Computer Sign Language Interpreter Design

Mikhail Grif, Yulia Manueva, Andrey Kozlov, Natalie Ganelina and Evgeniy Tsoy

Novosibirsk State Technical University, Novosibirsk, Russia
grifmg@mail.ru, juleno4eknot1@rambler.ru, andrey.n.kozlov@gmail.com, ganelina@corp.nstu.ru, ebcoi@inter.nstu.ru

Abstract. Present systems for computer sign language translation are considered in the article, their advantages and disadvantages are identified. The general case of translation (from sounding Russian into Russian sign language and vice versa) is discussed. A new method for constructing a semantic unit for computer sign language interpretation is proposed. Lexical meanings of words are defined to match the correspondence "word-gesture". For each word certain lexical meaning from a set of alternatives based on the semantic analysis algorithm is confirmed. Semantic analysis algorithms for simple sentences are developed and introduced. A new method for translating Russian text into Russian sign language based on a comparison of syntactic structures is proposed. Special library to determine syntax constructions is developed. Present hardware and software to design the architecture of the prospective system for gesture recognition are considered.
Keywords: Russian sign language, computer sign language interpretation, semantic analysis, syntax constructions, sign language recognition tools.

1 Introduction

According to the Ministry of Public Health of Russian Federation a number of hard of hearing or deaf people in Russia is increasing every day and now is about 13 millions. It means that nearly every tenth person has different problems with hearing. A special sign language as well as certain schools, associations help such people in social rehabilitation [1]. Hard of hearing people in Russia are welcome to use a human sign language interpreter provided by the government for free, and amount of the service is 40 hours per year [4]. Obviously, it is not enough. Therefore multimedia computer systems to interpret speech into sign language and vice versa are of great interest and need. Fundamental parameters of sign languages are similar to sounding ones though sign languages use visual and kinetic ways instead of sound for information transfer. This allows to consider sign languages as natural human languages and analyze them by tools and concepts developed for and based on sounding languages [2].

Let's consider diachronic aspects of computer sign language interpretation systems evolution, review each of systems and define its highs and lows. Zardos is a system for translating English into sign language by interlingua - intermediate language. The advantage of applying logical propositions and named slots allows to obtain semantic information easily by common sense and other reasoning system components. Due to the system complexity it can be applied to the limited areas. Zardos infrastructure is introduced currently at parsing, interlingua, generating and component animation. The research now is focused on universal grammar design, morphology and lexicon for Irish sign language [4].

TEAM (TranslationfromEnglishtoASLbyMachine) is a system for machine translation form English to American sign language. Translating in TEAM includes two stages. The first one is translating an input sentence from English to intermediate representation taking into account syntactical, grammatical and morphological information. At the second stage the intermediate representation is displayed as a motion with a few parameters, which further are transformed into a big set of parameters controlling a human model presenting gestures. System flexibility makes it possible to modify the system for other sign languages [5].

ViSiCAST project is a lite system recording gestures and motions of a human sign language interpreter. Then coordinates of an interpreter's hands are analyzed to achieve a realistic avatar. ViSiCAST has innovations of two kinds: «BAF» format for a simplified system based on motion capture and SiGML - Signing Gesture Markup Language for the recorded animation transfer based on a gesture mark [6].

None of foreign systems can process voice input. It is a significant disadvantage for the systems developed for oral interpretation. Sign languages use spatial information around a speaker and this differs them from sounding languages. Therefore such a system should take into account specific character of sign language. The specifics of gesture presentation is included only in TEAM. Morphological and syntactical information is not enough to make a high-quality translation. Zardos tries to consider a semantic part of sign language. Translating by ViSiCAST involves a human interpreter and this is the main disadvantage of the system. All the systems mentioned above use an avatar to show gestures. The most realistic avatar is only in ViSiCAST. Semantic component of sounding and sign languages is not implied in any of the systems discussed which can be considered as the main demerit of them. Taking into account semantic component in translating is a big advantage for a system. Translation quality improves significantly by this component. For the best result it is necessary to consider specifics in semantics both of input and output languages.

The purpose of the current paper is describing an approach to designing computer Russian sign language interpretation system based on the semantic relation dictionary by V.A. Tuzov, introducing algorithm for syntactical constructions comparison and review of present hardware and software for sign speech recognition.

