

М.Г. Гриф, М.К. ТИМОФЕЕВА
**ИНТЕРЛИНГВА В СИСТЕМАХ МАШИННОГО
ПЕРЕВОДА ДЛЯ ЖЕСТОВЫХ ЯЗЫКОВ**

Гриф М.Г., Тимофеева М.К. **Интерлингва в системах машинного перевода для жестовых языков.**

Аннотация. В статье представлена обзорная характеристика систем машинного перевода, предназначенных для жестовых языков и базирующихся на использовании семантического языка-посредника (интерлингвы). Рассматриваются принципы построения интерлингвы в двух разработках: ZARDOZ (многоязыковая система, ориентированная на ряд жестовых языков, в частности, американский, ирландский, японский) и Multi-path (система с многоходовой архитектурой, предназначенная для перевода с английского языка на американский жестовый язык). Обсуждение архитектуры этих систем и требований, обусловленных спецификой жестовых языков, проводится на содержательном уровне, без углубления в математические и технические детали. Цель статьи: способствовать пониманию проблем и стратегий формализации семантики жестовых языков в рамках систем автоматического перевода.

Ключевые слова: жестовый язык, автоматический перевод, интерлингва, язык-посредник, семантика, модель мира

Grif M.G., Timofeeva M.K. **Interlingua in machine translation systems for sign languages.**

Abstract. The article reviews those machine translation systems for sign languages that are based on intermediate semantic language (interlingua). The constructing principles of two systems are considered: ZARDOZ (multilingual system for translating spoken language into a number of different sign languages, in particular, Irish, American, Japanese) and Multi-path (system with multiple processing pathways for translating spoken English into American sign language). The architecture of those systems and the requirements stipulated by the peculiarities of sign languages are discussed on conceptual level, without going into mathematical and technical details. The purpose of this article is to contribute to perfect understanding of problems and strategies of formalizing semantics within machine translation systems for sign languages.

Keywords: sign language, machine translation, interlingua, intermediate language, semantics, world model

1. Введение. Жестовый язык — один из трёх типов средств коммуникации, предназначенных для людей с проблемами слуха. Остальные два типа не являются самостоятельными знаковыми системами, а лишь кодируют определённые составляющие звучащего языка¹: дактильная азбука кодирует буквы алфавита, калькирующая речь — слова. В отличие от них жестовый язык — полноценная самостоятельная

¹ Под термином «звучащий язык» здесь подразумевается язык, имеющий звуковую форму. При этом объектом рассмотрения не обязательно должна быть именно звуковая форма, в частности, автоматический перевод может быть ориентирован на преобразование письменной формы звучащего языка в жестовый язык.

знаковая система, отличающаяся от звучащих языков. Жестовые языки вошли в сферу интересов лингвистики по историческим меркам совсем недавно, в середине XX века, прежде всего благодаря деятельности американского лингвиста Уильяма Стоке. Лингвистическими исследованиями русского жестового языка занимаются, например, Г.Л. Зайцева, А.А. Кибрик, Е.В. Прозорова, В.И. Киммельман [5, 7, 8].

Одна из стратегий машинного перевода состоит в построении промежуточного языка (языка-посредника или интерлингвы), мыслимого как универсальное, не зависящее от конкретных языков, средство представления смыслов. Такая стратегия перевода восходит к идеям Уоррена Уивера, высказанным в 1955 г. [16].

Надо заметить, что корректнее было бы говорить только об относительной универсальности, так как реально это свойство оценивается лишь по отношению к определённому множеству известных, или принимаемых во внимание, языков. Нет оснований утверждать, что нам известно всё многообразие потенциально возможных типов языков. Одним из подтверждений этого оказывается включение в сферу интересов разработчиков систем машинного перевода весьма своеобразных знаковых систем — жестовых языков, предназначенных для общения глухих и слабослышащих людей. Анализ путей построения семантического языка-посредника в рамках систем машинного перевода для жестовых языков представляет интерес не только собственно для задачи автоматизации межязыкового перевода, но и для многообразных задач, связанных с изучением или компьютерной имитацией различных аспектов мыслительной и языковой деятельности человека.

Машинный перевод для жестовых языков начали создавать значительно позже, чем для звучащих языков. Первые системы такого рода появились в начале 90-х годов прошлого века [15]. Перевод в них, как правило, однопольный (текст звучащего языка преобразуется в текст жестового языка), результат перевода (жесты) визуализируется на экране монитора посредством движений анимированного изображения человека. Наиболее интенсивно велись разработки по автоматизации перевода с английского языка на американский жестовый язык ASL (амслен).

Для русского жестового языка действующих систем машинного перевода пока нет. В 1998 г. в Институте социальной реабилитации Новосибирского государственного технического университета был создан первый компьютерный сурдопереводчик с русской звучащей речи на калькирующую речь. В настоящее время планируется разрабо-

тать на этой основе систему компьютерного сурдоперевода для русского жестового языка [2, 3].

Работа по созданию автоматического сурдоперевода ведётся в Санкт-Петербургском институте информатики и автоматизации РАН [6]. Задача распознавания русского жестового языка в системе автоматизированного сурдоперевода обсуждается в [1].

Лингвистические проблемы, связанные с построением систем машинного перевода для жестовых языков, рассмотрены в [4]. Многие сложности обусловлены тем, что жестовые языки — это языки без письменной формы, и перевод со звучащего языка на жестовый состоит в преобразовании одной модальности (звуковой или графической) в другую (пространственную).

Одна из основных проблем, наиболее существенно затрагивающих семантический уровень, обусловлена наличием в жестовых языках классификаторных предикатов, то есть иконических знаков², служащих для представления пространственных движений объектов, обсуждаемых в переводимом тексте. В таком знаке одна и та же форма руки может обозначать целый класс внешне похожих объектов, а демонстрация знака в целом — отображать траекторию и скорость перемещения объекта, изменения его формы и т.д. На основании анализа различных стратегий машинного сурдоперевода в [11] делается вывод о том, что удовлетворительный перевод для жестового языка, содержащего классификаторные предикаты, возможен только при условии использования интерлингвы.

