Мая Йосифова
(България, Софийски университет „Св. Климент Охридски“)

Концептуален модел на уеб-базиран речник на българския жестов език

Резюме: Целта на статията е да представи основните етапи в разработването на концептуален модел на онлайн речник на българския жестов език (БЖЕ). Съществуват малко на брой дигитални ресурси, даващи възможност за езикови справки и обучения по БЖЕ. Това мотивира разработването на двуезичен речник на български вербален и жестов език. Към момента разработваните функционалности в проекта включват:

– Търсене на съдържание по ключови думи;

– Търсене на съдържание по семантични категории;

– Визуализация на съдържанието – видеоматериал на жестови единици със съответните им дефиниции на писмен български език; жестови варианти; артикулационни специфики; семантични категории.

Статията описва накратко и техническия процес по създаването на ресурса, както и лингвистичните данни, използвани в неговата демоверсия. Идеята на проекта е да предостави възможност на членовете на Глухата общност да изпращат видеофайлове с жестови варианти, които не се откриват в речника. Целта на концептуалния модел на речника е да подпомогне процеса по създаването на технически ресурси, необходими за популяризирането и изучаването на БЖЕ. Предвижда се съдържанието на платформата да бъде разработвано както от специалисти по БЖЕ, така и от носителите на езика от цялата страна. В разработването на представения проект участва интердисциплинарен екип.
Ключови думи: концептуален модел, онлайн речник, Български жестов език, жестов език.

Maya Yosifova
(Bulgaria, Sofia University *St. Kliment Ohridski*)

A Conceptual Model for a Web-based Sign Language Dictionary

Abstract: The article aims to outline the key stages in developing a conceptual model for a web-based dictionary of the Bulgarian Sign Language (BGSL). It describes the technical process involved in creating the resource and the linguistic data used in its demo version. The dictionary will be developed collaboratively by BGSL experts and language users from across the country. Its primary function is to provide essential linguistic information about BGSL and support its study. The project has been carried out by an interdisciplinary team.
Keywords: conceptual model, web-based dictionaries, Bulgarian Sign Language, sign languages

1. Introduction

This paper introduces a conceptual model for a digital web-based dictionary in sign language. In 2021, the Bulgarian Sign Language was officially recognized by government law. Currently, there are only a handful of digital resources available to aid in language comprehension, training, and research. This scarcity motivated the development of a bilingual web dictionary that encompasses both Bulgarian verbal and sign language. The key functionalities of this project include:

- Keyword-Based Content Search;
- Semantic Category-Based Content Search;
- Visualization of Content: Video materials containing sign language word units alongside their corresponding written Bulgarian definitions, different sign variants, sign articulation specifics, and sign category definitions.

Additionally, the project allows members of the Deaf community in Bulgaria to contribute video content with sign variants not yet included in the dictionary. Ongoing development includes statistical processing, such as tracking the most frequently searched (used) signs from the Bulgarian Sign Language (BGSL). The overarching goal of this conceptual model for a digital dictionary is to facilitate the creation of technical resources necessary for the popularization and learning of BGSL.

2. Rationale for the Idea

Bulgarian Sign Language stands as a natural language, expressing its unique linguistic features spatially—a stark contrast to verbal languages (Tisheva et al. 2017). As the primary mode of communication within the Deaf community in Bulgaria, BGSL has become a subject of extensive research across various fields. Beyond its scientific applications, BGSL is increasingly finding its place in translation services and artistic expressions. The growing interest in sign language raises important questions about its study, particularly for hearing individuals and those who acquire hearing impairments later in life.

Additionally, the development of sign and verbal competence plays a crucial role in the education of active language users. The adoption of the Bulgarian Sign Language Act in 2021 signifies its formal recognition and paves the way for its integration into the educational system. Consequently, interdisciplinary teams are now conducting in- depth scientific research to further understand and promote the language.

BGSL thrives in its various linguistic variants, disseminated among social groups inhabiting different regions of the country (Lozanova, Stoyanova 2022). However, collecting and processing empirical data on BGSL necessitate meticulous preparation during field studies. Active participation from BGSL users within the research area, thoughtful selection of research materials, and adherence to established principles and standards in video material creation are essential components of this endeavor. The study of natural Bulgarian Sign Language within scientific circles has only recently gained momentum (Lozanova, Stoyanova 2022). Researchers face challenges such as establishing academic traditions for the active participation of Deaf community members, insufficient scientific research on the language's structure, and the need to develop accessible digital resources. Hence, the creation of an online dictionary becomes essential. This resource aims to provide accessible information about BGSL, benefiting both researchers and language users. The dictionary's main objectives include:

- Active Participation: Engaging the Deaf community in language studies;
- Linguistic Information: Providing basic details about BGSL;
- Facilitating Study: Supporting research and learning related to BGSL.

