting special emphasis on the importance of concordances. In case a concordance had an obscure or ambiguous part (with respect to meaning, grammar, style, imagery etc.) we provided special *explanatory notes* for the given concordance in the same format which we used for the main content units of the learning material. Note that the notes often contain references to other concordances and explicit concordances as well.

The explanatory notes attached to certain phrases in the S05 content unit proved to be very useful for the students. Using these notes, the students could understand better those collocations which included abstract concepts such as „*an aesthetic of smallness and perfection*", „*his expression of what constituted excellence*", „*drive their wagons*" etc. In these explanatory notes we used concordances in order to support the language learners to understand and memorize better the meaning of certain collocations and broaden their vocabulary at the same time. Note that only at this point of the learning material we could easily integrate concordances into teaching. As a consequence, we found that concordances, which play an important role in activating language, would principally be associated with a given context.

H. Assessment of students' language competence

After 6 teaching weeks the students wrote a test paper on the vocabulary and grammar of the learning material in a traditional form. The so-called "word test" contained new words, collocations, parts of sentences or whole sentences that the students had to translate from English to Hungarian and vice versa. They had to provide English definitions of five terms that were selected from the studied texts of the learning material. The students found this exercise a bit difficult, but they could cope with it successfully. In addition, they had to solve a grammatical exercise related to advanced English grammar [22]

The students found this exercise a bit difficult, but they could cope with it successfully. In addition, they had to solve a grammatical exercise related to advanced English grammar [22] which occurred in the studied text about ancient Greek literature [CLA].

On the whole, we can conclude that the majority of the students could acquire successfully the vocabulary and the grammar of the primary and secondary texts of the virtual library because a relatively small proportion (15%) of them received unsatisfactory grade for their test paper. (Note that those students who failed at first could later try again to improve their results.). Due to the relatively small number of students we consider the students as one homogenous group and did not examine the test results according to different aspects. As regards the further assessment of the learning material, we plan to use the Google Translate service [26] as an AI agent that can be metaphorically considered as a virtual language learner at an intermediate (or advanced) level. GT can translate selected parts of the learning material and the evaluation of the results can provide unlimited number of empirical data which we can use continuously to assess and improve our learning material.

On the basis of the students' opinions and suggestions (e.g. on the spelling and pronunciation of proper names, the explanation of special terms etc.) we intended to **add further vocabulary items and explanations to our learning material** to make it more usable and understandable for those language learners who use our material to improve their English (preferring either classroom or online learning environment, or independent learning).

III. Content Development for Classroom Use

On the basis of our classroom experiences we decided to further develop the knowledge base of the virtual library in order that it could be much more effectively used in language learning and teaching. Because the use of thesaurus pages, and especially concordances caused considerable difficulties for students we focused on the full revision and development of the structure and content of **thesaurus pages**. In the following, we would like to describe a selected page in detail on the one hand, and explain the applied page elements and tools on the other hand.

Thesaurus pages are organized around given *microcontexts*, where the term 'microcontext' means a group of semantically related words (i.e. synonyms, antonyms, related or contrasted words) that fits into a given grammatical or collocation pattern. We present the words in a way following the structure of well-known thesauri, synonyms dictionaries and lexicons [i-vi] where most of the related words were selected from (creating several subgroups, sometimes in two or three hierarchical levels, according to the meaning of the words). The collocation patterns, in turn, can be found in collocations dictionaries [vii]. The microcontexts are illustrated by carefully selected concordances, similarly to the example sentences of the entries of monolingual dictionaries [e.g. vii-xii] where most of the concordances were selected from, although we used production and quotation dictionaries as well [xiii-xv] and various literary and other texts (e.g. [xvi-xvii]). Here we use the general term 'concordance' because in our case concordances include not only selected sentences, but quotations or short extracts from various texts as well which we think can be useful for language learners to develop their vocabulary and related skills (e.g. grammar, composition, style etc.).

We selected the thesaurus page entitled [T06] which is related to the keywords *account, description, report; narrative, story; article, document, essay, review, study* etc. The page is organized around the collocation pattern [**adj**+noun] where the possible nouns (i.e. the keywords which can be placed into the "noun" part of the pattern) can be seen in Fig. 2.


