

## СОВРЕМЕННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 004.82

### ИСТОРИЧЕСКИЕ АСПЕКТЫ РАЗВИТИЯ СИСТЕМ КОМПЬЮТЕРНОГО СУРДОПЕРЕВОДА\*

М.Г. ГРИФ<sup>1</sup>, Ю.С. МАНУЕВА<sup>2</sup>

<sup>1</sup> 630073, РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, доктор технических наук, заведующий кафедрой АСУ. E-mail: grifmg@mail.ru

<sup>2</sup> 630073, РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, магистр информатики и вычислительной техники, аспирант кафедры АСУ. E-mail: juleno4eknot1@rambler.ru

В данной работе представлена проблема компьютерного сурдоперевода. Определена актуальность темы сурдоперевода, обусловленная тем, что почти каждый десятый россиянин имеет проблемы со слухом. Общая схема работы компьютерного сурдоперевода описана на основе трехчленной модели. Рассмотрен диахронический аспект развития систем перевода. В данном разделе приведено краткое описание основных свойств существующих систем. На основе характеристик рассмотренных систем составлена сравнительная таблица компьютерных сурдопереводчиков. В результате анализа систем были выявлены их преимущества и недостатки. Рассмотрен синхронический аспект данной проблемы. Основным недостатком рассмотренных систем является отсутствие учета семантической составляющей как звучащего, так и жестового языка. Для осуществления качественного перевода необходимо учитывать специфику двух языков. С этой целью был проведен анализ возможных подходов для решения данной проблемы. Основным недостатком рассмотренных выше систем является отсутствие учета семантической составляющей как звучащего, так и жестового языка. Для осуществления качественного перевода необходимо учитывать специфику двух языков. Рассмотрим возможные подходы для решения данной проблемы. В качестве разрешения проблемы омонимии звучащего языка приведена система машинного перевода Диалинг. В данной системе основополагающим семантическим понятием является понятие семантического отношения. Основная идея авторов списка отношений заключается в том, что связи в тексте можно определить через предложенные отношения. На основе первичного семантического анализа определяются соответствующие отношения исходного языка на переводящий язык. Таким образом решается проблема семантической неоднозначности в процессе перевода для звучащего языка. Специфика жестового языка заключается в наличии конструкций, которые называются классификаторными предикатами. Модель

---

\* Статья получена 27 марта 2015 г.

правильной синхронизации времени создания элементов жестового языка, а также выражение лица, взгляд, положение тела, жесты в ходе беседы являются важными для получения грамматически правильного предложения, которое передает правильный смысл. В заключение предложено представление для разработки семантической модели русского языка в системе компьютерного сурдоперевода.

**Ключевые слова:** компьютерный сурдоперевод, семантический анализ, классификаторные предикаты, временная синхронизация, омонимия, межъязыковые преобразования, виртуальный персонаж, язык жестов

DOI: 10.17212/2307-6879-2015-2-68-79

## **ВВЕДЕНИЕ**

Инвалиды по-прежнему остаются одной из наиболее социально уязвимых групп населения. По данным Министерства здравоохранения Российской Федерации, количество людей с ограниченными возможностями здоровья по слуху в России растет с каждым днем и уже составляет порядка 13 млн человек. В настоящий момент существуют язык жестов, специальные школы, различные объединения для реабилитации слабослышащих людей. Язык жестов – это способ коммуникации слабослышащих людей, в котором информация передается за счет движения рук, тела, мимики. Жестовый язык представляет собой естественный человеческий язык, и для его анализа используются методы, основанные на исследованиях звучащих языков [1]. В повседневной жизни жесты применяются каждым, но особое значение имеют для слабослышащих людей. Несмотря на то что жестовые языки задействуют не звуковой, а визуально-кинестический канал передачи информации, по своим фундаментальным свойствам они схожи со звучащими языками, что позволяет причислять их к естественным человеческим языкам и анализировать, используя методы и понятия, разработанные на материале звучащих языков. Но данные программы не решают в полной мере проблему коммуникации людей с дефектами слуха с обществом и государственной властью. Несмотря на все сложности взаимодействия с обществом, среди слабослышащих есть много талантливых людей. В качестве примера можно привести Виктора Гюго, Бетховена, Жан Жака Руссо. Этим людям ничто не мешало реализовать себя. В связи с этим возникает необходимость в использовании мультимедийных компьютерных систем, которые будут переводить речь на язык жестов и наоборот [2].

