

# Rule-based Machine Translation into Ukrainian Sign Language Using Concept Dictionary

Olga Lozynska <sup>1</sup>[0000-0002-5079-0544], Maksym Davydov <sup>2</sup>[0000-0001-7479-8690], Volodymyr Pasichnyk <sup>3</sup>[0000-0002-5231-6395] and Nataliia Veretennikova <sup>4</sup>[0000-0001-9564-4084]

<sup>1, 2, 3, 4</sup> Information Systems and Networks Department, Lviv Polytechnic National University, Lviv, Ukraine

Olha.V.Lozynska@lpnu.ua, Maksym.V.Davydov@lpnu.ua,  
vpasichnyk@gmail.com, nataver19@gmail.com

**Abstract.** This paper describes rule-based machine translation into Ukrainian Sign Language using concept dictionary. Ukrainian Sign Language has its own grammar rules that are different from the Ukrainian spoken language. That is why, it is necessary to develop a system for recording all the elements of sign language, to create appropriate translation dictionaries and grammar rules for parsing and translating sign language correctly. The translation from spoken language to sign language is not an easy task. Sometimes a single sign means a whole phrase, but more often several signs are used to explain a single word. To solve this problem, it is used an approach based on concepts and relationships between them. We identified five main cases of relationships between words, signs and concepts used for translating Ukrainian Sign Language. It is proposed an algorithm for translation from Ukrainian Spoken Language to Ukrainian Sign Language based on concepts. The algorithm was tested using database of 360 sentences, which contained 60 concepts. As a result, 87% of sentences were translated correctly, 32% of which contained concepts, 13% were not translated due to the lack of word to sign correspondence.

**Keywords:** Ukrainian Sign Language, Concept dictionary, Infological model, Translation algorithm.

## 1 Introduction

Nowadays, the most important problem in the world is the creation of various information and communication technologies for people with disabilities [1]. Automatic translation helps people communicate and overcome linguistic and cultural barriers. Today, one of the most famous translation system is Google Translate, which combines neural nets, rules and statistical methods to translate into a lot of languages. However, the problem of translation into sign language (SL) has not been resolved yet and communication with deaf people remains uncovered by machine translation. Sign language is a visual-spatial language that has its own structure and is ubiquitously used by people with hearing impairments. Sign language uses hand gestures, lips articulation and facial expressions for communication. Ukrainian Sign Language (USL), like other well-

known sign languages, has its own grammar and rules, different from the Ukrainian Spoken Language (USpL).

There is no universal sign language in the world. Sign Languages from different countries have evolved independently, and therefore differ from each other. In addition, each country has a local sign language or a variety of languages that reflect the culture of people with hearing impairments. For example, in Switzerland there are four local spoken languages used by deaf French, Italian, German, and Swiss communities [2]. Austrian Sign Language (ÖGS) and German Sign Language (DGS) are two different languages despite the use of a single spoken language in Germany and Austria. A similar situation is in the US and UK. The official spoken language of these countries is English. However, people with hearing impairments in America communicate in American Sign Language – ASL, and in the UK – in British Sign Language (BSL) [3].

The complexity of developing an automatic translation system for Ukrainian Sign Language is compounded by the absence of large dictionaries and corpuses of the USL [4]. Sign Languages use various gestures and facial expressions instead of sounds for information presentation. To adapt the translation systems for written languages to the translation for sign language, it is necessary to develop a system for recording all elements of a sign language, to create appropriate translation dictionaries and grammar rules for parsing sign language. In addition, a significant linguistic difference between sign language and spoken language complicates the translation process.

Well-known machine translation systems of sign languages, based on statistical models, use direct relations between signs and language words. Such an approach can provide a high-quality translation only with the use of large training corpus for sign language and spoken language. The absence of such corpus for the Ukrainian Sign Language requires the use of an alternative approach – introduction of concepts and the study of the relationships between these concepts.

Concept is a meaning of sign or spoken expressions that denote the same notion (process, action, sign), and can be reduced only if a word or sign is restored from the context. For example, the phrase “clock goes” is considered as a concept, as in sign language the one sign “CLOCK\_GOES” is used to display this phrase, and the phrase “man goes”, we do not consider as a concept (it is used the signs “MAN” and ”GOES”).