[G154] paper, *document*, *record*, register; ⇒ minutes, notes;
[G155] letter, note, label; ⇒
[G156] notebook, excercise book, folder, file; ⇒
[G157] *account*, *description*, *report*, version; ⇒ *testimony*;
[G157] *story*, *tale*, *anecdote*; ⇒ *narrative*;
[G158] *essay*, composition, *paper*; ⇒ (piece of) *work*, *study*, *survey*;
→ [G65] speech, talk; lecture; ⇒ presentation;
[G159] *article*, feature, sketch, editorial, headline; ⇒ *review*;
[G160] material, copy, print; ⇒
→ [G242] *text*; ⇒
[G161] summary, synopsis, abstract, abridgement, précis, brief; ⇒
[G162] book, manuscript, scroll, script, draft, textbook, manual, diary; ⇒

Fig. 2. The list of keywords used in thesaurus page [T06].

Some of the possible adjectives (which can be placed into the "adj" part of the pattern) can be seen in Fig. 3. For example, valid combinations are *"accurate account"*, *"apt description"*, *"comprehensive report"* etc.

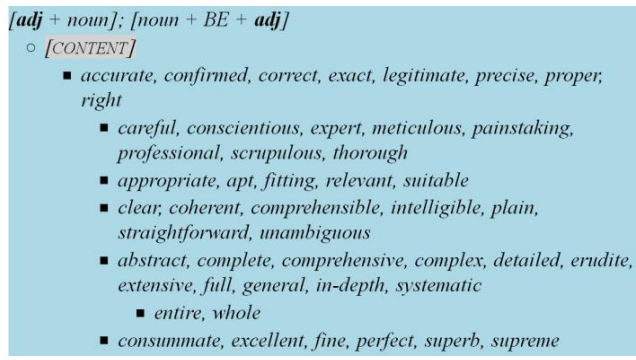
***[adj** + noun]; [noun + BE + **adj**]*
- *[CONTENT]*
  - *accurate, confirmed, correct, exact, legitimate, precise, proper, right*
    - *careful, conscientious, expert, meticulous, painstaking, professional, scrupulous, thorough*
    - *appropriate, apt, fitting, relevant, suitable*
    - *clear, coherent, comprehensible, intelligible, plain, straightforward, unambiguous*
    - *abstract, complete, comprehensive, complex, detailed, erudite, extensive, full, general, in-depth, systematic*
      - *entire, whole*
    - *consummate, excellent, fine, perfect, superb, supreme*

Fig. 3. The first part of the list of adjectives used in thesaurus page [T06].

At the end of the thesaurus page [T06] all the available combinations are listed in two tables. In the first table the collocations are ordered by nouns (e.g. 'account', 'anecdote' etc.; see Fig. 4), and in the second table the collocations are ordered by adjectives (e.g. 'accurate', 'ancient' etc.). Note that currently there are 18 nouns (including plural forms) and 221 adjectives which make possible, at least theoretically, 3978 combinations of the corresponding words according to the given collocation pattern. In the thesaurus page there are currently 58 combinations that are validated (and illustrated) by selected concordances.

COLLOCATION PATTERNS (ORDERED BY NOUNS)

adj	noun	frequency
account, accounts		
accurate	account	1
contradictory	accounts	1
detailed	account	2
detailed	accounts	1
different	accounts	1
firsthand	account	1
vivid	account	1

Fig. 4. The collocations of the word 'account' listed in the first table at the end of the thesaurus page [T06].

In the thesaurus page we arranged and grouped the selected concordances according to the adjectives used in the applied collocation pattern (formatted in bold font type). In the concordances the keywords were emphasized using small-caps font variants. In some cases we presented the context of the concordance as well, inserting a short text immediately after the concordance. Where a rare or possibly "difficult" word occurred in the concordance (denoted by an asterisk) we added a link to a vocabulary entry at the end of the thesaurus page explaining the meaning of the given word (see Fig. 5.1-5.2).

[accurate]
accurate description = a description which is correct, exact and precise in all details

- *At a committee meeting we need to have an* ***accurate*** ACCOUNT *recorded * of our discussions.*

  Have you ever been to a ***committee meeting****? An important thing has to be settled at the very outset: who is going to take the minutes? We need to have an accurate account recorded of our discussions and sadly our memories are not always that reliable so we have someone take minutes. These can later be consulted so that we can follow through on the decisions we've taken and appropriate action can be taken.*
- *Quantum mechanics has to be used to provide an* ***accurate*** DESCRIPTION *of the microscopic system.*
- *A more* ***thorough****,* ***accurate*** DESCRIPTION *of the manuscript is under preparation.*
- *He has always kept an* ***accurate*** RECORD *of his spending.*

Fig. 5.1. The concordances containing the adjective *'accurate'* in the given collocation pattern.