Цель статьи: введение в проблему формализации семантики жестовых языков в системах автоматического сурдоперевода, основывающихся на использовании интерлингвы.

2. Типы систем машинного перевода для жестовых языков.

Существующие стратегии автоматического сурдоперевода (как и в случае звучащих языков) принято подразделять на два типа: перевод на основе правил (rule-based), и перевод на основе данных (data-based). В рамках стратегий первого типа система правил перевода составляется разработчиками на основе знания обоих языков, в рамках стратегий второго типа — она выводится автоматически на основе анализа языковых данных, опирающегося на минимальные предварительные знания об устройстве рассматриваемых языков (например, статистический перевод на основе корпуса параллельных текстов).

² Согласно классификации Ч. Пирса иконический знак – это знак, связанный с обозначаемым объектом отношением подобия (по каким-либо наблюдаемым признакам: очертанию, характеру движения и т.д.).

Для жестовых языков первый тип стратегий в настоящее время превалирует, что объясняется, в частности, сложностью создания корпусов жестовых языков. Например, ни одна из систем автоматического сурдоперевода, представленных в обзоре [11], не использует статистический перевод. Автор объясняет это сложностью аннотирования (разметки) жестовой речи и недостаточным объемом корпусных данных для жестовых языков. Однако по мере накопления корпусов ситуация меняется, и стратегии второго типа привлекают всё больший интерес. Обзор таких методов перевода приведён в уже указанной выше диссертации Сары Морриссей [15], автор которой предлагает также свой вариант перевода на основе данных (*data-based*). Существуют варианты автоматического сурдоперевода, совмещающие оба типа стратегий, например, в [19] описана тайваньская система, использующая и трансфер, и статистический перевод. Для русского жестового языка стратегии второго типа в настоящее время не представляются реализуемыми, поскольку корпусов этого языка пока нет.

Системы, реализующие стратегию первого типа (*rule-based*), могут различаться по степени глубины анализа переводимого текста. При *прямом переводе* синтаксический и семантический анализ не проводятся, перевод осуществляется на основе установления лексических соответствий. Перевод типа *трансфер* базируется на более глубоком анализе, затрагивающем синтаксис и, возможно, семантику. Синтаксическое / семантическое представление переводимого текста преобразуется (на основе правил, заложенных в систему перевода) в синтаксическое / семантическое представление другого языка, которое затем служит основой генерации выходного текста.

Ещё более глубокий анализ текста происходит при переводе, использующем искусственный *язык-посредник*, или *интерлингву*. Предполагается, что этот промежуточный язык позволяет представить любой смысл, выраженный на любом языке. Содержание переводимого текста сначала выражается на языке-посреднике, а затем переводится с языка-посредника на жестовый язык.

Опыт разработки систем машинного перевода показывает, что чем более универсален язык-посредник, тем сложнее создать качественную и работоспособную систему перевода. По сути, достаточно, чтобы язык-посредник был универсальным средством описания смыслов только для тех языков, для которых система строится. Например, если она ориентирована только на перевод со звучащего языка на жестовый, то достаточно создать язык-посредник, являющийся универсальным именно для этих двух языков. Такую интерлингву нельзя назвать

нейтральной по отношению к выбору описываемого языка (language-neutral), поэтому в [12] она именуется «почти-интерлингва» (“near-interlingua”).

Ограниченно универсальный язык-посредник («почти-интерлингва») используется в системе с многоходовой архитектурой перевода Multi-path [11, 12].

Авторы многоязыковой системы машинного перевода ZARDOZ [16–18] называют созданный ими язык-посредник универсальным. Однако, как уже говорилось, представляется более корректным говорить лишь об универсальности относительно определённого множества языков.

Языки-посредники, используемые в указанных двух системах, фактически моделируют мир, обсуждаемый посредством языка. Одно из преимуществ такого подхода состоит в том, что формализация знаний о мире позволяет включать в процесс перевода процедуры логического вывода, основанные на соображениях из области здравого смысла.

В [15] к числу разработок, использующих интерлингву, отнесена также система TEAM (Translation from English to ASM by Machine), описанная в [20] и предназначенная для машинного перевода с английского языка на амслен. Вместе с тем, в обзоре [11] система TEAM рассматривается как пример реализации стратегии перевода типа трансфер на уровне синтаксиса (для представления которого используются древесные грамматики STAG — Synchronous Tree Adjoining Grammar). В TEAM синтаксическая схема входного английского текста преобразуется в «промежуточное представление» (intermediate representation, IR) перевода данного текста на амслен. Это промежуточное представление уже организовано в соответствии с синтаксисом амслена и далее интерпретируется как описание движений анимированного компьютерного изображения человека (для чего в промежуточное представление дополнительно включаются параметры, определяющие мимические движения). Система TEAM не ориентирована на перевод классификаторных предикатов, что и позволяет ограничиться межязыковым преобразованием на уровне синтаксиса. Промежуточное представление — это результат такого преобразования, поэтому позиция, высказанная в [11] и следующая традиционному пониманию интерлингвы как средства описания смысловой стороны текстов, представляется предпочтительной.

Далее рассматриваются принципы построения языка-посредника в двух системах: ZARDOZ и Multi-path.

3. Система ZARDOZ. Система многоязыкового машинного перевода ZARDOZ задумывалась её разработчиками как инструмент перевода текстов со звучащего языка (условно говоря, английского) на некоторое множество жестовых языков, в частности, амслен, ирландский (ISL), японский (JSL). Эта система носит преимущественно экспериментальный характер, хотя, конечно, при её создании имелись в виду и прагматические цели, связанные с прикладными задачами из области искусственного интеллекта.

Разработка межмодального перевода, преобразующего звуковую (или графическую) модальность в пространственную, даёт возможность продвинуться в изучении языковых универсалий, когнитивных процессов и структур. Эта перспектива была существенна для создателей системы ZARDOZ. Особое внимание они уделяли вопросам, связанным с творческими аспектами языков, в частности, метафорам и созданию новых знаков.