## 1. ОБЩАЯ СХЕМА КОМПЬЮТЕРНОГО СУРДОПЕРЕВОДА

Системы компьютерного перевода основываются на трехчленной модели Шалепиной. Данная модель состоит из трех подсистем: анализа исходного текста, анализа межкузыковых преобразований и синтеза жестовой речи. Компьютерный переводчик также должен обеспечивать загрузку текста, визуализацию результатов перевода, возможность изменения структуры системы (расширение, изменение словарей).

Модель системы компьютерного сурдоперевода представлена на рис. 1. Рассмотрим более подробно этап межкузыковых преобразований. Трансфер заключается в преобразовании синтаксической и семантической информации об исходном тексте и семантико-синтаксического представления о переводных эквивалентах в синтактико-семантическое представление выходного. Сначала анализируется синтаксическая структура предложения и его семантика. Затем по правилам машинной грамматики структура входного предложения преобразуется в промежуточную структуру. На этапе синтеза задается синтаксическая структура выходного предложения, происходит выбор эквивалентов для языка перевода по синтаксическим и семантическим признакам [3]. Заключительным этапом перевода является визуализация жестового языка. Существует два варианта реализации: фотосъемка модели, демонстрирующей определенные жесты, и сохранение их в базе данных системы; создание для демонстрации жестов виртуального персонажа [4].

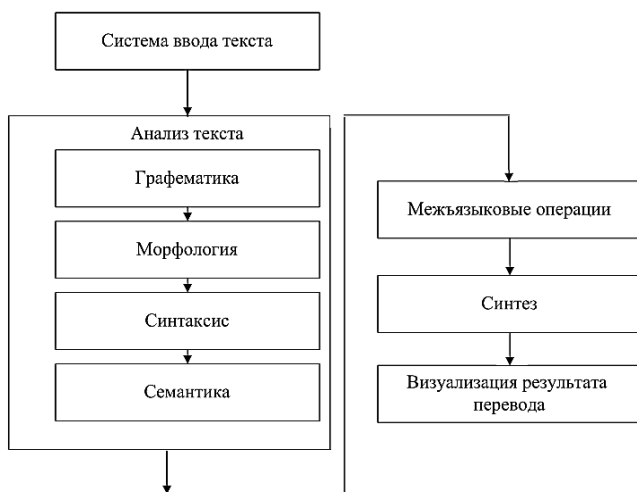


Рис. 1. Общая схема компьютерного сурдоперевода

## 2. ДИАХРОНИЧЕСКИЙ АСПЕКТ

Рассмотрим системы компьютерного перевода и проведем их сравнительный анализ. Система Zardoz представляет собой систему перевода с английского на язык жестов с использованием интерлингвы. Архитектура системы Zardoz представляет собой модульную систему, работающую в определенной структуре, которая представляет собой фреймовую структуру. Архитектура системы представлена на рис. 2.