[4] *record* [rɪ'kɔ:d] an account or facts = to set down * an account (of a meeting etc.) or facts (about historical events etc.) in writing or some other permanent form for later reference

- *An official report* ***records*** *that at least half the nation's monuments are in need of repair.*
- *At a committee meeting we need to have an accurate account* ***recorded*** *of our discussions.*
- *An authentic document* ***records*** *that the battle took place six years earlier.*

Fig. 5.2. The vocabulary entry explaining the verb *'record'* by a short definition and several concordance samples.

In the learning process it is strictly necessary for a student to understand the structure and aim of the learning material in order to (repeatedly) read and memorize its content. But to memorize almost 60 concordances certainly needs additional support. Therefore we created a semantic map of the most important collocation patterns and concordances which shows the semantic relationships of the patterns and helps visually memorize them (Fig. 6).

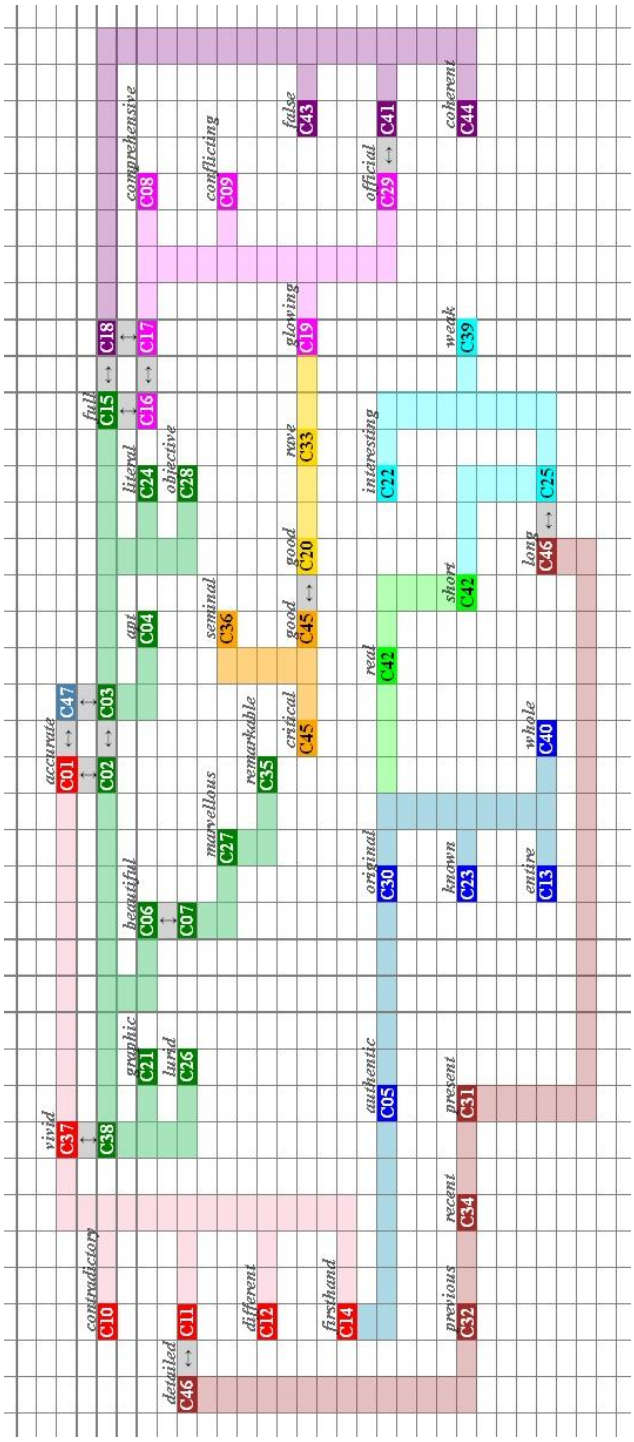

Fig. 6. The semantic map of the most important collocations represented in the thesaurus page [T06]. Each abbreviation refers (and is linked) to a specific concordance in the page; e.g. clicking on [C46] (labeled with the adjective 'long') we can go to the concordance "*He made a* ***long*** *and detailed study of how animals adapt to their environment.*" [x].

Note that we deliberately used different colors for different keywords in the patterns (e.g. red for 'account', green for 'description', etc.). We linked the patterns which contained either the same or syntactically related adjectives (i.e.

synonymous / antonymous or related / contrasted words as in thesauri), e.g. 'beautiful' ~ 'marvellous' ~ 'remarkable'; 'authentic' ~ 'original' ~ 'real'; 'long' ~ 'short' etc.

Finally, we created 50 **tests** for the students to practise their skills and improve the knowledge they have acquired. Note that doing tests repeatedly is also a very effective way of *learning and memorizing* the content provided by the tests.