A person who speaks fluently in both languages understands the meaning of sentences and concepts in the sentence quickly and rephrases it using language translation tools. Such a person is an expert who possesses the knowledge necessary for translation. A computer program, that can correctly translate concepts, should have the means to represent the expert knowledge in the form of an appropriate database, containing information about concepts and their translation.

This article describes rule-based machine translation into Ukrainian Sign Language using concept dictionary and the infological model of the concepts that used for Ukrainian sign language translation. In the infological modeling the data is presented in the form of entities (objects), which are interconnected by certain relationships that express dependencies between them. Characteristics or properties of entities are attributes.

The infological model is developed on the basis of analysis of linguistic relations between concepts in Ukrainian Spoken and Ukrainian Sign Languages. The correctness of the infological model was investigated using the parallel sentence corpus "USL –

USpL". The assessment of the translation system quality using the infological model of concepts was carried out.

## 2 Related Work

Sign Languages began to be investigated since 1960 not only in the US [5], but also in Europe [6], Africa [7] and Asia [8]. Ukrainian Sign Language was studied by S.V. Kulbida, I. I. Chepchina, N. B. Adamyuk, N. V. Ivanusheva [9].

Sign language translation systems [3, 4, 10] for experiments use corpus of parallel texts from a particular subject field. Scientists [11] proposed translation system based on grammar rules and parallel English-ASL corpus. Translation mistakes arose in those sentences that were not available in parallel corpus.

The translation of concepts is available in the works [6, 12]. In the work [13] scientists proposed a semi-automatic method for associating a Japanese lexicon with a semantic concept taxonomy, using a Japanese-English bilingual dictionary for machine translation. They described three algorithms to associate a Japanese lexicon with the concepts of the ontology automatically and tested these algorithms for 980 nouns, 860 verbs and 520 adjectives as preliminary experiments. The algorithms are found to be effective for more than 80% of the words.

The relationship between the concepts of Spanish Spoken and Sign Languages is analyzed for Spanish Sign Language translation in studies of Spanish scholars as R. San-Segundo [6]. The expert described the basic rules of translation, based on these relationships. A machine translation module is developed based on these rules. The translation process is carried out in two stages. In the first step, each word is matched with one or more syntax tags. After that, taking into account the translation rules, the marked words are transformed into signs according to the relationship between the concepts. The rule-based translation module contains 153 rules. Sign Error Rate (SER) is used to evaluate the translation result, which is 31.60%.

S. Baldassari with other researchers [12] have developed a system for translating Spanish into Spanish Sign Language. The rule-based machine translation module uses the syntactic and morphological characteristics of the words and their semantic value to generate the corresponding signs. The system was tested using 92 sentences containing 561 words. The translation result is 96% of correctly translated words.

Ukrainian scientists [4] described the algorithmic implementation of information technology for translation from inflectional languages to sign language. Infological model of Ukrainian dictionary and sign language, related to generalized grammatical constructions for automatic translation is built. Scientists have developed information technology based on generalized grammatical constructions of simple sentences of Ukrainian and USL. The experiment results showed that 64% of sentences were translated automatically. However, the researches [4] do not take into account the fact that each statement has a certain semantic sense. For example, in the grammatical constructions "I go" and "time goes", the sign "GO" is shown in different ways. Therefore, we cannot translate this sign unambiguously without taking into account the context. To

solve this problem, it is necessary to create a concept dictionary and their correspondences in USL and USpL.

### 3 Main Part

#### 3.1 Relationships Between Concepts in the Dictionary “Ukrainian Spoken Language – Ukrainian Sign Language”

The main task of Ukrainian Sign Language translation, as for all other cases of translation from one language to another, is a correct content transfer of the translated text. It is a difficult task, because you need to understand fully the source text. During the construction of machine translation systems for sign languages, it is important to determine the relationship between the concepts of spoken and sign languages that express the message content. It should be noted that one sign can denominate the phrase, and some words can be explained using several signs. In addition, most of the statements in the sign language have several meanings that need to be clarified in spoken language.