2 The general scheme of computer translation from Russian speech (text) to Russian sign language

All modern computer translation systems are based on trinomial model suggested by Shalyapina [7]. This model consists of three subsystems: analysis of the original text, analysis of interlingual transformations and sign speech design.

A computer translator should also provide:

1. text upload;
2. translation results visualization;
3. system structure flexibility and changeability (extending, vocabulary changing) [8].

Sign language visualization is the final stage of translation. It can be realized by two ways: taking photos of a person performing certain gestures which are stored at the system data base; designing a virtual character for gesture showing [9]. Computer sign language translation system includes the following modules [10]:

1. Module for Russian sounding language recognition (RSdL). The basic version of the system uses Google service;
2. Module for Russian text analysis including subsystems for morphological, syntactical and semantic analysis. This module to a large extent is based on free software (DIALING system) [11];
3. Module for translating Russian text to instructions (commands) for its transformation in Russian sign language (RSL) [12]. This module uses interrelation between grammatical systems of the languages mentioned above.
4. Module for RSL visualization based on computer model of a human interpreter (avatar). The avatar is controlled by the special control system based on Hamburg notation system [13].

3 Semantic analysis of Russian text

The meaning of a unit in Russian language depends on relations between it and the rest of units as well as on lexical and syntactical collocation with them. Language units are combined into groups by common features. For example, a group named 'a man' can include subgroups 'relative', 'nationality', 'occupation' [14]. The problem of lexical ambiguity is one of the priority tasks in translating from one language to another, because incorrect work of semantic module effects the meaning of the entire text. Homonyms and phraseological units are of special interest in this context.

O.S. Ahmanova describes a homonym as two words or a set of words which have identical phonemic presentation and differ by semantics or grammar [15]. The purpose of semantic analysis is to design meanings of words in sentences. To achieve this it is necessary to resolve lexical ambiguity (homonymity). Meanings of words are defined by the dictionary created by V.A. Tuzov [16]. Module for semantic analysis starts after morphological and syntactical analysis. The input information is a text formed after speech recognition. The process of semantic analysis includes two parts: initial analysis and semantic analysis. The first stage is based on the "Dialing machine translation from Russian into English" system [11]. The aim of preliminary stage is to get initial form of every word, morphological features and semantic relations. Semantic relation is basic semantic concept in Dialing. We determine semantic relation as a universal connection that is implied by a native speaker in a certain context. The basic idea of a relation list is the possibility to define connections in the text by the suggested relations or their composition. The majority of known semantic relations are considered to be universal [17].

In "Computer semantics of Russian language" by V.A. Tuzov each word is presented as a valence structure, which consists of a number of actants [16]. Each of actants is described by a set of parameters presented in (1):

$$A_i = \{CN_i, BL_j, SD_k, MD_l, C_m, SP_p\} \tag{1}$$

where CN_i – class number, $i = 1..N$, BL_j – basic lexeme, $j = 1...M$, SD_k – semantic description, $k = 1..P$, MD_l – morphological description, $l = 1..S$, C_m – comment, $m = 1..L$, SP_p – part of speech, $p = 1..W$.

Basic semantic information is in the class number. Additional information can also be included [17]. A classifier fragment is presented in the table 1. There is a number of words for a certain class number in the table.

Table 1. Classifier fragment

Class number	Class name	Number of words
\$1[a]	Noun	88
\$110	Noun. Abstract notion(AN)	4
\$1100/01	Noun. (AN) – abstract–concrete	2

Let's consider a verb "to go" ("ехать" in Russian) as an example. Semantic description for this verb is the following (in Russian):

1. \$15402(N%~ВЕРХОМ\$0(Oper01(!Им,ПОЕЗДКА\$15402(ПОДАТ:НЕЧТО\$1~!поДат,ОТКУДА:НЕЧТО\$1~!Откуда,КУДА:НЕЧТО\$1~!Куда,НАПРЕД:ТРАНСПОРТ\$121324(НАПРЕД:ЖИВОТНОЕ\$12422~!наПред))))) {г16нН 24947} <1>
2. N%~ПОЕЗДКА\$15402(Oper01(!Им,ПОЕЗДКА\$15402(ПОЧЕМУ:ПРИЧИНА\$1/37/05\ПРИКАЗ\$1526031~!Почему,ПОДАТ:НЕЧТО\$1~!поДат,ОТКУДА:НЕЧТО\$1~!Откуда,КУДА:НЕЧТО\$1~!Куда,ТВ:НАПРЕД:ТРАНСПОРТ\$121324~!Тв\!наПред))) {г16нН 24947} <2>