Every test is based on a given concordance but the adjective and noun parts of the concordance are replaced by 4-4 randomly generated words in addition to the "valid" words which provide the right solution of the test (thus a student should choose the correct words from 5 options for each variable part of the sentence). So there are fixed and variable parts of a given sentence in a test, and sometimes, for either grammatical or didactic purposes, there are more than two variable parts (e.g. a student should choose the right article, the correct form of the verb etc.). We tried to create the algorithm which generates the random tests that as many combinations appear to be more or less acceptable as possible. We can always find the offered solution of every test clicking on the OK button at the end of the sentence (see Fig. 7-8).

1. TEST

At a committee meeting	*we need to have*	*– a an*	*accurate comprehensive fair inadequate introductory*	*account anecdotes essay story works*	*recorded of our discussions.*	**ok**

Fig. 7. The test based on the first concordance which can be seen in Fig. 4.

1. TEST

At a committee meeting	*we need to have*	*– **a** **an***	***accurate*** *comprehensive fair inadequate introductory*	***account*** *anecdotes essay story works*	*recorded of our discussions.*	**ok**

Fig. 8. The solved test which can be seen in Fig. 7.

Note that, in addition to grammar rules, we need to get realistic alternatives for the words which appear in the variable parts of the sentences. In this respect, the semantically related nouns (keywords) and adjectives are crucial. But we cannot be quite sure that all the combinations are correct unless we have a real occurrence which is provided by one or more concordances (which validate a possible combination of words).

IV. Presentation of the Virtual Library Content in the 3D Space

As mentioned before, the current implementation of the 3DVLM exploits the excellent 3D features of the MaxWhere Seminar System including, in the first place, the effectively and decoratively arranged *smartboards* in ready-made 3D virtual spaces where the main content (e.g. *cabinets,* corridors as well as cabinet walls) and the navigation / organization tools (index, thesaurus, reference etc. *pages*) of the virtual library [16–17] can be presented. There are a lot of well-designed and spectacular 3D virtual spaces available on the MaxWhere site [14] which can be used in almost every context, although each space has its unique and distinguished features. For the arrangement and presentation of the virtual library content, in our previous publications [16–17,21] we chose the 3D Castle virtual space. Due to the flexibility of the 3DVLM, we can use other 3D spaces as well. For the new implementation we decided to try and use another 3D space, namely the 3D Library virtual space which contains its smartboards in a two-storey virtual library building. In the following, we would like to present some screenshots, along with detailed explanations, which would illustrate how to access selected virtual library content in the 3D Library space.

Let the starting point be the navigation page [17] (Fig. 9).

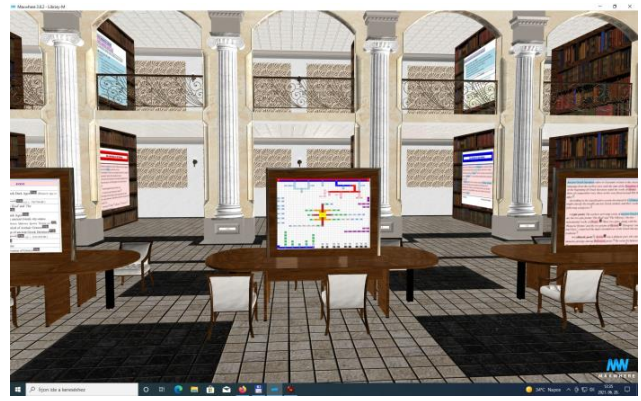

Fig. 9. A screenshot focusing on the *navigation page* of the virtual library located on the ground floor in the MaxWhere 3D Library space.

In the foreground of the image presented in Fig. 9 we can see smartboards which jointly serve as an "information desk" of the 3D library. They offer "smart" access to the navigation / organization tools of the virtual library:

– the *navigation page* is located at the centre of the image; it contains the 2D map of the virtual library where the primary and secondary texts, and other content units are identified by corresponding codes (cf. section II/B. and Fig. 1);

– on the left side we can see a small part of the page which provides a *timeline* of some historical milestones of the ancient era;

–on the right side a part of the *category page* [17] can be seen which contains descriptions of the main classification categories and shows their hierarchical arrangement.

In the background of the image presented in Fig. 9 we can observe some additional smartboards as follows:

– the smartboards on the ground floor of the 3D library present the content of the *main cabinets* of the virtual library containing primary texts about Callimachus (see later);

– the smartboards on the first floor of the 3D library present the *thesaurus pages* of the virtual library including the thesaurus page [T06] introduced in detail in Section III.

Because the main function of the information desk is to help the users find the relevant information, we placed the content of the navigation / organization pages on the wall of the 3D library as well (Fig. 10). Here at the centre of the image we can see again the 2D map of the virtual library which provides entry points to the texts and other content units of the library. Note

that as the collection of the virtual library grows new codes and links will appear in the map.