Considering the peculiarities of relations between words, signs and concepts (see Fig. 1), there are five main cases used for translating Ukrainian Sign Language.

1. Concept is presented by one word of USpL and one sign of USL (Table 1).

In this case, the word is translated directly into one sign. The translation is simple, as one sign of USL has one meaning.

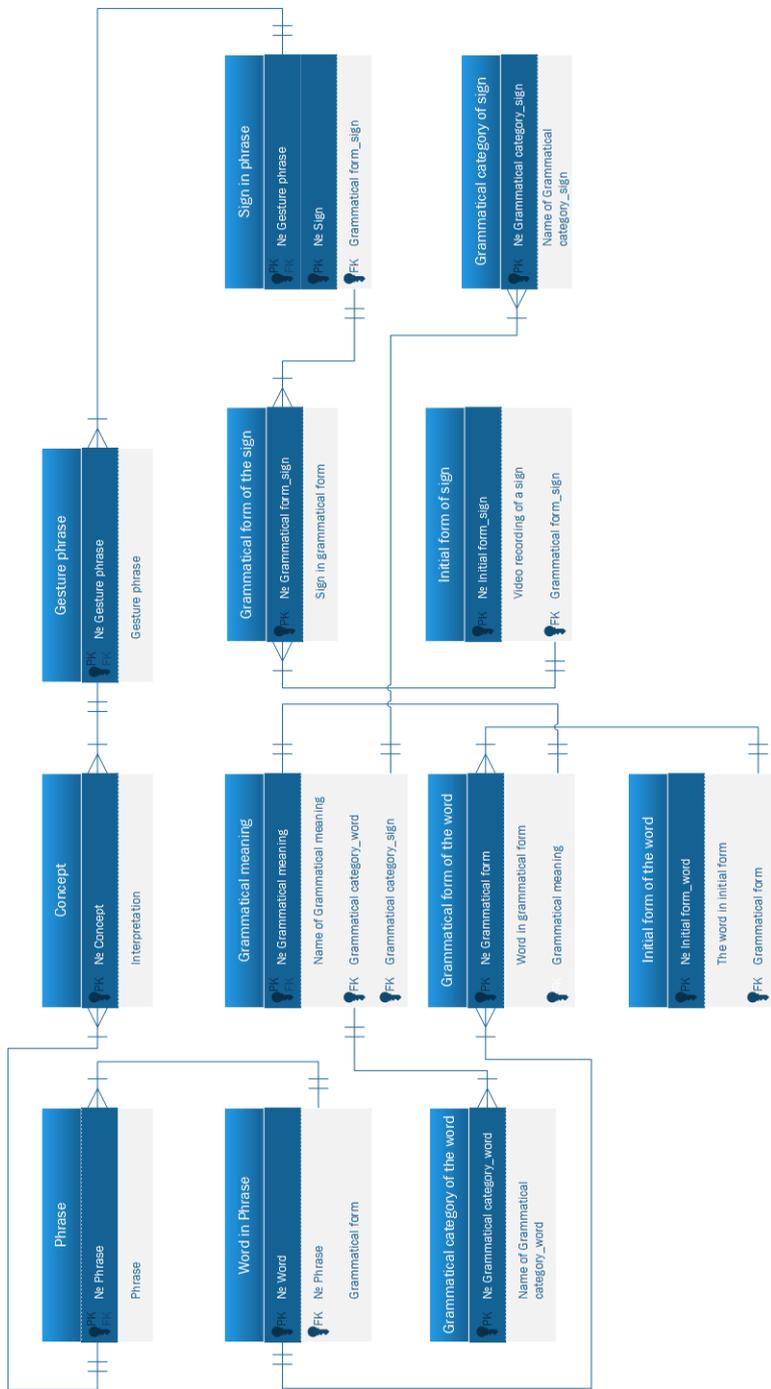
**Table 1.** One word express one sign.

№	Ukrainian Sign Language	Ukrainian Spoken Language
1.	Я ЛЮБИТИ СВОЯ РОБОТА (I LOVE MY JOB)	Я люблю свою роботу (I love my job)
2.	МОЯ МАМА ГОТУВАТИ ВЕЧЕРЯ (MY MOTHER COOK THE DINNER)	Моя мама готує вечерю (My mother cook the dinner)

2. The concept, which is given in a few words, corresponds to one sign (Table 2). For example, the USpL phrase “The heart beats” is translated using one sign “PALPITATION”.