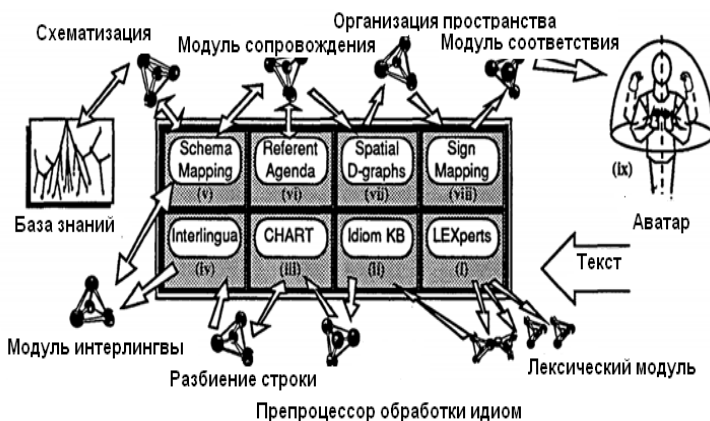


Рис. 2. Архитектура системы Zardoz

На сегодняшний день внедрена инфраструктура системы Zardoz, включающая разбор, интерлингву, генерацию и анимацию компонентов, но еще только предстоит внедрить всеобъемлющую грамматику и лексикон [5].

Система TEAM – это первая система 3D анимационного машинного перевода с английского на американский жестовый язык, учитывающая не только языковые, но и визуальную и пространственную информацию, связанную с жестовым языком. В данной системе происходит учет специфики жестового языка. Человек в процессе перевода в данной системе перевода не привлекается. Система является универсальной, поэтому возможно ее применение для других языков. Семантическая составляющая языка не учитывается данной системой. Для вывода предложения на американском жестовом языке необходима детально проработанная 3D модель человека, представленного на рис. 3 [6].

Проект ViSiCAST направлен на удовлетворение потребностей людей с ограниченными возможностями по слуху. Модель человека системы ViSiCAST представлена на рис. 4 [7].

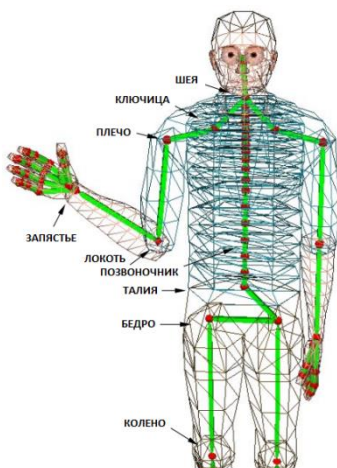


Рис. 3. Модель человека системы TEAM



Рис. 4. Модель человека системы ViSiCAST

Система ASL Workbench – это система машинного перевода на американский жестовый язык, разработанная Диамондом Спирсом. ASL Workbench основана на современных лингвистических исследованиях жестового языка. Для сопоставления английского слова фразы с аналогичными фразами американского языка жестов в системе используется передача конкретной лексики.

В случае возникновения выбора из нескольких лексических значений или при возникновении других проблем при переводе для разрешения неоднозначных ситуаций отправляется запрос к пользователю [8].

Ни одна из зарубежных систем не может обрабатывать входную информацию, поступающую в виде голоса. Для систем перевода, которые направлены именно на устный перевод, этот недостаток является существенным. Жестовые языки от звучащих отличаются тем, что используют пространственную информацию вокруг говорящего. Следовательно, в данных системах необходим учет специфики жестового языка. Специфика воспроизведения жестов учитывается только в системе Team.

Для более качественного перевода недостаточно только морфологической и синтаксической информации. В системе Zardoz делаются попытки учета семантической составляющей жестового языка. Технология перевода в системе ViSiCAST включает привлечение человека в процесс перевода, что является основным недостатком данной системы. Все системы отображают жесты с использованием аватара, но только в системе ViSiCAST достигнута максимальная реалистичность.

В таблице приведен сравнительный анализ рассмотренных систем.

**Сравнительный анализ систем**

| Характеристика                    | Zardoz             | Team                      | ViSiCAST     | ASL Workbench |
|-----------------------------------|--------------------|---------------------------|--------------|---------------|
| Универсальность                   | Ведутся разработки | Применима к другим языкам | Отсутствует  | Отсутствует   |
| Учет специфики жестового языка    | Отсутствует        | Учитывается               | Отсутствует  | Отсутствует   |
| Семантическая информация          | Учитывается        | Отсутствует               | Отсутствует  | Учитывается   |
| Привлечение человека при переводе | Отсутствует        | Отсутствует               | Используется | Используется  |

### 3. СИНХРОНИЧЕСКИЙ АСПЕКТ

Основным недостатком рассмотренных выше систем является отсутствие учета семантической составляющей как звучащего, так и жестового языка. Для осуществления качественного перевода необходимо учитывать специфику двух языков. Рассмотрим возможные подходы для решения данной проблемы.