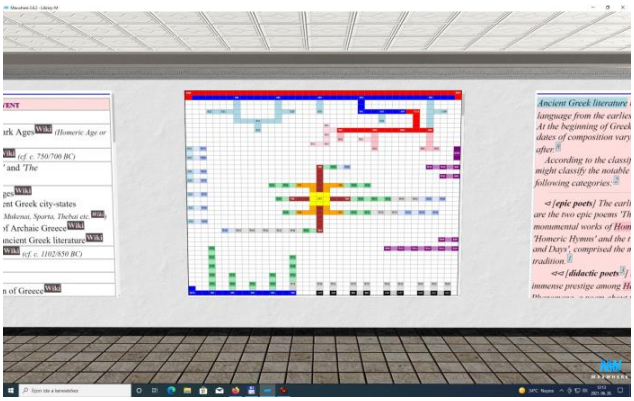

Fig. 10. A screenshot of three *navigation / organization pages* of the virtual library located on the wall in the MaxWhere 3D Library space.

The content of the main cabinets is organized around selected slides about the life and work of Callimachus (labeled as Callimachus / S01, Callimachus in Alexandria / S02, The Great Library of Alexandria / S03, The Pinakes / S04, The works of Callimachus / S05 etc. [16–17,23]). They can be found on the ground floor of the 3D library just behind the information desk. The slide about Callimachus, and that about the Great Library of Alexandria can be seen in Fig. 11.

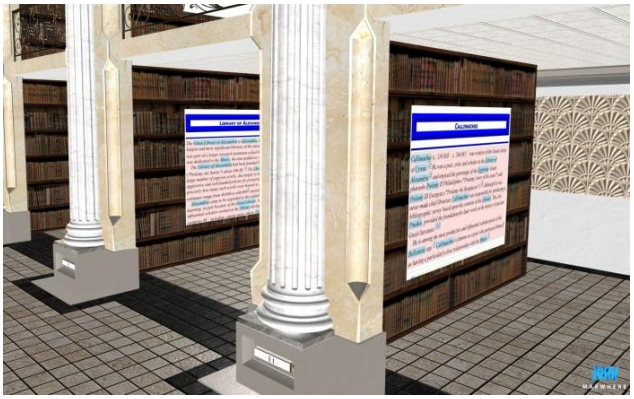

Fig. 11. A screenshot of the *cabinets* which contain slides about Callimachus and the Library of Alexandria located on the ground floor in the MaxWhere 3D Library space.

From a different angle, the slide about the Pinakes can be seen in Fig. 12.

Note that the slides about the life and work of Callimachus can serve as the starting point for the language learners to the full content of the library. After them they can explore other texts some of which are directly or indirectly connected to the primary texts about Callimachus and his work. For example, the presented text about Pinakes, which "was one of the first known documents that lists, identifies, and categorizes a library's holdings" [xvi], involves both a selected text about the collection of the Great Library of Alexandria [S41] and a brief summary about Callimachus' classification system [S42].

Fig. 12. A screenshot of the *cabinet* which contains a slide about the Pinakes located on the ground floor in the MaxWhere 3D Library space.

For those who are interested in the content of the virtual library that we have presented in this study (including the texts, the thesaurus pages, additional supporting materials etc.), the current implementation of the virtual library project can be accessed in 2D web page form via the internet [xviii-xix].

V. Conclusion

Obviously, the linguistic content which the thesaurus page [T06] represents covers only a small piece of the practically infinite variations the natural language can offer. Nevertheless, we think that it can teach and broaden very important skills for the students, e.g.

– the semantically related words (either the adjectives or the keywords), as well as the additional words which occur in the concordances, considerably improve the students' vocabulary;
– the students can get "ready-made" patterns which they can apply directly, and they learn the way how such patterns can be created;
– some concordances (e.g. quotations) and/or their attached context convey useful and valuable meanings which are well worth memorizing;
– hopefully the structure of the material motivates the students to observe and collect similar patterns.

In order to support language learners, we developed a knowledge base the content of which we try to gradually improve with additional items and tools that might be useful for language learners (e.g. preprocessed texts, dictionary and encyclopedia entries, maps, elaborated microcontexts, selected concordances, quotations, selected passages from literary works, generated tests etc.). Because the arrangement of the various items and their relationships play a significant role in the learning process, we took full advantage of both the hypertext-based 2D and the virtual 3D environment *mapping and visualizing* the compiled material of the virtual library using the MaxWhere Seminar System with its excellent features [23].