В ZARDOZ используется семантический язык-посредник Doll Control Language (DCL).

При переходе от входного языка к DCL все метафоры, использованные во входном тексте, устраняются. На этом этапе язык DCL рассматривается авторами как универсальное представление смысла, не зависящее от языка (*language independent*). Однако затем к данному представлению применяется ряд преобразований, приближающих его к жестовому языку (тем самым, оно, строго говоря, перестаёт быть интерлингвой). Эти преобразования, названные «процессом схематизации», дают представление смысла, трактуемое далее как описание движений анимированного компьютерного изображения человека.

Как уже говорилось, разработчики системы ZARDOZ акцентируют своё внимание на одном из существенных для автоматического сурдоперевода аспектов: соотношении метафор звучащего и жестового языка. Метафорические и метонимические средства звучащих и жестовых языков, вообще говоря, не совпадают³. Следовательно, язык-посредник, будучи средством нейтрального описания смыслов, таких средств содержать не должен. Поэтому при переходе от входного (звучащего) языка к DCL все случаи использования метафор устраняются. Этот процесс опирается на имеющуюся в системе ZARDOZ базу соот-

³ В частности, форма иконического знака может препятствовать образованию метафор, вполне обычных для звучащего языка. Эта особенность жестовых языков детально рассмотрена в [14].

ветствий между лексемами (которые могут использоваться в метафорическом смысле) и нейтральными концептами.

На следующем этапе подключаются средства, предназначенные для разрешения анафорических ссылок, генерации синтаксических конструкций выходного языка, построения соответствий типа концепт-знак. На основе этих соответствий структурные составляющие интерлингвы преобразуются в знаки выходного языка. В результате могут снова появляться метафоры, но уже метафоры, свойственные выходному (жестовому) языку.

Для представления смысла в DCL используются фреймовые структуры. Проиллюстрируем процесс устранения метафор на примере перевода с английского языка предложения *I have a terrible headache* (*У меня ужасная головная боль*).

Форма выражения мысли в данном предложении метафорична. Эта метафоричность отразится в первоначально созданном фрейме **HAVE** (ИМЕТЬ), в котором позиции *POSSESSOR* (ОБЛАДАТЕЛЬ) и *POSSESSION* (ОБЪЕКТ ОБЛАДАНИЯ) будут заняты соответственно концептами *SPEAKER* (ГОВОРЯЩИЙ) и *HEADACHE* (ГОЛОВНАЯ БОЛЬ). Затем вместо этого будет построен другой фрейм, более адекватный данной ситуации и лишённый метафоричности: **SUFFER-FROM-AILMENT** (СТРАДАТЬ-ОТ-НЕДОМОГАНИЯ). Концепты *SPEAKER* и *HEADACHE* будут перемещены в более подходящие слоты *SUFFERER* (СТРАДАЮЩИЙ) и *AILMENT* (НЕДОМОГАНИЕ). Поскольку в ZARDOZ строится интерлингва, основанная на формализации знаний о мире, оказывается возможным заменить первый фрейм на второй. Элемент смысла, обозначающий интенсивность и выражаемый во входном тексте словом *terrible*, будет отображён в выходном языке как модификатор (*ASL-INTENSE*) жестового знака *HEADACHE* (ГОЛОВНАЯ БОЛЬ), а не как самостоятельный жестовый знак. Таким образом, при переходе от звучащего языка к жестовому языку метафора утрачивается.

3.1. Синтаксическая информация в составе языка-посредника.

Несмотря на то, что интерлингва системы ZARDOZ представляет собой описание семантики, она отчасти сохраняет и синтаксическую информацию. Дело в том, что одна и та же ситуация может быть описана разными текстами входного (звучащего) языка. Если считать, что смысл текста — это представление ситуации, описанной в тексте, то все тексты, эту ситуацию описывающие, должны иметь один и тот же смысл. Однако языковой способ описания ситуации тоже может быть важен при переводе. Например, в симметричном предикате типа *по-*

знакомиться, помириться, встретиться оба участника действия (занимающие позиции аргументов данного предиката) равноправны: фразы *Иван познакомился с Петром* и *Петр познакомился с Иваном* могут трактоваться как описывающие одну и ту же ситуацию знакомства. Однако в конкретном языковом предложении в роли активного субъекта выступает лишь один из названных людей, поэтому, вообще говоря, эти две фразы следует переводить по-разному. Интерлингва системы ZARDOZ сохраняет подобную синтаксическую информацию о роли аргументов предиката. Так, предложения *Why did Mary marry Bill?* (*Почему Мэри вышла замуж за Билла?*) и *Why did Bill marry Mary?* (*Почему Билл женился на Мэри?*) в интерлингве будут отображаться в виде одной и той же структуры, но в этой структуре будет указано, какой из участников выступает в роли субъекта [17].

3.2. Новые концепты и новые обозначения. Любая достаточно сложная система машинного перевода сталкивается с проблемой недостаточности словаря, используемого при переводе. При автоматизации сурдоперевода эта проблема также актуальна. В частности, для некоторых смыслов, выразимых посредством слов в звучащем языке, в жестовом языке лексические средства выражения могут отсутствовать. Для разрешения таких ситуаций в системе ZARDOZ предусмотрен специальный инструмент создания новых знаков (неологизмов).

Вообще говоря, при жестовой коммуникации всегда есть возможность отобразить новый смысл, для которого нет готового знака, посредством дактильной азбуки, побуквенно воспроизводящей слово звучащего языка. Однако помимо этого в ZARDOZ используются алгоритмы создания новых знаков, имитирующие естественное словотворчество носителей жестового языка. Рассмотрим несколько типов таких знаков.