By addressing these goals, the online dictionary aims to enhance understanding, accessibility, and appreciation of BGSL.

3. Technical Specifications for the Presented Dictionary

In the process of creating the online dictionary, a collaborative team comprising software engineers from France and a linguist with research interests in BGSL is actively involved. The stages of building the online platform involve researching existing similar resources and selecting suitable technologies that align with the project's requirements and goals. Properly structuring the technical process and documenting the necessary details are crucial for determining the appropriate technologies to create the dictionary. Let's delve into the technical solutions adopted by the team, along with their advantages. The dictionary's architecture consists of two main components: the frontend and the backend, both developed using familiar programming languages known to the team.

3.1. Frontend:

The frontend encompasses the elements visible to users on the screen, with which they interact. Among the popular frontend frameworks, Angular stands out due to its ease of use, scalability, and robust features. Maintained by Google, Angular boasts comprehensive documentation and a large community of software developers. Leveraging this framework allows building user applications with high performance and an improved user experience.

3.2. Backend:

The backend, often referred to as the server side of the website, operates behind the scenes and remains invisible to users. It manages the application's data (content). For robust database management, the team has opted for Microsoft SQL Server, known for its security features. However, the Angular front-end cannot directly interact with the database information. To bridge this gap, an Application Programming Interface (API) will be developed. This API, built on the .NET framework, ensures secure, reliable, and high-performance communication between the frontend and the backend services, providing essential data to the application.

By combining these technologies, the team aims to create an efficient and user- friendly digital dictionary for Bulgarian Sign Language. The framework also integrates robust security features, safeguarding applications against suspicious users. Additionally, it offers external authentication options, including Google sign-in. With these measures in place, the next critical step is deployment—the process of making the software accessible to end users and other programs[1].

Currently, we're engaged in discussions regarding the deployment strategy for the presented web-based dictionary applications. Two excellent options under consideration are Microsoft Azure and Heroku. Deploying on either of these cloud platforms ensures internet accessibility for the website.

Our ongoing research involves exploring existing dictionary resources to select state-of-the-art concepts. As we fine-tune the platform, various technical changes are possible to optimize performance.

4. Content Structure of the Sign Language Dictionary

4.1. Related Work

On a global scale, linguistic research on sign languages began in the early 1960s, leading to the creation of numerous online dictionaries and corpora. The ongoing digitalization processes over the past two decades have fueled this growth. Web platforms and applications have emerged, serving educational, popularization, and research purposes for various sign languages. For our current project, we thoroughly examined available resources, both within and outside a unified platform[2]. Notably, American Sign Language (ASL)[3], German Sign

[1] https://www.ibm.com/docs/en/zos/2.4.0?topic=task-deploying-software
[2] http://deafeducationworldwide.weebly.com/online-sign-language-dictionaries.html
[3] https://www.handspeak.com/

Language (DGS)[4], and French Sign Language (FSL)[5] have well-established digital dictionaries. However, for Bulgarian Sign Language, several paper-based dictionaries exist. Among them, the most comprehensive and recently published is *Rechnik na balgarskiya zhestov ezik* (Tisheva et al. 2017), which is also available in electronic format.

In addition to these resources, an online sign language dictionary has been specifically developed to aid communication with deaf individuals during crisis situations (Lozanova, Stoyanova 2022). Lexical units from BGSL are also featured in an international website[6] that aggregates data from various sign languages worldwide. Furthermore, several online resources cater to the educational needs of people with hearing impairments. Specialized video materials on Bulgarian Sign Language are freely accessible. These materials form part of the manual titled *Development of Skills for Linguistic Analysis of Sign Language*[7], which has been adapted for deaf researchers. Its purpose is to acquaint scholars with the methodology of organizing language data within linguistic corpora.

The main goal of the developed conceptual model is to create a bilingual online dictionary that provokes active participation of the Deaf community in language studies, provides basic linguistic information, and facilitates the study of BGSL. In this sense, the administration of the website is planned to be carried out by deaf and hearing experts. Together, they can create a rich and inclusive resource that benefits learners, researchers, and the broader community. As we continue our work, more specialized resources with diverse purposes for presenting BGSL are actively under development.

4.2. Data collection

The conceptual model for the dictionary is an integral part of a broader study on natural Bulgarian Sign Language. Our language data collection and processing occur through field studies conducted in the cities of Montana and Burgas. Participants in these studies are members of the Deaf community who primarily use BGSL for communication.

To ensure minimal interference from Bulgarian verbal language, we meticulously examined materials used globally for similar research. These selected resources are tailored for presentation to BGSL users. Our primary stimulus during BGSL studies is the animated film *Canary Row*, complemented by various images suitable for extracting grammatical information.