Because "cognition, metacognition, procedural skills, and motivational factors are important determinants of learning activity" [27], cognitive aspects of the virtual library model are

of utmost importance in language learning. In this respect, we can differentiate several levels, such as

– language level which, among others, makes possible to build words, expressions (phrases, idioms etc.) and sentences using grammatical (e.g. collocation or verb) patterns, stylistic and rhetorical figures or devices etc. (also called syntactic level [28]);

– textological level which, among others, makes possible to build texts (also called semantic level [28]);

– intertextual or hypertextual level which, among others, makes possible to build coherent (or semi-coherent) texts from different texts (also called pragmatic level [28]);

– presentation level which, using interactive 2D browsers and/or various 3D VR or AR (Virtual or Augmented Reality) tools, presents grammatical, textual, intertextual, hypertextual etc. relationships between the lexical items or units of the presented knowledge base and makes possible to have access to them in the virtual 2D or 3D space.

Note that the higher levels are based on the lower ones, and the resulting complex structure is one of the key factors that contribute to the overall effectiveness of the learning process. With regard to the fourth level, cognitive infocommunications research shows that using digital 3D VR and AR technology in education can support cognitive processes such as finding, processing, memorizing, and recalling information. Moreover, 3D environments are also "capable of providing users with a much higher level of comprehension when it comes to sharing and interpreting digital workflows" [29], e.g. when using e-learning tools or participating in a collaborative learning process [30].

At the first level the virtual library offers various solutions which can support the language learning process. As we have seen above, a thesaurus page is organized around separate but interwoven elements of knowledge called microcontext (i.e. two separate groups of semantically related words, illustrated by a list of selected concordances). We added tests to help the learners memorizing vocabulary and lexical items which seem to be necessary for them "to become a long-term part of the learner's own store of English" [24]. The distinction between learning lexical items and their context by reading the pages repeatedly and practicing them by using tests raises the question whether it is worth presenting those parts of the knowledge base simultaneously, in separate smartboards. For example, one smartboard can display the tests, and another (practically the one that is next or opposite to the other one) can display the vocabulary, collocations, concordances etc. which the tests are based on. It can be a real benefit which only the 3D environment can provide.

As feedbacks are always important, future works could focus on evaluating the syntactic and semantic level using AI tools (e.g. the Google Translate service) and/or assessing the motivation and performance of language learners when using the 2D and 3D forms of the virtual library.

Acknowledgment

The results presented in this paper have partially been achieved in the Virtual Reality Laboratory of the Faculty of Informatics of the University of Debrecen, Hungary. This work has been supported by QoS-HPC-IoT Laboratory and project TKP2021-NKTA of the University of Debrecen, Hungary. Project no. TKP2021-NKTA-34 has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the TKP2021-NKTA funding scheme

Sources

[i] Webster's Collegiate Thesaurus. Springfield, MA: G. & C. Merriam Company, 1976.

[ii] Roget's Thesaurus of English Words and Phrases. (rev. by Dutch, Robert A.,) Penguin Books, 1978.

[iii] Roget's Thesaurus of English Words and Phrases. (ed. by Kirkpatrick, Betty) Penguin Books, 1987.

[iv] The Penguin Dictionary of English Synonyms and Antonyms. (ed. by Fergusson, Rosalind,) Penguin Books, 1986.

[v] Oxford Learner's Thesaurus. A Dictionary of Synonyms. Oxford: Oxford University Press, 2012.

[vi] Tom McArthur, Longman Lexicon of Contemporary English. Longman, 1984.

[vii] Oxford Collocations Dictionary for Students of English. Oxford University Press, 2002.

[viii] Longman Dictionary of Contemporary English. Longman, 2001.

[ix] Longman Dictionary of American English. Pearson-Longman, 2009.

[x] Hornby, A. S., Oxford Advanced Learner's Dictionary of Current English. Oxford: Oxford University Press, 2004.

[xi] Macmillan English Dictionary for Advanced Learners. Oxford: Macmillan Education, 2002.

[xii] Cambridge International Dictionary of English. Cambridge University Press, 1995.

[xiii] Longman Language Activator. The World's First Production Dictionary. Longman, 1993.

[xiv] The Oxford Dictionary of Quotations. Oxford: Oxford University Press, 1964.

[xv] Bartlett, John, Familiar quotations. A collection of passages, phrases, and proverbs traced to their sources in ancient and modern literature. Boston-Toronto-London: Little, Brown and Company, 1992.