Большое количество слов, передаваемых посредством дактильной азбуки, делает речь довольно неуклюжей и трудной для восприятия. Поэтому со временем такие знаки имеют тенденцию упрощаться, сохраняя лишь часть составляющих их букв (обычно из начала слова). Например, исполняемое побуквенно левой рукой слово J-A-P-A-N (*Япония*) уменьшается до демонстрации левой рукой одной лишь буквы J. Это самый простой способ образования нового знака в ZARDOZ.

Использование иерархии понятий — ещё один способ образования знаков, имитирующий интуицию носителей жестового языка. Например, если в жестовом языке нет обозначений для *аспирина* и *тайленола*, то они могут быть образованы посредством сочетаний знаков А + МЕДИЦИНА и Т + МЕДИЦИНА соответственно.

На основе иерархии понятий, используемой в интерлингве, смысл слова *сэндвич* отображается как [ЕДА В ХЛЕБЕ] ([FOOD INSIDE BREAD]). (Квадратные скобки указывают на то, что содержащиеся в них концепты составляют единый знак.)

Для отображения конкретного пространственного представления могут вводиться знаки, указывающие локализацию. Например, *головная боль* — это ЛОБ + БОЛЕТЬ, а *поясничная боль* — это ПОЯСНИЦА + БОЛЕТЬ.

Интересные варианты словотворчества — метонимия и пространственная метафора. Один из вариантов создания знака — подбор последовательности подходящих метонимий, имеющих жестовые обозначения. Например, обозначение смысла словосочетания *сэндвич из свинины* в амслене таково: ASL-PIG + ASL-MEAT + ASL-INSIDE + ASL-BREAD (то есть последовательность знаков амслена СВИНЬЯ + МЯСО + ВНУТРИ + ХЛЕБ).

Пространственные метафоры часто трактуются как проявления базовых концептуальных структур, свойственных человеческому мышлению. Эта мысль, связываемая прежде всего с именами Дж. Лаккоффа и М. Джонсона, очень популярна в современной когнитивной лингвистике. Суть её в следующем. Концептуализация (осмысление) человеком мира основывается на эмпирически доступной физической реальности, поэтому истоки всех, в том числе абстрактных, понятий лежат в сенсорном опыте и, следовательно, структурируются в тех же терминах, какие человек создаёт применительно к физической реальности. Основной вид сенсорного опыта — пространственный, поэтому пространственная метафора позволяет сводить многие абстрактные понятия к физическому опыту.

Как считают авторы [17], базовый характер сенсорного опыта свойствен не только звучащим, но и жестовым языкам, причём вторым — в большей степени, так как в них пространство является не только основой концептуализации мира, но и средой, в которой реализуется план выражения языка.

В системе ZARDOZ используется специальное инструментальное средство поддержки пространственных метафор (*scaffolding model of metaphor*), представляющее собой своеобразное пространственное исчисление, определённое на конструктивных метафорообразующих элементах: *Вверх, Вниз, Соединение, Разъединение (Up, Down, Contact, Discontact)*. Эти элементы используются как строительные блоки для конструирования смыслов. Таким образом может быть построена семантика многих языковых единиц (как абстрактных, так и конкрет-

ных), используемых в повседневной речи, например, *Зарплаты (цены на продукты, протестные настроения и т.д.) повышаются*. Конструктивный элемент *Соединение / Разъединение* усматривается в содержании текста, сообщающего о сближении / расхождении мыслей, а также в смыслах слов типа *знакомство, дружба, женитьба*. На основе указанного пространственного исчисления система ZARDOZ может динамически создавать новые, основанные на метафоре, знаки, не существовавшие ранее в жестовом языке. Как уже говорилось, между метафорами жестового языка и метафорами звучащего языка имеются различия, однако авторы системы полагают, что пространственные метафоры в большей степени универсальны, чем прочие виды метафор, поэтому использование их в качестве одного из средств порождения новых знаков жестового языка вполне возможно.

3.3. Сегментация жестового пространства как составляющая семантического описания. При использовании звучащего языка пространственные характеристики обсуждаемых объектов могут указываться более или менее детально, а могут и не указываться вообще. В силу специфики жестового языка его носители обладают гораздо более развитыми навыками пространственного видения и рассуждения. Этот факт должен учитываться при разработке системы автоматического сурдоперевода. В системе ZARDOZ жестовое пространство виртуального «говорящего»⁴ разделено на несколько областей. Самые значимые — *Левая область (Left-Space)*, *Правая область (Right-Space)*, *Срединная область (Middle-Space)*, располагающиеся вдоль стандартной линии обозначений жестового языка. Другие важные области включают *Верхне-левую (Upper-Left-Space)*, *Верхне-срединную (Upper-Middle-Space)*, *Верхне-правую (Upper-Right-Space)* области, лежащие на верхней линии обозначений, а также *Нижне-левую (Lower-Left-Space)* и *Нижне-правую (Lower-Right-Space)* области на нижней линии обозначений.

Локализация рук «говорящего» в одной из этих областей отчасти зависит от семантики текста: определённые синтактико-семантические роли локализуются в определённых областях жестового пространства. Например, в амслене на стандартной линии обычно располагаются позиции агенса и пациенса, на верхней / нижней линиях — явно ука-

⁴ Термин *говорящий* обозначает человека, использующего звучащий язык. Аналогичного термина, который обозначал бы человека, использующего жестовый язык, нет, поэтому будет использоваться тот же термин, но заключённый в кавычки.

занные ориентиры для локализации описанных в тексте объектов и действий. В системе ZARDOZ коррелированность семантики с сегментацией жестового пространства учитывается на том этапе, когда представление смысла на языке DCL трактуется как описание движений анимированного компьютерного изображения человека (то есть после применения к интерлингве преобразований, приближающих её к выходному языку).

В системе ZARDOZ необходимость интерлингвы обусловлена возможностью использования в переводимых текстах метафор: поскольку способы образования метафор в звучащем и жестовом языках не совпадают, необходимо нейтральное представление смысла, лишённое метафоричности. Другое языковое явление, обуславливающие необходимость интерлингвы, — это классификаторные предикаты, имеющиеся в жестовых языках. Этой проблеме автоматического сурдоперевода уделено существенное внимание при разработке системы Multi-path, основывающейся на пространственной интерлингве.