The abridgements above denote cases, prepositions, collocation words and other information according to the description notation. We can see that according to the semantic description the verb "to go" has two alternatives. A fixed expression "ехать верхом" in Russian meaning "to ride" was defined as a separate sense.

An entry in the computer semantic dictionary includes a title word and its interpretation at a semantic language. Most of words in the dictionary have several semantic descriptions. Among all parts of speech the most polysemantic is preposition. For example, the number of meanings for the preposition 'in' is 239. According to the semantic dictionary a data base model was designed (fig.1).

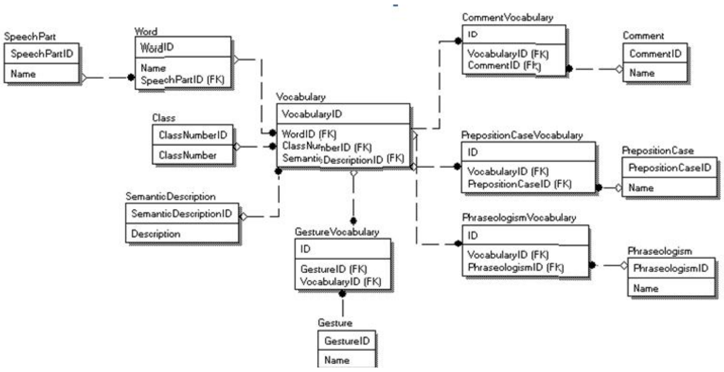

Fig. 1. System data base model

A model of semantic analysis system is presented at fig.2. The algorithm of semantic analysis includes two basic stages. At the stage of preliminary processing every sentence of the text is analyzed word-by-word. This step results in the list of initial forms for words of the sentence and in their morphological features and semantic relations. The information achieved at the first stage is an input for the second stage.

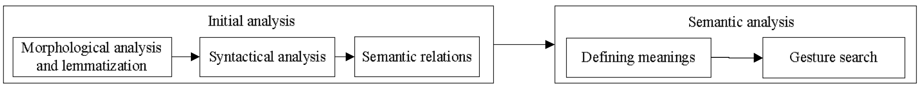

Fig. 2. Semantic analysis of text

The main task of the second stage is designing independent alternatives in words description and defining semantic and grammatical type of each of alternatives included in the description. These transformations are made by several steps. At the first step alternative descriptions for each of words are determined. The second step includes the following auxiliary activity: all alternatives of every word are numbered and identified; the number of the word semantic class is indicated; all arguments of semantic description are removed.

The description designed consists of a set of alternatives, each of them includes two basic parts: morphological part with the semantic class of a word indicated and semantic part. The first part of the alternative includes information about words that the current word can be collocated with, the second part defines the words which the current word can connect to itself. When assembling two constructions next to each other interact.

The next stage for semantic module is phraseological units processing. Phraseological units are defined as different alternatives in the semantic dictionary. To reduce a number of alternatives it is necessary to process phraseological units in the first place and then to remove redundant alternatives for these words. It will simplify further work of the module.

Prepositions processing includes two stages. Prepositional-case collocations are searched in the first place. The correct choice of semantic description for a preposition depends on the noun connected. As a result we have the only semantic description to every preposition.

The further actions of the analyzer depend on the type of a sentence. We define two types of a sentence: the first one includes sentences with only one word having several alternatives, sentences with several words having a set of alternatives form the second one. In the first case only one word is analyzed and there is no cycles. All wards are divided into 4 groups: verbs, nouns, adjectives, infinitives. Analyzing depends on the part of speech of the polysemantic word. Algorithms are different for each of cases. The second type implies a cycled analysis. Each iteration begins with checking a number of words with a set of alternatives. The sentence is under processing till each of words has the only semantic description. When every word corresponds to the only semantic feature a corresponding gesture is searched. Semantic analysis system results in the list of corresponding words and gestures. Software complex for semantic analysis of sentences is developed and introduced.