**Table 2.** A few words correspond to one sign.

№	Ukrainian Sign Language	Ukrainian Spoken Language
1.	Я НЕСТИ_СУМКУ (I CARRY_THE_BAG)	Я несу сумку (I carry the bag)
2.	СЕРЦЕБИТТЯ (PALPITATION)	Серце б’ється (The heart beats)



**Fig. 1.** Infological model of concept dictionary.

3. The concept, which is presented in one word, is presented by several signs (Table 3). It is in the following cases:

1) translation of verbs, when a concept that corresponds to one verb, generates several signs. In order to express the verb of the present time in Ukrainian Sign Language, the infinitive is used; for the expression of the verbs of the past or future time, the infinitive and auxiliary verbs “був” (“was”) and “буду” (“will”) are used respectively (Table 3). For example, the USpL concept “писав” (“wrote”) corresponds to sign “писати + був” (“write + was”), and the concept “писатиму” (“will write”) – sign “писати + буду” (“write + will”);

2) translation of general and specific nouns. In sign language there is not always a corresponding sign for a certain concept that can be expressed in one word of Ukrainian Spoken Language. For example, in Ukrainian sign language, there are corresponding sign for the words “gold”, “silver”, “copper”, etc., but there is no sign for the word “metal”. Therefore, in order to solve this problem, the sign language uses several signs to express the general concept (Table 3);

3) translation of a noun in plural. There are several ways to translate a sign in plural: repeating the sign, using additional signs such as: “many”, “different”, “few” or gesturing with both hands. For example, the word “apples” can be shown in the following ways: a two-handed sign “APPLE”, a repetition of the sign “APPLE APPLE” or “APPLE + MANY” (see Fig. 2.).

**Table 3.** One word is presented by several signs.

№	Ukrainian Sign Language	Ukrainian Spoken Language
1.	Я ЗНАЙТИ БУВ ЦЕЙ КНИГА (I FIND WAS THIS BOOK)	Я знайшла цю книгу (I found this book)
2.	ВІН НЕ ЛЮБИТИ ІСТИ ЯБЛУКО БАГАТО (HE DO NOT LOVE TO EAT APPLE MANY)	Він не любить яблука (He does not like apples)
3.	Я МАТИ СОБАКА ТРИ ШТУКИ (I HAVE A DOG THREE THING)	Я маю три собаки (I have three dogs)
4.	ЗОЛОТО, СРІБЛО, МІДЬ (GOLD, SILVER, COPPE,)	Метал (золото, срібло, мідь)



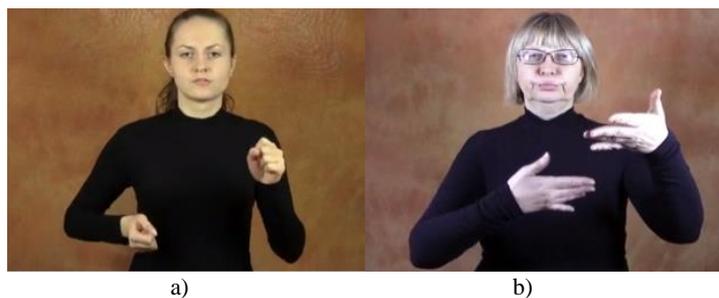
**Fig. 2.** Reproduction of the signs: a – “MANY”; b – “APPLE”.

4) translation of a noun by gender. To express the gender of an object or a subject in USL to the main sign added supplementary signs “man” or “woman”. In most cases, there is no need for such gender clarifications, because it will complicate translation. However, in some cases gender clarification is necessary for the correct transfer of content.

4. Several concepts (several words) correspond to several signs.

This case is the most complicated in terms of translation, since it is necessary to generate signs of USL which would correspond to the content of several words of the spoken language. For example, the verb “go”, depending on the subject of action (“Go to school” and “Time goes fast”), shows various signs: “SCHOOL\_GO” AND “TIME\_FLY” respectively. In this case, the sign verb depends on the subject of action.