4. Система Multi-path. В системе Multi-path основное внимание уделяется классификаторным предикатам движения и локализации. Более сложные виды классификаторных предикатов данной версией системы не охвачены. Например, пока не рассматриваются «телесные» предикаты «body part classifier predicates» и «body classifier predicates». Применяя такие знаки, «говорящий» в качестве означающего использует своё собственное тело (а не только руки). В первом случае («body part classifier predicates») — одна часть тела (обычно верхняя) указывает на расположение или изменение другой части (обычно нижней). Во втором случае («body classifier predicates») человек описывает движения обсуждаемого персонажа посредством движений своего собственного тела.

Особую сложность при построении данной системы представляют случаи, когда разные типы классификаторных предикатов следуют непосредственно друг за другом, или когда приходится переключаться с одного способа видения ситуации на другой.

Классификаторный предикат характеризуется двумя составляющими: первая задаёт форму и ориентацию руки, вторая — траекторию движения в 3-мерном пространстве. В речи такому предикату обычно предшествует номинативная группа, указывающая, какой объект будет далее описан посредством классификаторного предиката. Классификаторный предикат позволяет локализовать точки референции обсуждаемых в тексте объектов, показать форму, ориентацию, движение этих

объектов. «Говорящий» как бы рисует в пространстве воображаемую ситуацию и показывает в ней отдельные элементы.

Количество классификаторных предикатов потенциально бесконечно. Поэтому коллекция анимированных изображений должна включать конструктивные компоненты, достаточные для спецификации всевозможных позиций в пространстве, ориентаций предметов относительно друг друга (*ниже, выше, слева, справа, над, под, рядом* и т.д.), траекторий движения.

4.1. Виртуальная реальность как интерлингва. Идея использования виртуальной реальности в качестве интерлингвы состоит в следующем [13]. Входной текст (например, на английском языке) трактуется как естественно-языковая инструкция (Natural Language Instructions, или NLI), служащая для управления динамически меняющимся поведением определённой системы физических объектов, представляемой в виде трёхмерной модели. Интерлингва в этом случае соответствует этапу визуализации динамической сцены, описанной входным текстом и включающей все упомянутые в нём объекты.

Такой подход к построению интерлингвы достаточно необычен для систем машинного перевода, но именно он позволяет описать те компоненты смысла входного текста, которые значимы для перевода классификаторных предикатов.

Задача построения универсального (для определённого множества языков) средства представления смыслов чрезвычайно сложна. В реально действующих системах автоматического перевода для звучащих языков интерлингва используется тогда, когда тематика и грамматическое разнообразие переводимых текстов сильно ограничены (например, сводки погоды). Однако автор системы Multi-path [11, 12] сделал одно полезное наблюдение, показав, что при переводе на жестовый язык интерлингву можно ограничивать иным образом, сужая не тематическое и грамматическое разнообразие текстов, а область применения интерлингвы. В Multi-path в эту область попадают прежде всего классификаторные предикаты.

Классификаторные предикаты выражают значения определённого, **ограниченного**, типа: они обозначают движения, местоположения, ориентации, физические характеристики сущностей в трёхмерном пространстве. Если сузить область действия интерлингвы, используя её преимущественно для перевода текстовых фрагментов, выражающих подобные смыслы, то отпадает необходимость предусматривать в языке-посреднике средства выражения для многих гораздо более

сложных смыслов (абстрактных понятий, верований, интенций, квантификаций и т.д.).

В Multi-path предлагается система перевода с многоходовой архитектурой, где траектория перевода определяется составом элементов переводимого текста и может использовать три стратегии: перевод с использованием интерлингвы, перевод типа «трансфер», прямой перевод. Проблема выбора стратегии перевода при этом не возникает, поскольку в любом случае система сначала пытается перевести текст на основе наиболее глубокого анализа (то есть через интерлингву) и лишь в случае неуспеха переходит к более простым стратегиям.

Схема работы пространственной интерлингвы, используемой в многоходовой системе Multi-path, такова: программы визуализации преобразуют текст в трёхмерное изображение, затем это изображение накладывается на область жестового пространства, после чего переводится на жестовый язык.

4.2. Язык описания виртуальной реальности. Система Multi-path использует библиотеку иерархически упорядоченных параметризованных представлений действий (Parameterized Action Representations, или PARs), производимых в виртуальном пространстве [9]. Эта библиотека задаёт **анимированные примитивы**, то есть простые действия и под-действия, из которых составляются более сложные действия. Библиотека PAR используется также в указанной ранее системе [13], ориентированной на управление динамическими анимированными объектами посредством инструкций на естественном языке. Первоначально библиотека PAR создавалась в рамках проекта AnimNL (animation from natural language), реализованного в военной сфере (в частности, был создан тренажёр, предназначенный для тренировки действий военнослужащих в ситуациях разрешения конфликтов и выполнения миротворческих функций). Это направление использования PAR описано, например, в [10].

В системе Multi-path библиотека действий PAR была расширена таким образом, чтобы охватить как можно более широкий спектр пространственных действий, описываемых текстами английского языка.

Этап преобразования входной текстовой информации в виртуально-пространственную форму осуществляется посредством специальной программы визуализации (AnimNL scene visualization software), на вход которой поступает инструкция на английском языке, предписывающая выполнить некое действие с объектами в 3-мерном пространстве. Реакция системы — выполнение полученной инструкции.

В AnimNL действиями всех объектов управляет специальный планирующий процесс, благодаря которому достигается естественность движений. Кроме этого, Multi-path содержит подборку образцов «исполнения» для различных типов классификаторных предикатов (движения гуляющего человека, автомобиля при парковке и т.д.). Обычно такой образец определяет не только траекторию движения руки, но и её форму.