4 Translating Russian text into Russian sign language based on comparison of syntactical constructions

Let's consider the module for translating Russian text to Russian sign language (RSL) more carefully. We'll be guided at this stage by the results of linguistic research of Russian sign language used by deaf and hard-of-hearing citizens of Russian Federation. The results applied to developing computer sign language translator include peculiarities of vocabulary, word composition, morphology, syntax and semantics of RSL [18]. The purpose of the following syntactic and semantic transformations is to simplify Russian text by dividing sentences describing complicated situations into sequences of simple sentences. A unit which should be transformed is a sentence with full participle.

Applying any of transformation rules to the original sentence results in dividing the sentence Π into two parts that are formed further as separate sentences Π_1 and Π_2. Each of them denotes fewer number of situations than in Π and thus it is less complicated than the original one (Π).

While transforming the original sentence Π is not just divided into two parts but additionally modified. In particular, participle is replaced by the corresponding verb. Pronouns are added to denote identity of the objects mentioned.

The $ПРИЧ\_СУЩ$ (participle-noun) group consists of two parts: a noun group (denoted as Z) and a fragment of the type $ПРЧ$ (participle), which includes a participle agreed with the main word of the group Z.

A noun group can be just a noun (denoted as S) or a group of the $ПРИЛ\_СУЩ$ (adjective-noun) type. In a group of such a kind the main word is a noun also denoted as S.

P denotes full participles included in a ПРЧ (participle) fragment, which are agreed between each other on gender, number, time. Symbol P denotes each of these agreed participles. Transformation rules are applied to each of the participles in the same way.

Transformation of the participle to the verb it was generated from is presented in the formula (2).

$$P \to V(p) \tag{2}$$

where P is a participle and $V(p)$ is a verb.

In case a transformation (2) is needed it is applied to both agreed participles. The form of the verb $V(p)$ is identified by the type of the rule applied. Signs Z, S, P can be defined as a certain set of values for the indices: X - case (with standard symbols: $им$ - nominative, $род$ is for genitive, etc.), y– number, z – gender, l – person, t – tense, a – voice ($д$– active, c – passive).

$V(P)_{y,z,l,t,a}$ - a function, which value is a verb that a participle ($P_{y,z,l,t,a}$) was generated from; this verb has the same values of parameters y,z,l,t,a as P.

Applying transformation rules to a sentence $П$ is preceded by recognizing in S a certain syntactical construction including participle. So $П$ is divided into three parts: the beginning of the sentence - $НП$, identified syntactical construction, the end of the sentence - $КП$. In the sentence *Splash of raindrops falling down to its surface took the hum far away (Плеск дождевых капель, ниспадавших на его поверхность, далеко относил гул)* the $ПРЧ$ fragment corresponds to the part *Splash of drops falling down to its surface* (*Плеск капель, ниспадавших на его поверхность*) that follows a noun S = drops (капель), $НП$ - *splash* (*плеск*), $КП$ - *took the hum far away* (*далеко относил гул*). This situation is denoted as: [$НП$ S $ПРЧ$ $КП$].

The rules can replace or delete some parts of input text as well as change a sequence order of text parts. The general requirements for the rules to be applied: the rules of this type can be applied only if the previous analysis (morphological and syntactical) has recognized a $ПРИЧ\_СУЩ$ group and P is a full participle. This type of the rules includes six variants (2 rules for active participles and 4 rules for passive participles). Applying the rule to the sentence *Splash of raindrops falling down to its ebullient surface took the hum far away (Плеск дождевых капель, ниспадавших на его поверхность, далеко относил гул)* results in dividing this sentence into two simple: [*raindrops fell down to its surface (дождевые капли ниспадали на его поверхность)*] [*splash of these drops took the hum far away (плеск этих капель далеко относил гул)*].

We should mention that it is possible to translate using basic sequence of gestures in a sentence. Let's consider a software complex for computer sign language translating Russian text into written spoken Russian sign language [10]. Any Russian text (sentence) entered to the program is provided for morphological, syntactical and semantic analysis. Then a standard scheme of sentence in colloquial Russian sign language is formed: subject, attribute, adjunct, predicate, object.