The video resource we employ has also been utilized in studies on other sign languages, including Russian Sign Language (RSL), German Sign Language (DGS), Dutch Sign Language (NGT), Bali's Kata Kolok (KK), and American Sign Language (ASL) (Kimmelman et al. 2019; Casey, Emmorey 2009). The language data derived from this stimulus holds promise for future comparative analyses across different sign languages.

During video recording processing, we discovered regional sign variants corresponding to verbal components within the semantic category *Animals*. While our primary focus remains on studying grammatical relationships in BGSL, the collected video material also provides insights into regional lexical units. However, for the scope of this article, we emphasize the digital resources' structure.

These sign variants have been thoughtfully incorporated into the developed dictionary, showcasing its fundamental functions and capabilities.

4.3. Functionalities of the Developed Sign Language Dictionary

[4] https://web.dgs-korpus.de/dictionary-dw-dgs.html; https://signdict.org/
[5] https://www.sematos.eu/index.html
[6] https://www.spreadthesign.com/en.us/search/
[7] https://www.univercity-project.eu/

In this section, we outline the functionalities implemented in the dictionary up to this point. The main menu introduces options such as Home, Categories, Corpus, and Contacts. This standardized structure enables users to access various BGSL-related information.

The search for a specific lexeme can be accomplished in two ways:

1. Standard Search Engine: Users can type the desired word into a search bar;
2. Semantic Categories: Alternatively, users can select a specific semantic category (e.g., Family, Housing, Animals, Plants) to refine their search.

Currently, the Corpus option is in development. Its primary purpose is to systematize video resources for extracting data across different language levels within the BGSL system. Sign language users will have the opportunity to submit videos of spontaneous speech, contributing to research on natural sign language.
Guidelines for using the platform, its goals, and information about administration (Home, Contacts) will also be included. The overarching aim of the developed conceptual model for the online dictionary is to provide accessible BGSL information in the most suitable manner. This includes video materials featuring sign units and their descriptions.

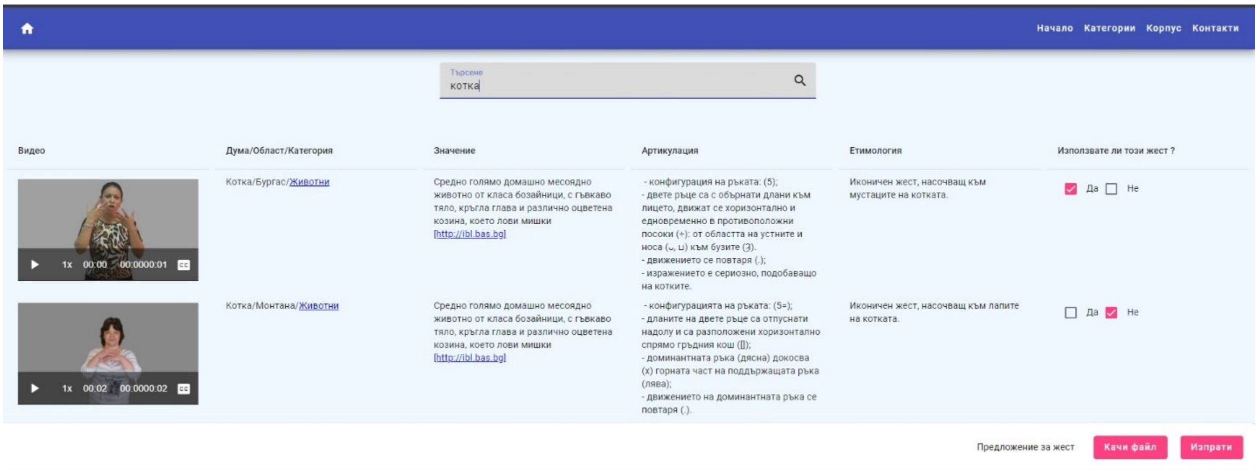

Figure 1: A screenshot of the Search function in the online dictionary.

Figure 1 illustrates one of the possible search methods within the dictionary. The video material showcases the sign for CAT, as used in the cities of Montana and Burgas. Each unit's description contains:

1. Word/Area/Category:
   - Indicates the area where the sign is observed and specifies the semantic category it belongs to.
2. Meaning:
   - Provides an explanation of the concept, sourced from the Bulgarian Academy of Sciences.
3. Articulation:
   - Offers basic linguistic information related to the motor execution of the sign in space.
4. Etymology:
   - Explores the origin of the sign and potential structural correspondences with extralinguistic referents.
5. Do You Use This Sign?:
   - Allows users to mark the sign as known or unknown. This data contributes to statistical analysis of commonly used signs.