Каждое действие описывается шаблоном с набором слотов, задающих действующий объект, траекторию, способ, скорость перемещения, конечное / начальное положение, результат, необходимые условия совершения и т.д. Библиотека PAR организована иерархически, аргументы и свойства наследуются при переходе от более общих действий к более конкретным. На верхних уровнях иерархии находятся обобщённые PAR-схемы, каждая из которых задаёт базовую аргументную структуру для всей группы подчинённых вершин, соответствующих различным вариантам обобщённого действия. Семантические свойства глаголов движения выявляются на основе иерархии английских глаголов VerbNet (<http://verbs.colorado.edu/verb-index/index.php>). Ниже кратко описаны параметры действий в PAR [10].

Объекты в PAR характеризуются своим типом и упорядочены иерархически. Тип объекта определяется перечнем действий, осуществляемых с объектом (схватить, достичь, передвинуть и т.д.), и изменений, возникающих в результате осуществления этих действий.

Агент есть особого типа объект: объект, выполняющий действия. Свойства агентов хранятся в иерархической базе объектов.

Условия применимости действия — это булево выражение, определяющее, какие предикаты должны быть истинны для того, чтобы выполнить действие. Условия могут включать способности агента, конфигурацию объекта и прочие свойства окружающей среды. Например, для действия *идти* одним из условий может быть наличие у агента способности ходить.

Подготовительные спецификации — это список предусловий, необходимых для осуществления действия. Система сначала проверяет, выполнены ли предусловия, и, если они не выполнены, изменяет положение дел таким образом, чтобы предусловия выполнялись, и лишь после этого осуществляет само действие. Например, *стоять* — это одно из предусловий действия *идти*, прежде чем выполнить действие *идти*, агент, если он ещё не находится в положении, должен встать.

Стадии выполнения характеризуют детали выполнения действия, описывая сложное действие в терминах составляющих его простых под-действий. Порядок выполнения под-действий может быть параллельным или последовательным.

Способ выполнения действия агентом. Например, инструкция *Подойди к двери и медленно поверни ручку* задаёт способ осуществления действия (*медленно*).

Условия завершения — это список условий, выполнение которых свидетельствует о том, что действие осуществлено.

Завершение — список утверждений, которые должны быть реализованы после того, как система установит, что условия завершения действия выполнены. Данные утверждения определяют, какие изменения среды должны произойти в результате осуществления действия (причиной этих изменений может быть как само действие, так и его побочные эффекты).

Часть нужной для доопределения этих параметров информации извлекается из входного английского текста. Таким образом могут быть определены, например, начальный и конечный пункты движения, выбраны местоположения объектов (на основе анализа наречий и предлогов) и т.д.

Следующий иллюстративный пример [13] поясняет способ представления действий. Обобщённое действие *контакт* (*contact*) предполагает наличие одного действующего агента и двух объектов. Действие считается законченным тогда, когда два объекта входят в контакт. Предварительное условие действия: агент должен обладать объектом. Существует много более конкретных действий, представляющих собой подвиды действия *контакт*. Одно из них — действие *hit* ('ударить'). Обобщённое описание действия *hit* не определяет, какой объект должен выступать в роли инструмента действия. Эта информация содержится в описаниях ещё более конкретных действий, например, *kick* ('пнуть', 'лягнуть') и *hammer* ('прибивать'), являющихся подвидами действия *hit*. В первом случае инструмент — нога, во втором — молоток. Вместе с тем, характеристика способа совершения действия (*manner: forcefully*, 'способ: сильно') не указывается в описаниях действий *kick* и *hammer*, а наследуется от более общего глагола *hit*. Действие *touch* ('прикасаться') также относится к типу *contact*, однако имеет другую характеристику: *manner: gently* ('способ: осторожно, бережно').

4.3. Невидимое изображение ситуации как составляющая семантики. Один из важных аспектов понимания текста — установле-

ние референции, то есть тех объектов, которые являются составляющими обсуждаемой ситуации. При использовании жестового языка позиции референции локализуются в жестовом пространстве [7]. В системе Multi-path такая возможность обеспечивается двумя визуальными представлениями.

Первое (visualization scene) визуализирует общий вид ситуации: определяет пространственное расположение объектов и их ориентацию. Здесь же эти объекты связываются с теми языковыми обозначениями, которые использовались для указания на них.

Второе (depicting canvas) рисует картину, задающую определённую точку зрения на ту же ситуацию (позицию, с которой её видит в жестовом пространстве «говорящий»). Эта картина невидима: на ней обозначены лишь местоположения (placeholders) объектов, составляющих обсуждаемую ситуацию. Сами эти объекты — точнее, их «призраки» (ghosts), — присутствуют незримо. Позиция каждого «призрака» задаёт место референции соответствующего объекта. Характеристика «призрака» содержит информацию о соответствующем ему объекте и включает:

- место расположения (центр тяжести) объекта в рассматриваемой 3-мерной ситуации;
- ориентацию объекта;
- указание на его языковые обозначения (установленные в ходе разрешения референции);
- бинарные признаки объекта (является ли он человеком, животным, транспортным средством и т.д.), предопределяющие форму руки, которая используется для референции к этому объекту в составе классификаторного предиката.

Переход от первого визуального представления ко второму — это переход от нейтрального видения ситуации к видению, приближающемуся к жестовому языку. Поэтому языком-посредником в Multi-path считается первое визуальное представление. Как полагает автор данной системы, эта трёхмерная ситуация не вполне соответствует понятию интерлингвы, используемому в литературе по машинному переводу. Одно из общепринятых требований к интерлингве — нейтральность задаваемого ею семантического описания, отсутствие привязки к какому-либо конкретному языку. Визуализированная ситуация не является в полной мере «языко-независимой» (language-neutral), поскольку содержит некоторую информацию, специфичную для английского языка и необходимую для представления классификаторных

предикатов амслена. Поэтому визуализированная ситуация характеризуется как «почти-интерлингва» (“near-interlingua”).

5. Заключение. Опыт использования интерлингвы в системах автоматического сурдоперевода можно резюмировать следующим образом.