В качестве разрешения проблемы омонимии звучащего языка рассмотрим систему машинного перевода Диалинг. В системе Диалинг, разработанной

А. Сокирко, основополагающим семантическим понятием является понятие семантического отношения. Определим семантическое отношение как универсальную связь, которая усматривается носителем языка в конкретном контексте.

Любой текст представляет собой набор различных отношений. Основная идея авторов списка отношений заключается в том, что связи в тексте можно определить через предложенные отношения. На основе первичного семантического анализа определяются соответствующие отношения исходного языка на переводящий язык. Трансфер заключается в поиске соответствий между языками. Таким образом решается проблема семантической неоднозначности в процессе перевода для звучащего языка [9].

Отличительной особенностью жестового языка является наличие специальных конструкций, называемых классификаторными предикатами. Американский ученый М. Хунерфауф разработал программу компьютерной анимации американского языка жестов (ASL) с учетом классификаторных предикатов. Общая схема создания таких конструкций представлена на рис. 5.

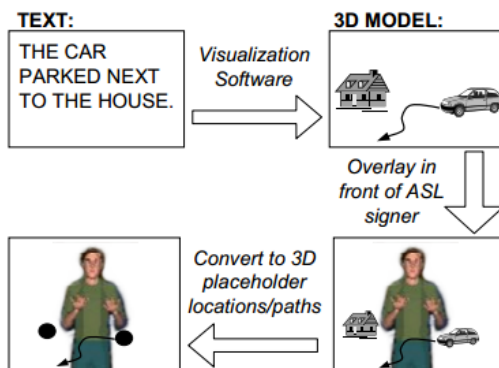


Рис. 5. Общая схема создания классификаторных предикатов

Использование 3D-визуализаций в системе машинного перевода с английского на американский язык жестов решает проблему учета специфики жестового языка [10].

Модель правильной синхронизации времени создания элементов жестового языка является важной для получения грамматически правильного предложения, которое передает правильный смысл. Такая деталь, как взгляд, направленный на правильное место в классификаторном предикате, может оказать существенное влияние на передаваемое значение.

В дополнение к информации, содержащейся в речи, участники могут передать информацию с помощью изменений в голосе, интонации и паузах. Вы-

ражение лица, взгляд, положение тела, жесты в ходе беседы также существенно влияют на смысл в жестовой речи [11].

Исследователи естественного языка рассуждали о коммуникативном сигнале в различных плоскостях: как написание строки текста, как речевой аудиосигнал (с просодией, паузами, глубиной и интонацией) и как смешанный тексто-речевой сигнал с согласованными визуальными элементами (картами, схемами, диаграммами и т. д.). Например, исследователи изучают, как координировать представление текста с диаграммами и фотографиями [12, 13].

Другие изучают «воплощение разговорными агентами», компьютерными персонажами, которые взаимодействуют с пользователями с помощью речи, взгляда, выражения лица, тела и жестов [14, 15].

## **ЗАКЛЮЧЕНИЕ**

В данной работе приведена общая схема работы компьютерного сурдоперевода на основе трехчленной модели. Рассмотрен диахронический аспект развития систем перевода, в котором представлено краткое описание основных свойств существующих систем. На основе характеристик рассмотренных систем составлена сравнительная таблица компьютерных сурдопереводчиков.