5. Concept expresses one sign of USL and few words of Ukrainian. For example, one and the same sign “SKI” (Fig. 3a) can mean: ski, skiing, walking (skiing) on the ski, skier, etc. (of course, in different contexts). Analogical, sign “LIST” (Fig. 3b) is used to define the following concepts: menu, timetable, schedule, program, repertoire, etc. [14]. Therefore, in order to understand the content of the sentence, we need to make a parsing analysis and identify all the members of the sentence (noun, pronoun, verb, adjective, etc.).



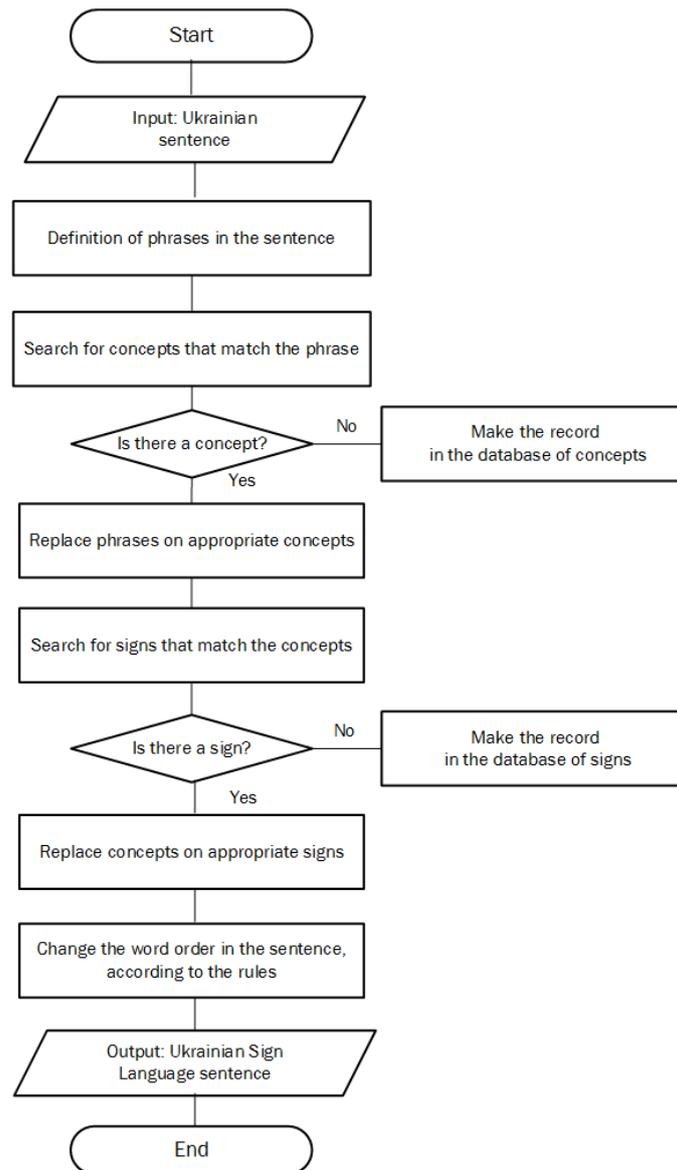
**Fig. 3.** Reproduction of the signs: a – “SKI”; b – “LIST”.