This scheme is consecutively applied to simple statements. If a word in a sentence has a corresponding gesture than it is replaced by Hamburg notation system for this gesture. If there is no corresponding gesture - the word is translated by dactyl alphabet in the notation system. To remove polysemy in gesture performing a base of ontology is used. Possibility of parallel performing of statements is analyzed. Software for syntactical transformation of sentences with full participle constructions is a library that encapsulates text process. The logic of the process is isolated in the .NET library (.dll integration) which lets different applications with any interfaces (console, desktop, web etc.) address to the library regardless of the language.

For analysis of the text in Russian natural language the library provides:

1. text download and initialization of necessary DIALING components;
2. dividing text into sentences and organizing the procedure of its transformation;
3. morphological and syntactical analysis of each sentence (building dependency tree of a sentence);
4. translating each sentence using rules for syntactical transformation (building translation tree of a sentence).

5 Review of hardware and software tools for sign speech recognition

It is not enough just to recognize separate gestures and this is the main problem in sign speech recognition. It is necessary to recognize a continuous sign speech and its meaning. Let's review present hardware and software tools and basic approaches applied to resolve this problem. All devices used for gesture recognition can be divided into two groups: devices based on video data and tools using sensors, accelerometer, gyroscope and magnetometer. The best results can be achieved by systems that besides RGB sensor include IR Depth sensor for recognizing gestures at a small distance.

One of successful engineering solutions is the camera produced by Creative. It has two microphones, which allow to improve the quality of speech recognition, RGB and infrared cameras integrated.

The Kinect contact-free controller designed by Microsoft, includes RGB sensor, infrared emitter and sensor. Kinect can be applied at the distance from 1.8 to 3 meters, while Creative camera works at 15 centimeters to 1 meter. Although Creative camera is good for gesture recognition, identification and tracking of objects, hands, faces, its synthesis and voice recognition. And Kinect is focused on identification and recognition of gestures and lines of a human body to mark out a human in a house space [19, 20].

Leap Motion is another controller of this type. Leap Motion is a small device not bigger than a 3G-modem, includes 2 Monochromatic IR cameras and 3 Infrared LEDs. It can track all 10 fingers with precision at 1/100 millimeter. Viewing angle of the device is 150 degrees, also there is axis Z to define height [20].

One more interesting application of Leap Motion is virtual reality glasses Oculus Rift (fig.3). Controller is attached to the glasses and recognizes hands of the owner.

S. Loshkovska, S. Koceski (Editors): ICT Innovations 2015, Web Proceedings, ISSN 1857-7288

Fig. 3. Virtual reality glasses OculusRift and LeapMotion controller.

Fig. 4. The MYO bracelet by Thalmic Labs.

The second group of devices consists of tools based on accelerometer, gyroscope and magnetometer data. These tools are used basically for human-machine interaction. At present time new systems which receive accelerometer and gyroscope data and can be applied to sign language translation are coming into the market. Most of them are still prototypes, for example, rings, gloves, bracelets with sensors. Companies developing these tools try to achieve investments at crowd funding platforms. The products can be presented not earlier than the end of 2015.

The MYO bracelet designed by Thalmic Labs might be the device that resolve the problem of motion recognition (fig.4). The bracelet includes a three-axled accelerometer, three-axled gyroscope and three-axled magnetometer. But the most significant peculiarity of the device is a special sensor which receives information by electro pulses in muscles of a moving hand. The device is programmed to ignore random movements [20].

6 Conclusion

The method for designing a computer Russian sign language interpretation system is suggested in the paper. The diachronic aspect of computer sign language interpretation systems evolution is considered. Present sign language translation systems are considered, their advantages and disadvantages are discussed. A new method for constructing a semantic unit for computer sign language interpretation is proposed. Semantic analysis system results in the list of corresponding words and gestures. For each word certain lexical meaning from a set of alternatives based on the semantic analysis algorithm is confirmed. Semantic analyses algorithms for simple sentences are developed and introduced. The semantic analysis module can be improved and modified in the following priority directions: gesture base extension, composite sentences parsing, including the specifics of Russian sign language into analyzing algorithm. A new translation approach based on comparison of syntactical constructions is suggested. Special library to determine syntax constructions is developed. Present hardware and software to design the architecture of the future system for gesture recognition are considered.