Интерлингва — это единый метаязык описания семантик двух языков: входного звучащего и выходного жестового. Основная проблема построения такого метаязыка состоит в необходимости уточнения и обогащения семантики звучащего языка.

Входной текст, по сути, задаёт инструкцию, определяющую поведение динамического изображения человека. Эта инструкция в ряде отношений является семантически недоопределённой.

Недоопределённость обусловлена, в частности, тем, что многие из пространственных параметров (способ движения, локализация точек референции и т.д.) для звучащей речи несущественны и могут мыслиться самым общим образом, но в жестовом языке должны быть конкретизированы и «показаны» посредством жестов. Соответственно при построении интерлингвы семантика звучащего языка должна быть перестроена так, чтобы учесть все параметры, необходимые для пространственного моделирования ситуаций.

Недоопределённость звучащего языка проявляется также в отсутствии необходимости отражать все пространственные перемещения обсуждаемых объектов. Эти объекты обладают относительной автономией: они могут менять своё расположение и форму, даже если во входном тексте об этом не сказано, то есть просто благодаря течению времени. Например, идущий человек по прошествии времени оказывается в другой точке жестового пространства. При переходе к интерлингве семантика звучащего языка должна быть дополнена средствами моделирования (автономных, не обозначенных входным текстом) перемещений и изменений формы объектов в трёхмерном пространстве.

Формальным средством воплощения таких требований может служить пространственная интерлингва, позволяющая задавать трёхмерные динамические модели обсуждаемых ситуаций. Детальное обсуждение возможного устройства и функционирования такой интерлингвы, а также оценки качества перевода содержатся в диссертационном исследовании [12], полный текст которого доступен по адресу <http://eniac.cs.qc.edu/matt/research.html>.

При построении интерлингвы необходимо не только доопределить семантику звучащего языка недостающими параметрами, но и произвести ревизию метафорических выражений, так как возможности

построения метафор в звучащих и жестовых языках, вообще говоря, различны. Интерлингва служит средством нейтрального описания ситуации и может содержать лишь базовые виды метафор (предположительно таковы пространственные метафоры).

Вероятность появления новых для системы автоматического перевода слов неустраима. Система сурдоперевода, включающая в себя модель словотворчества и поиск синонимов для новых (отсутствующих в системе) слов, делает жестовую речь более естественной.

Литература

1. *Воскресенский А.Л., Ильин С.Н., Zelezny M.* О распознавании жестов языка глухих // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 26–30 мая 2010 г.). Вып. 9 (16). М.: РГГУ, 2010. С. 76–81.
2. *Гриф М.Г., Королькова О.О.* Разработка компьютерного сурдопереводчика звучащей речи на разговорный русский жестовый язык /Материалы XI межд. науч.-метод. конф. Информатика: проблемы, методология, технологии 10–11 февраля 2011. Т. 1. Воронеж: ВГУ. 2011. С.206–208.
3. *Гриф М.Г., Демьяненко Е.А., Королькова О.О.* Разработка технологий компьютерного сурдоперевода непрерывной русской речи на разговорный русский жестовый язык // Сборник науч. труд. Российской науч.-практ. конф. Автоматизированные системы и информационные технологии 22–23 сентября 2011. Новосибирск: НГТУ. 2011. С.59–68.
4. *Гриф М.Г., Тимофеева М.К.* Проблема автоматизации сурдоперевода с позиции прикладной лингвистики // Сибирский филологический журнал. 2012. № 1. С. 211-219.
5. *Зайцева Г.Л.* Жестовая речь. Дактилология: Учебник для студентов. М.: ВЛАДОС, 2000. 192 с.
6. *Карпов А.А., Касиров И.А.* Формализация лексикона системы компьютерного синтеза языка жестов. СПб: Труды СПИИРАН. 2011. Вып. 1(16). С. 123–140.
7. *Кибрик А.А., Прозорова Е.В.* Референциальный выбор в русском жестовом языке // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог 2007» (Бекасово, 30 мая – 3 июня 2007 г.). М.: РГГУ, 2007. С. 220–230.
8. *Прозорова Е.В.* Российский жестовый язык как предмет лингвистического исследования // Вопросы языкознания, 2007. № 1. С. 44–61.
9. *Badler N.I., Bindiganavale R., Allbeck J., Schuler W., Zhao L., Lee S.-J., Shin H., Palmer M.* Parameterized Action Representation and Natural Language Instructions for Dynamic Behavior Modification of Embodied Agents. 2000. AAAI Technical Report SS-00-02.
10. *Bindiganavale R., Schuler W., Allbeck J.M., Badler N.I., Joshi A.K., Palmer M.* Dynamically Altering Agent Behaviors Using Natural Language Instructions // Proceedings of the fourth international conference on Autonomous agents. ACM New York, NY, USA 2000. P. 293–300.
11. *Huenerfauth M.* A Survey and Critique of American Sign Language Natural Language Generation and Machine Translation Systems. Technical Report MS-CIS-03-32, Computer and Information Science, University of Pennsylvania. 2003. 36 p.
12. *Huenerfauth M.* Generating American sign language classifier predicates for English-to-ASL machine translation. Doctoral Dissertation in Computer and Information Science. University of Pennsylvania Philadelphia, PA, USA, 2006. 312 p.