### 3.2 Experiments

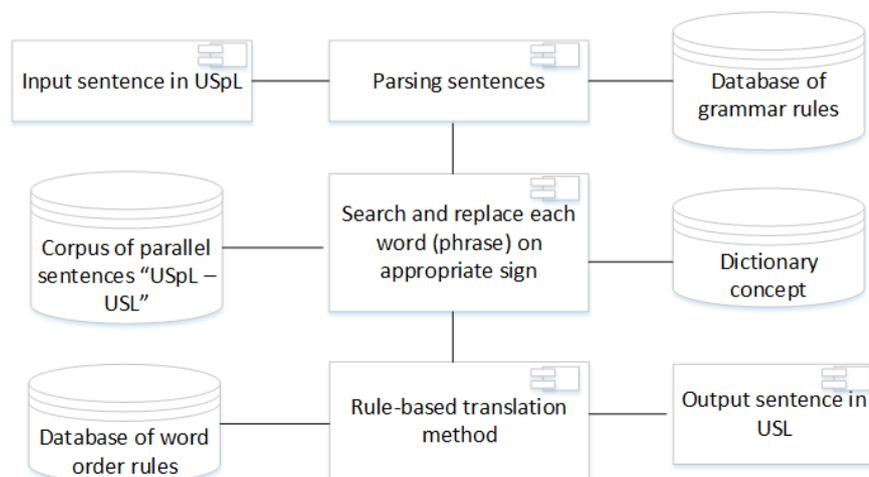
To test the machine translation system we used parser [15], the grammar rules of translation [16], corpus of parallel sentences “USpL – USL” and dictionary of concepts. The parser builds the parsing tree of a sentence. After parsing sentences, searching and replacing each word (phrase) in the Ukrainian Spoken Language with an appropriate sign using the infological model of concepts are performed. Then, the translation rules and the strict word order in the sentence, described in [16], are used for translating into the Ukrainian Sign Language. The algorithm of translation from Ukrainian Spoken Language to Ukrainian Sign Language is depicted in Fig. 4. The architecture of the rule-based machine translation system is shown in Fig. 5.

Let's consider the translation with the use of concepts for the sentence “Моя мама несе сумку з магазину” (“My mother carries the bag from the store”) (Table 4). The words “моя” (“my”) and “мама” (“mother”) correspond to the signs “МОЯ” (“MY”)

and “МАМА” (“MOTHER”) respectively, and the phrase from two words “несе сумку” (“carries the bag”) is a concept, which is translated in one sign “НЕСТИ\_СУМКА” (“CARRY\_BAG”). Because, the prepositions in sign languages do not designate (they are used only in calk sign language), then we do not translate it. Distinguishing the phrase “з магазину” (“from the store”) and “до магазину” (“to the store”) in the sign language indicates a direction of movement. The word “магазин” (“store”) in the sign language corresponds to the sign “МАГАЗИН” (“STORE”).



**Fig. 4.** Translation algorithm from USpL to USL



**Fig. 5.** The architecture of the rule-based machine translation system into Ukrainian Sign Language

For convenience, the phrases, that have the same structure, can be separated in one group. For example, the phrases are “from the cinema”, “from the store”, “from the country”, etc., “to the cinema”, “to the store”, “to the country”, etc., can be separated as “from (to) the cinema (store, country, etc.)”. Selection of such groups can be used to develop ontology with a grammatical component for Ukrainian Sign and Spoken Languages.

**Table 4.** Translation using concept dictionary

№	Ukrainian Sign Language	Ukrainian Spoken Language
1.	“МОЯ” (“MY”)	“моя” (“my”)
2.	“МАМА” (“MOTHER”)	“мама” (“mother”)
3.	“НЕСТИ_СУМКА” (“CARRY_BAG”).	“несе сумку” (“carries the bag”)
4.	“МАГАЗИН” (“STORE”)	“з магазину” (“from the store”)

To test the machine translation system using grammar rules and concept dictionary, 360 sentences were taken from the corpus of parallel texts “Ukrainian Spoken Language – Ukrainian Sign Language”. The corpus consists of 1056 words of USpL and 984 words of USL. The dictionary of concepts is developed in accordance with the described relationships between concepts and contains 60 sign concepts and their corresponding translations in the Ukrainian Spoken Language.

The algorithm was tested using database of 360 sentences, which contained 60 concepts. As a result, 87% of sentences were translated automatically, 32% of which contained concepts, 13% were not translated due to lack of word – sign corresponding.

## 4 Conclusions

The rule-based machine translation into Ukrainian Sign Language using concept dictionary are described. The infological model of the concepts for the construction of Ukrainian Sign Language translation based on the analysis of linguistic relations between concepts in the Ukrainian and Ukrainian Sign Language was developed. The translation algorithm from Ukrainian Spoken Language to Ukrainian Sign Language based on concepts proposed. This algorithm was tested using database of 360 sentences from the corpus “USpL– USL”, which contained 60 concepts. The evaluation of the quality of the rule-based machine translation using the concepts has shown an improvement in the translation result on those sentences, which contain concepts. As a result, 87% of sentences were translated accurately, 32% of which contained concepts, 13% were not translated due to the lack of word to sign correspondence. The disadvantage of the system is a small amount of testing data (360 sentences).