References

1. Resolution of the RF Government №175 from 17.03.2011 "On the state program of Russian Federation "Accessible environment" for 2011 - 2015 years"

2. Prozorova E.V.: Russian sign language as a linguistic research subject: Problems of Linguistics, Moscow, pp. 44-61 (2007)
3. Resolution №608 from 25.09.2007 "On the order of sign language interpretation services for the Deaf at the expense of federal budget"
4. Cross modal comprehension in ZARDOZ an English to sign-language translation system, http://dl.acm.org/citation.cfm?id=1641450
5. A Machine Translation System from English to American Sign Language, http://dl.acm.org/citation.cfm?id=749243
6. Wakefield M.: VisiCAST Milestone: Final Report, pp. 1-97 (2002)
7. Shalyapina Z.M.: To the problem of designing a formal model for translating: In: Proc. "Translation theory and scientific foundations for interpreters training", P. 2. Moscow, M. Torez MSPIFL, pp. 165-172 (1975)
8. Grif M.G., Korolkova O.O.: [Development of computer sign language interpreter sounding speech at conversational Russian sign language]. In: Informatika:problemi, metodologiya, tehnologii. Materiali XI meshdunarodnoi nauchno-metodicheskoi konferencii [Computer science: problems, methodology, technology. Proceedings of the XI International Scientific Conference], pp. 206-208 (2011) (In Russ)
9. Grif M.G.: Metodi I tehnologii kompyuternogo surdoperevoda: uchebnoe posobie [Methods and techniques of computer sign language: a tutorial]. Novosibirsk: NGTU, p. 71 (2012) (In Russ).
10. Grif M.G., Korolkova O.O., Panin L.G., Timofeeva M.K.,Tsoi E.B.: Leksicheskie i grammaticheskie aspect razrabotki kompyuternogo surdoperevodchika russkogo yazika: monografiyu [Lexical and grammatical aspects of the development of computer sign language Russian language: monograph]. Novosibirsk: NGTU, p. 292 (2013) (In Russ).
11. Sokirko A.: Semantic vocabularies in automatic text processing (based on DIALING system papers). PhD thesis, MSPIFL. Moscow, p. 108 (2000)
12. Grif M.G., Demyanenko Y.A., Korolkova O.O., Tsoy Y.B.: Development of Computer Sign Language Translation Technology for Deaf People. In: Proc. of the 6th Internation Forum of Strategic Technology (IFOST 2011), pp. 674-677 (2011)
13. Grif M.G., Korolkova O.O., Tsoy Ye.B.: On Peculiarities of the Russian Language Computer Translation into Russian Sign Language for Deaf People. In: Proc. of the 2-nd Indo-Russian Joint Workshop on Computational Intelligence and Modern Heuristics in Automation and Robotics. Novosibirsk, NSTU, pp. 219-222 (2011)
14. Korolkova O.O.: Features of derivation in Russian Sign Language. In: Modern linguistics and intercultural communication: monograph: O.A. Berezina, E.N. Grushetskaya, T.V. Itskovich and others. Krasnoyarsk, Research and Innovation Center, pp. 98-152 (2012)
15. Ahmanova O.S.: Homonym dictionary for Russian language. Moscow, Soviet Encyclopedia, p. 451 (1974).
16. Tuzov V.A.: Computer semantics of Russian language, tutorial. Saint-Petersburg, SPSU, p. 391 (2003)
17. Collocation dictionary for Russian language, ed. P.N. Denisov and V.V. Morkovkin, 2nd ed. Moscow, Russian language, p. 685 (1983)
18. Korolkova O.O.: Russian Sign Language Morphology. In: Modern linguistics and intercultural communication: monograph, Book 3.: O.A. Berezina, N.V.Bizyukov, O,O, Korolkova and others. Krasnoyarsk, Research and Innovation Center, pp. 60-87 (2011)
19. Grif M.G., Kozlov A.N.: [Comparative analysis of software and hardware in patternrecognition of sign language]. In: Sbornik trudov NGTU – Proc. of the NSTU. Novosibirsk, pp. 63-72. (2014) (In Russ)
20. The MYO bracelet and world of tomorrow, http://www.3dnews.ru/645483