6. Suggested Signs: Upload File/Send:

● Encourages active involvement of the Deaf community by enabling users to submit videos. These authentic materials enhance research on natural sign language and aid in determining sign meanings

4.4. Characteristics of Sign Variants

The task of the conceptual model is to provide lexical and semantic information about sign units and introduce fundamental concepts related to the grammatical structure of Bulgarian Sign Language. The primary components of sign languages are broadly described in terms of word formation (morphology), word order, and grammatical relationships within sign sentences (syntax), as well as the compositional organization of signs concerning their articulatory features (phonology). In studies of American Sign Language, William Stokoe identified three fundamental phonological parameters that can be combined in various ways to form lexical units: handshape [dez], movement [sig], and location of execution [tab] (Stokoe 1960). For BGSL, as well as other sign languages, five basic categories are used to represent signs:
Handshape/Configuration [dez];

- Articulation Location of the Sign [tab];
- Linguistic Movement [sig];
- Palm Orientation [ori];
- Finger Orientation [ori] (Tisheva et al. 2017).

The structure and meaning of each sign depend on how individual articulators shape, move, and localize. Therefore, it is essential to uncover fundamental information about the way a given sign unit is structured. Following examples from some of the researched online resources and for demonstration purposes in the developing dictionary, basic articulatory characteristics of the included signs were manually introduced. Figures 2 (CAT_B) and 3 (CAT_M) present synonymous variants of the word CAT in BGSL, observed in the cities of Montana and Burgas. The primary descriptions during the production of both signs are outlined in the Articulation section. They include handshape, linguistic movement, and hand position concerning the articulation location. The textual description of the structure is supplemented with symbols from the notation system adopted and further developed for BGSL (Tisheva et al. 2017). Marking of facial expressions, defined as non- manual articulators, is also anticipated. These play a crucial role in conveying meaning and various grammatical relationships.

Figure 2: CAT_B

Figure 3: CAT_M

The execution of the sign CAT_B involves the basic handshape 5, with palms facing the articulation location (from the area near the lips and nose toward the cheeks - ᴗ, ʒ ;). The movement is horizontal, repetitive (.), and in opposite directions relative to the face (÷). This sign represents the whiskers of the animal. For CAT_M, the observed handshape is 5= (a variant of 5), with downward-facing palms, the right hand as the active hand, and the left is as the supporting hand. The articulation location is horizontal relative to the chest ([]); the active hand performs the motor act through repeated touches (x) on the supporting hand. The sign signifies the animal's paws.

Providing similar descriptions of the structure of meaningful sign units in the dictionary allows for comparing execution methods and segmenting the participating articulators. Including notation symbols aids users in familiarizing themselves with fundamental linguistic data applied in BGSL research and facilitates orientation within the structural model of signs. Currently, the team is discussing possibilities for developing functionalities that allow marking syntactic and morphological features in the Corpus section.

5. What's next?

The presented conceptual model of the web-based online dictionary for Bulgarian Sign Language is an attempt to create a resource that supports language development and popularization. The proposed version includes fundamental linguistic information, aiming to introduce deaf individuals to the language research process and assist researchers. The project is currently under development, with the following functionalities planned:

1. Automated Search for Sign Units by Handshape:
- The system will allow users to search for sign units based on the specific handshape used;

2. Automated Description Generation:
- Descriptions of sign units will be automatically generated, providing essential linguistic information;

3. Notation System Symbols Explanation:
- The dictionary will include explanations of the symbols used in the notation system, aiding users in understanding the linguistic details

4. Integration with External Resources:
- The online dictionary will be linked to external resources, enhancing its functionality and providing additional context.

REFERENCES

Emmorey, Shannon 2009: *Emmorey. K., C. Shannon*. Co-speech gesture in bimodal bilinguals. In: Language and Cognitive Processes, 24 (2): 290–312.

Kimmelman et al. 2019: *Kimmelman. V. et al*. Argument Structure of Classifier Predicates: Canonical and Non-canonical Mappings in Four Sign Language. In: Open Linguistics, 5: 332–353.

Lozanova 2023: *Lozanova, S*. Semiotic aspects of bimodal bilingualism. St. Kliment Ohridski University Press, Sofia.

Lozanova, Stoyanova 2022: *Lozanova. S., I. Stoyanova*. WordNet-Based Bulgarian Sign Language Dictionary of Crisis Management Terminology. In: Proceedings of the Fifth International Conference Computational Linguistics in Bulgaria (CLIB 2022), pages: 59-67, Publisher: Bulgarian Academy of Sciences, Sofia.

Stokoe 1960: *Stokoe, W. C*. Sign Language Structure: An Outline of the Visual Communication System of the American Deaf, (1993 Reprint ed.) Silver Spring, MD: Linstok Press.

Tisheva et al. 2017: *Tisheva. Y. et al*. Teoretichno opisanie na gramatikata na balgarskiya zhestov ezik. Sofia: MON. Tisheva et al. 2017: *Tisheva. Y. et al*. Rechnik na balgarskiya zhestov ezik. Sofia: Nauka i obrazovanie.