It was investigated on the basis of the conducted experiments that it is necessary to group the concepts into grammatically correct constructions reproducing the expression content for high-quality translation into Ukrainian sign language.

## References

1. Veretennikova, N., Lozytskyi, O., Kunanets, N., Pasichnyk V. Information and technological service for the accompaniment of the educational process of people with visual impairments. In: CEUR Workshop Proceedings, Vol. I, pp. 290–301. ICTERI (2018).
2. Stone, C., Russell, D. Comparative Analysis of Depicting Signs in International Sign and Natural Sign Language Interpreting. *International Sign: Linguistic, Usage, and Status Issues*. Gallaudet University Press, pp. 65-83 (2015).
3. Forster, J., Stein, D., Ormel, E., Crasborn, O., Ney, H. Best Practice for Sign Language Data Collections Regarding the Needs of Data-Driven Recognition and Translation. In: LREC 2010 Proceedings. Malta (2010).
4. Barmak, O., Krak, Iu., Romanyshyn, S. Automated translation from inflectional languages to sign languages. *Information Content and Processing: International Journal* 1(1), 37-43 (2014).
5. Stokoe, W. C. The use of sign language in teaching English. *American Annals of the Deaf*. 120(4): 417-21 (1975).
6. San-Segundo, R., Pardo, J.M., Ferreiros, J., Sama, V., Barra-Chicote, R., Lucas, J.M., Sánchez, D., García, A. Spoken Spanish generation from sign language. *Interacting with Computers*, 22(2), 123–139(2010).
7. Nyst, V. Sign Languages in West Africa. In *Sign Languages - A Cambridge Language Survey*. Cambridge University Press, Cambridge, 405-432 (2010).
8. Abdel-Fattah, M. On the Translation of Modals from English into Arabic and Vice Versa: The Case of Deontic Modality. *Babel*, 51(1), 31-48 (2005).
9. Kulbida, S.V., Chepchina, I.I., Adamyuk, N.B., Ivanyusheva, N.V. *Ukrainian Sign Language. Pedagogical thought*, Kiev (2012).
10. Dreuw, P., Forster, J., Gweth, Y., Stein, D., Ney, H., Martinez, G., Verges, J. L., Crasborn, O., Ormel, E., Du, W., Hoyoux, T., Piater, J., Moya, J.M., Wheatley, M. *SignSpeak - Un-*

- derstanding, Recognition, and Translation of Sign Languages. In: 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010), pp. 65-72. Valletta, Malta (2010).
11. Hamandi, A. M., Karim, A., Bahjat, H. Proposed English to American Sign Language Translation System. *International Journal of Advances in Engineering Research* 10, pp. 53-72 (2015).
  12. Baldassarri, S., Royo-Santas, F. An automatic rule-based translation system to Spanish Sign Language. *New Trends on Human-Computer Interaction*, pp. 1–11 (2009).
  13. Okumura, A., Hovy, E. Building Japanese-English dictionary based on ontology for machine translation. In: *HLT '94 Proceedings of the workshop on Human Language Technology*, pp. 141-146. Association for Computational Linguistics, Plainsboro (1994).
  14. Tishchenko, O., Kulbida, S. Ukrainian sign language as an object of lexicographic parametrization. *Lexical Bulletin* 12, pp. 37-42. Institute of the Ukrainian language of the National Academy of Sciences of Ukraine, Kiev (2005).
  15. Davydov, M., Lozynska, O. Mathematical method of translation into Ukrainian sign language based on ontologies. *Advances in Intelligent Systems and Computing* 2, pp. 89-100 (2018).
  16. Davydov, M., Lozynska, O. Machine translation based on rules for translation into Ukrainian Sign Language. *International scientific and technical journal "Information Technologies and Computer Engineering"*, 1 (29), pp. 11-17 (2014). Vinnytsya. [in Ukrainian]