[xvi] *Wikipedia*.
[Online]. Available: https://en.wikipedia.org/wiki/Callimachus etc. (2019-07-15)

[xvii] *Project Gutenberg*.
[Online]. Available: http://www.gutenberg.org/ (2020-06-04)

[xviii] *3DVLM | thesaurus [T6]*.
[Online]. Available:
https://bodaistvan.hu/callimachus/_thes/accurate_description.html (2023-04-29)

[xix] *Interactive map of the three dimensional virtual library (3DVLM)*.
[Online]. Available: https://bodaistvan.hu/callimachus/map.html (2021-06-29, 3D version; 2023-04-28, 2D version)

References

[1] P. Baranyi and Á. Csapó, "Definition and synergies of Cognitive Infocommunications," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 67–83, 2012.

[2] P. Baranyi, A. Csapó, and Gy. Sallai, Cognitive Infocommunications (CogInfoCom). Springer International, 2015.

[3] A. Gilányi, A. Rácz, A. M. Bólya, J. Décsei, K. Chmielewska, "A Presentation Room in the Virtual Building of the First National Theater of Hungary," in P. Baranyi Ed. CogInfoCom 2020. Proc. of the 11th IEEE International Conference on Cognitive Infocommunications, (Mariehamn, Finland, 23-25 Sept 2020,) pp. 519–523.

[4] A. Rácz, A. Gilányi, A. M. Bólya, J. Décsei, K. Chmielewska, "On a Model of the First National Theater of Hungary in MaxWhere," in P. Baranyi Ed. CogInfoCom 2020. Proc. of the 11th IEEE International Conference on Cognitive Infocommunications, (Mariehamn, Finland, 23- 25 Sept 2020,) pp. 575–576.

[5] I. Horváth, "How to Develop Excellent Educational Content for 3D VR," in P. Baranyi Ed. CogInfoCom 2019. Proc. of the 10th IEEE International Conference on Cognitive Infocommunications, (Naples, Italy, October 23-25, 2019,) pp. 483–489.

[6] I. Horváth, "Personalized Learning Opportunity in 3D VR," in P. Baranyi Ed. CogInfoCom 2020. Proc. of the 11th IEEE International Conference on Cognitive Infocommunications, (Mariehamn, Finland, 23-25 Sept 2020,) pp. 425–439.

[7] B. Berki, "Sense of Presence in MaxWhere Virtual Reality," in P. Baranyi Ed. CogInfoCom 2019. Proc. of the 10th IEEE International Conference on Cognitive Infocommunications, (Naples, Italy, October 23-25, 2019,) pp. 91–94.

[8] B. Berki, "Navigation Power of MaxWhere: a Unique Solution," in P. Baranyi Ed. CogInfoCom 2020. Proc. of the 11th IEEE International Conference on Cognitive Infocommunications, (Mariehamn, Finland, 23-25 Sept 2020,) pp. 511–515.

[9] I. Boda, M. Bényei, and E. Tóth, "New dimensions of an ancient Library: the Library of Alexandria," in CogInfoCom 2013. Proc. of the 4th IEEE International Conference on Cognitive Infocommunications, (Budapest, Hungary, December 2-5, 2013,) pp. 537–542.

[10] I. Boda, E. Tóth, M. Bényei, and I. Csont, "A three-dimensional virtual library model of the ancient Library of Alexandria," in ICAI 2014. Proc. of the 9th International Conference on Applied Informatics, (Eger, Hungary, January 29-February 1, 2014,) vol. 1, pp. 103–111.

[11] I. Boda, E. Tóth, I. Csont, and L. T. Nagy, "Toward aknowledge base of literary content focusing on the ancient Library of Alexandria in the three dimensional space," in CogInfoCom 2015. Proc. of the 6th IEEE International Conference on Cognitive Infocommunications, (Győr, Hungary, October 19-21, 2015,) pp. 251–258.

[12] I. Boda, E. Tóth, I. Csont, and L. T. Nagy, "Developing a knowledge base of ancient literary texts in virtual space," in P. Baranyi Ed. CogInfoCom 2016. Proc. of the 7th IEEE International Conference on Cognitive Infocommunications, (Wrocław, Poland, October 16-18, 2016,) pp. 263–270.

[13] I. Boda, E. Tóth, I. Csont, and L. T. Nagy, "The use of mythological content in virtual learning environment," in P. Baranyi Ed. CogInfoCom 2017. Proc. of the 8th IEEE International Conference on Cognitive Infocommunications, (Debrecen, Hungary, September 11-14, 2017,) pp. 307–314.

[14] *MaxWhere VR Even more*. [Online]. Available: http://www.maxwhere.com/ (2021-01-05).