В результате анализа систем были выявлены их преимущества и недостатки. Рассмотрен синхронический аспект данной проблемы. Основным недостатком рассмотренных систем является отсутствие учета семантической составляющей как звучащего, так и жестового языка. Учет семантической составляющей в процессе перевода является большим преимуществом системы, обладающей таким свойством. Качество перевода заметно повышается за счет этого улучшения. Для достижения наилучшего результата необходимо учитывать особенности семантики исходного языка и языка перевода. Были рассмотрены две системы: Диалинг и американская система сурдоперевода.

Условно необходимая модель должна состоять из двух частей: первая часть представляет собой семантику звучащего языка, вторая отражает специфику жестового языка. Основное назначение семантической модели звучащего языка заключается в разрешении проблемы омонимии. Для естественности перевода необходимо учесть классификаторные предикаты – специальные конструкции, использующиеся в жестовом языке, а также временную и пространственную составляющую генерации жестов (взгляд, мимика, положение тела). На основе модели, представляющей классификаторные предикаты в американской системе перевода, требуется разработать аналогичную систему для системы перевода с русского языка на русский язык жестов. Учет озвученных требований приведет к более качественному и естественному переводу, понятному своей целевой аудитории.



## СПИСОК ЛИТЕРАТУРЫ

1. *Прозорова Е.В.* Российский жестовый язык как предмет лингвистического исследования // Вопросы языкознания. – 2007. – № 1. – С. 44–61.
2. *Гриф М.Г., Мануева Ю.С., Козлов А.Н.* Построение системы компьютерного сурдоперевода русского языка // Труды СПИИРАН. – 2014. – Вып. 6 (37). – С. 170–183. – doi: <http://dx.doi.org/10.15622/sp.37.11>.
3. *Гриф М.Г., Тимофеева М.К.* Проблема автоматизации сурдоперевода с позиции прикладной лингвистики // Сибирский филологический журнал. – 2012. – № 1. – С. 211–219.
4. *Гриф М.Г.* Методы и технологии компьютерного сурдоперевода: учебное пособие. – Новосибирск: Изд-во НГТУ, 2012. – 71 с.
5. *Veale T., Conway A.* Cross modal comprehension in ZARDOZ an English to sign-language translation system // Proceedings of the 7th International Workshop on Natural Language Generation, INLG'94, Kennebunkport, Maine, USA, June 21–24, 1994. – Kennebunkport, 1994. – P. 249–252.
6. A machine translation system from English to American sign language / L. Zhao, K. Kipper, W. Schuler, C. Vogler, N. Badler, M. Palmer // Lecture Notes in Computer Science. – 2000. – Vol. 1934. – P. 54–67. – doi: 10.1007/3-540-39965-8\_6.
7. *Wakefield M.* VisiCAST Milestone: final report N IST-1999-10500 / Information Societies Technology. – [S. l.], 10 December 2002. – 97 p.
8. *Speers A.L.* Representation of American sign language for machine translation: doctoral dissertation / Georgetown University, department of linguistics. – Washington, 2001. – 136 p.
9. *Сокирко А.* Семантические словари в автоматической обработке текста: по материалам системы Диалинг: дис. ... канд. техн. наук: 05.13.17 / Российский государственный гуманитарный университет. – М., 2000. – 120 с.
10. *Huenerfauth M.* Improving spatial reference in American sign language animation through data collection from native ASL signers Proceedings of the 5th International Conference on Universal Access in Human-Computer Interaction, UAHCI 2009. Pt. 3: Applications and Services, San Diego, California, USA, July 19–24, 2009. – Berlin; Heidelberg: Springer, 2009. – P. 530–539. – doi: 10.1007/978-3-642-02713-0\_56.
11. *Huenerfauth M.* Spatial, temporal, and semantic models for American sign language generation: implications for gesture generation // International Journal of Semantic Computing. – 2008. – Vol. 2, iss. 1. – P. 21–45. – doi: 0.1142/S1793351X08000336.
12. *Andre E.* The generation of multimedia presentations // A Handbook of Natural Language Processing: Techniques and Applications for the Processing of Language as text. – New York: Marcel Dekker, 2000. – P. 305–327.