[15] I. Boda and E. Tóth, "From Callimachus to the Wikipedia: an ancient method for the representation of knowledge in the WWW era," in Gy. Sallai Ed. CogInfoCom 2018. Proc. of the 9th IEEE International Conference on Cognitive Infocommunications, (Budapest, Hungary, August 22-24, 2018,) pp. 205–210.

[16] I. Boda and E. Tóth, "Text-based approach to second language learning in the virtual space focusing on Callimachus' life and works," in P. Baranyi Ed. CogInfoCom 2019. Proc. of the 10th IEEE International Conference on Cognitive Infocommunications, (Naples, Italy, October 23-25, 2019,) pp. 439–444.

[17] I. Boda and E. Tóth, "English language learning in virtual 3D space by visualizing the library content of ancient texts," in P. Baranyi Ed. CogInfoCom 2020. Proc. of the 11th IEEE International Conference on Cognitive Infocommunications, (Mariehamn, Finland, 23-25 Sept 2020,) pp. 305–311.

[18] R. Gairns and S. Redman, Working words: a guide to teaching and learning vocabulary. Cambridge: Cambridge University Press, cop. 1986. pp. 22–43.

[19] *Oxford Advanced Learner's Dictionary*. [Online]. Available: https://www.oxfordlearnersdictionaries.com/ (2021- 06-15)

[20] C. Tribble and G. Jones, Concordances in the Classroom.A resource book for teachers. Harlow: Longman, 1990.

[21] J. Harmer, The Practice of English Language Teaching. Harlow: Longman, 2009.

[22] M. Hewings, Advanced grammar in use: a self-study reference and practice book for advanced learners of English. Cambridge: Cambridge University Press, cop. 1999. Unit 91. pp. 182–183.

[23] I. Boda, and E. Tóth, "English language learning by visualizing the literary content of a knowledge base in the three-dimensional space," *Ann. Math. Inform.* vol. 53, pp. 45–59, 2021.

[24] J. Scrivener, Learning Teaching. A guidebook for English language teachers. Macmillan Education, 2005. Chapter 11. Lexis. pp. 226-251.

[25] I. Boda and E. Tóth, "Content development for second language learning in the 3D virtual space," in CogInfoCom 2021. Proc. of the 12th IEEE International Conference on Cognitive Infocommunications, (September 23-25, 2021, online on MaxWhere 3D web,) IEEE, 2021. pp. 321–327.

[26] I. Boda, E. Tóth and L. T. Nagy, "Improving A Bilingual Learning Material in the Three-Dimensional Space Using Google Translate," in CogInfoCom 2022. Proc. of the 13th IEEE International Conference on Cognitive Infocommunications, (Online, September 22-23, 2022,) IEEE, 2022. pp. 93–98.

[27] F. E. Weinert, "Introduction and overview: metacognition and motivation as determinants of effective learning and understanding," in F. E. Weinert and R. H. Kluwe, *Metacognition, Motivation and Understanding*. Hillsdale, NJ – London: Lawrence Erlbaum Associates, 1987. pp. 1–16.

[28] G. Tolcsvai Nagy, A magyar nyelv szövegtana [Textology of the Hungarian Language]. Budapest: Nemzeti Tankönyvk.

[29] I. Horváth, "An analysis of personalized learning opportunities in 3D VR," *Front. Comput. Sci.*, 3:673826, 2021.

[30] B. Lampert, A. Pongracz, J. Sipos, A. Vehrerand, I. Horváth, "MaxWhere VR-learning improves effectiveness over classical tools of e-learning," *Acta Polytechnica Hungarica*, vol. 15, no. 3, pp. 125–147, 2018.

István Károly Boda is a college professor at the Debrecen Reformed Theological University, Hungary. He is the head of the Department of Mathematics and Informatics. He was conferred the degree of Ph.D. in 2002; the title of his doctoral dissertation was "Use of Hypertext in Information Science: Concepts, Systems, Models and Applications". In 1994, he was declared a habilitated doctor ("dr. habil.") in the field of mathematics and computer science. His research interest includes computational linguistics and the applications of hypertext especially in text linguistics and stylistics. Presently he investigates the cognitive infocommunication aspects of developing knowledge bases in virtual 2D and 3D environments.

Erzsébet Tóth is an assistant professor at the Faculty of Informatics, University of Debrecen, Hungary. She had a degree in English language and literature and library and information science in 1995. In 2008 she obtained her Ph.D degree in information science and technology. In the doctoral dissertation she investigated the evaluation of search engine performance. She is involved in a virtual library project that focuses on the presentation of digitized library content in three-dimensional space, and she studies the enhanced possibilities of the English language teaching and learning in virtual learning environment. Since 2022 she has been an IEEE member.