13. *Foster M.E., White M., Setzer A.* Multimodal generation in the COMIC dialogue system // Proceedings of the ACL 2005 on Interactive Poster and Demonstration Sessions, ACLdemo'05, 12–14 June 2005. – Stroudsburg, Pennsylvania, USA: Association for Computational Linguistics, 2005. – P. 40–48. – doi: 10.3115/1225753.1225765.
14. *Kopp S., Tepper P., Cassell J.* Towards integrated microplanning of language and iconic gesture for multimodal output // 6th International conference on Multimodal interfaces, 6–9 August 2004. – USA: State College, 2004. – P. 136–144.
15. *Cassell J., Stone M., Yan H.* Coordination and context-dependence in the generation of embodied conversation // International Natural Language Generation Conference, 2–5 September 2000. – Stroudsburg, Pennsylvania, USA, 2000. – P. 171–178.

**Гриф Михаил Геннадьевич** – заведующий кафедрой автоматизированных систем управления Новосибирского государственного технического университета, доктор технических наук, профессор. Область научных интересов: методы проектирования и оптимизации человеко-машинных систем, системы компьютерного сурдоперевода. Число научных публикаций – 220. E-mail: grifmg@mail.ru

**Мануева Юлия Сергеевна** – аспирант кафедры автоматизированных систем управления Новосибирского государственного технического университета. Область научных интересов: семантические модели перевода. Число научных публикаций – 12. E-mail: juleno4eknot1@rambler.ru

## Historical aspects of computer interpretation system to sign language\*

M.G. Grif<sup>1</sup>, J.S. Manueva<sup>2</sup>

<sup>1</sup> Novosibirsk State Technical University, 20 K. Marx Prospekt, Novosibirsk, 630073, Russian Federation, D. Sc. (Eng.), professor. E-mail: grifmg@mail.ru

<sup>2</sup> Novosibirsk State Technical University, 20 K. Marx Prospekt, Novosibirsk, 630073, Russian Federation, Master student (Eng.), postgraduate student at the department of automated control systems. E-mail: juleno4eknot1@rambler.ru

This paper presents the problem of computer sign language interpretation system. Relevance of the topic of sign language translation is showed, because every tenth Russians have hearing problems. General scheme of the computer sign language interpretation system is described on the base of trinomial-model. Diachronic aspect of transfer systems is considered. In this sec-

---

\* Received 27 March 2015.

tion a brief description of the main characteristics of existing systems is shown. Based on the characteristics of the considered systems comparative table of computer sign language interpreters is composed. During an analysis of these systems their advantages and disadvantages have been identified. The synchronic aspect of this problem is considered. The main disadvantage of these systems is the lack of consideration of the semantic component of the sound language and sign language. To high-quality translation must take into account the specifics of the two languages. For this purpose, the analysis of possible approaches to solve this problem. The main disadvantage of the above systems is-is the lack of consideration of the semantic component as the sound and sign language. High-quality translation must take into account the specifics of the two languages. Possible approaches to solve this problem are considered. As permitted to problems of homonymy-sounding language machine translation system Dialing is described. In this system, the basic concept is a semantic relationship. The basic idea of the authors of the list of relations is that the links in the text can be determined by the proposed relationship. On the basis of primary semantic analysis are determined by the corresponding ratio of the source language to the target language. Thus, the problem of semantic ambiguity in the translation process for the sounding language is solved. Specificity of sign language is the presence of structures that are called classifier predicates. Model of correct timing synchronization of creation of elements of sign language and facial expression, gaze, body posture, gestures during a conversation are important for grammatical correct sentence, which sends the correct meaning. In conclusion, the representation is proposed for the development of the semantic model of the Russian language in the computer sign language. It is necessary to take into account specific sound language and sign language. To solve this problems homonyms are defined and classifier predicates are processed.

**Keywords:** computer interpretation system to sign language, semantic analysis, classifier predicates, timing synchronization, homonyms, virtual person, sign language, cross-language conversion

DOI: 10.17212/2307-6879-2015-2-68-79

## REFERENCES

1. Prozorova E.V. *Rossiiskii zhestovyi yazyk kak predmet lingvisticheskogo issledovaniya* [The Russian gesture language as subject of linguistic study]. *Voprosy yazykoznaniya – Topics in the study of language*, 2007, no. 1, pp. 44–61.
2. Grif M.G., Manueva Yu.S., Kozlov A.N. Postroenie sistemy komp'yuternogo surdoperevoda russkogo yazyka [Computer sign language interpretation system development of Russian language]. *Trudy SPIIRAN – SPIIRAS Proceedings*, 2014, iss. 6 (37), pp. 170–183. doi:<http://dx.doi.org/10.15622/sp.37.11>
3. Grif M.G., Timofeeva M.K. Problema avtomatizatsii surdoperevoda s pozitsii prikladnoi lingvistiki [Sign language machine translation as a problem of applied linguistics]. *Sibirskii filologicheskii zhurnal – Siberian Philological Journal*, 2012, no. 1, pp. 211–219.
4. Grif M.G. *Metody i tekhnologii komp'yuternogo surdoperevoda* [Methods and techniques of computer sign language]. Novosibirsk, NSTU Publ., 2012. 71 p.

5. Veale T., Conway A. *Cross modal comprehension in ZARDOZ an English to sign-language translation system*. Proceedings of the 7th International Workshop on Natural Language Generation, INLG'94, Kennebunkport, Maine, USA, June 21–24, 1994, pp. 249–252.
6. Zhao L., Kipper K., Schuler W., Vogler C., Badler N., Palmer M. A machine translation system from English to American sign language. *Lecture Notes in Computer Science*, 2000, vol. 1934, pp. 54–67. doi: 10.1007/3-540-39965-8\_6
7. Wakefield M. *VisiCAST Milestone: final report no. IST-1999-10500*. Information Societies Technology, 10 December 2002. 97 p.
8. Speers A.L. *Representation of American sign language for machine translation*. Doctoral dissertation. Georgetown University, department of linguistics. Washington, 2001. 136 p.
9. Sokirko A. *Semanticheskie slovari v avtomaticheskoi obrabotke teksta: po materialam sistemy Dialing*. Diss. kand. tekhn. nauk [Semantic dictionaries in automatic text processing: based on materials of the DIALING system. PhD eng. sci. diss.]. Moscow, 2000. 120 p.
10. Huenerfauth M. Improving spatial reference in American sign language animation through data collection from native ASL signers. Proceedings of the 5th International Conference on Universal Access in Human-Computer Interaction, UAHCI 2009. Pt. 3: Applications and Services, San Diego, California, USA, July 19–24, 2009, pp. 530–539. doi: 10.1007/978-3-642-02713-0\_56
11. Huenerfauth M. Spatial, temporal, and semantic models for American sign language generation: implications for gesture generation. *International Journal of Semantic Computing*, 2008, vol. 2, iss. 1, pp. 21–45. doi: 0.1142/S1793351X08000336
12. Andre E. The generation of multimedia presentations. *A Handbook of Natural Language Processing: Techniques and Applications for the Processing of Language as text*. New York, Marcel Dekker, 2000, pp. 305–327.
13. Foster M.E., White M., Setzer A. Multimodal generation in the COMIC dialogue system. *Proceedings of the ACL Interactive Poster and Demonstration Sessions*. Association for Computational Linguistics, 12–14 June 2005, pp. 40–48.
14. Kopp S., Tepper P., Cassell J. Towards integrated microplanning of language and iconic gesture for multimodal output. *Proceedings of 6th International Conference on Multimodal Interfaces*, USA, State College, 6–9 August 2004, pp. 136–144.
15. Cassell J., Stone M., Yan H. Coordination and context-dependence in the generation of embodied conversation. *International Natural Language Generation Conference*, Stroudsburg, USA, 2–5 September 2000. pp. 171–178.