13. *Kipper K., Palmer M.* Representation of Actions as an Interlingua // NAACL-ANLP-Interlinguas '00 Proceedings of the 2000 NAACL-ANLP Workshop on Applied interlinguas: practical applications of interlingual approaches to NLP, 2000. Vol. 2. P. 12–17.
14. *Meir I.* Iconicity and metaphor: Constraints on metaphorical extension of iconic forms // *Language*. Vol. 86. № 4. 2010. P. 865–896.
15. *Morrissey S.* Data-Driven Machine Translation for Sign Languages. PhD Thesis, Dublin City University, Dublin, Ireland, 2008. 192 p.
16. *Veale T., Collins B.* Space, metaphor and schematization in sign: Sign Language translation in the ZARDOZ system // Proceedings of the Second Conference of the Association for Machine Translation in the Americas (AMTA'96), Montreal, Canada, 1996.
17. *Veale T., Collins B., Conway A.* The Challenges of Cross-Modal Translation: English-to-Sign-Language Translation in the Zardoz System // *Machine Translation*, N 13. 1998. P. 81–106.
18. *Veale T., Conway A.* Cross modal comprehension in ZARDOZ an English to sign-language translation system // INLG '94 Proceedings of the Seventh International Workshop on Natural Language Generation, Kennebunkport, Maine, USA, June 21–24, 1994.
19. *Wu, C.-H., Su, H.-Y., Chiu, Y.-H., and Lin, C.-H.* Transfer-Based Statistical Translation of Taiwanese Sign Language Using PCFG // *ACM Transactions on Asian Language Information Processing (TALIP)*, 2007. Vol.6. No.1. P. 1–18.
20. *Zhao L., Kipper K., Schuler W., Vogler C., Badler N.I., Palmer M.* A Machine Translation System from English to American Sign Language // Proceedings of the 4th Conference of the Association for Machine Translation in the Americas on Envisioning Machine Translation in the Information Future (AMTA 00), 2000. Springer-Verlag. P. 54–67.

Гриф Михаил Геннадьевич — д.т.н.; проф., профессор кафедры автоматизированных систем управления Новосибирского государственного технического университета (НГТУ). Область научных интересов: методы проектирования и оптимизации человеко-машинных систем, системы компьютерного сурдоперевода. Число научных публикаций — 205. grifmg@mail.ru; пр. К.Маркса, 20, Новосибирск, 630092; п.т. +7(383) 346-15-59, факс +7(383)346-11-00.

Grif Mikhail Gennadyevich — Doctor of Science (Engineering); professor, Department of Automated Control Systems, Novosibirsk State Technical University. Research interests: methods of designing and optimization of man-machine systems, systems of sign language machine translation. Number of publications — 205. grifmg@mail.ru; Karl Marx Avenue, 20, Novosibirsk; p.t. +7(383) 346-15-59, факс +7(383)346-11-00.

Тимофеева Мария Кирилловна — д.фил.н.; с.н.с. лаборатории логических систем Института математики им. С.Л. Соболева СО РАН, профессор гуманитарного факультета Новосибирского государственного университета. Область научных интересов: логические проблемы семантики и прагматики естественного языка, функциональные математические модели естественного языка, философско-методологические основания лингвистики. Число научных публикаций — 79. timof@math.nsc.ru; пр. ак. Коптюга, 4, 630090, г. Новосибирск, Россия; +7(383)3634538.

Timofeeva Maria Kirillovna — Doctor of Philology; senior scientific researcher, Laboratory of Logical Systems, Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences; professor of the Faculty of Humanities, Novosibirsk State University. Research interests: logical problems of semantics and pragmatics of natural language, functional mathematical models of natural language, philosophical and methodological basis of linguistics.

tics. Number of publications — 79. timof@math.nsc.ru; 4 Acad. Koptyug avenue, 630090 Novosibirsk Russia; +7(383)3634538.

Рекомендовано лабораторией речевых и многомодальных интерфейсов, заведующий лабораторией Ронжин А.Л., д-р техн. наук, доц.
Статья поступила в редакцию 26.03.2012.

РЕФЕРАТ

Гриф М.Г., Тимофеева М.К. **Интерлингва в системах машинного перевода для жестовых языков.**

Жестовые языки предназначены для коммуникации глухих и слабослышащих людей. Такие языки возникли значительно позже звучащих языков и ещё позже стали предметом научного изучения. Системы автоматического перевода со звучащих языков на жестовые языки начали разрабатывать около двух десятилетий тому назад, для русского жестового языка пока таких систем нет. Цель статьи: проанализировать имеющийся опыт построения систем автоматического сурдоперевода, рассмотрев предложенные в них стратегии решения ряда семантических проблем, являющихся общими для жестовых языков и, тем самым, возникающих также и в связи с русским жестовым языком.

Жестовые языки являются самостоятельными знаковыми системами, отличающимися от звучащих языков как по грамматическому, так и по семантическому устройству. Это делает невозможным прямое использование многих алгоритмов и программ, разработанных для компьютерной обработки текстов, в частности, для машинного перевода. Поэтому, хотя создатели систем автоматического сурдоперевода используют те же базовые типы стратегий, что и для звучащих языков (статистический перевод, прямой перевод, трансфер, перевод на основе интерлингвы), реализация этих стратегий для жестовых языков представляет собой весьма нетривиальную задачу.

Сложность разработки систем автоматического сурдоперевода связана прежде всего с преобразованием входного текста из одной модальности (звуковой или графической) в другую (пространственную). Ряд факторов обуславливает необходимость промежуточного уровня представления смысла, нейтрального по отношению к указанным модальностям и реализуемого в стратегии перевода, использующей интерлингву.

Знаки, реализуемые в жестовом пространстве, обладают значительно большей вариативностью по сравнению со знаками звучащего языка. В этом плане наибольшие проблемы для перевода создают классификаторные предикаты, при изображении которых в жестовом пространстве фактически моделируется (посредством движений рук) динамическая ситуация, описанная во входном (переводимом) тексте. Теоретически возможное количество вариаций у таких знаков неограниченно, один из способов их задания — построение исчисления, определяющего конечный набор базовых примитивов и правила комбинирования примитивов для построения описаний сложных действий.

Другой ряд проблем обусловлен различиями в закономерностях образования метафор в жестовых и звучащих языках, а также несовпадением множеств лексических значений, выразимых в этих языках.

Опыт решения указанных общих семантических проблем в существующих системах сурдоперевода, основанных на использовании интерлингвы, может быть полезен при разработке систем автоматического сурдоперевода для русского жестового языка.

SUMMARY

Grif M.G., Timofeeva M.K. **Interlingua in machine translation systems for sign languages.**

Sign languages are languages for deaf people or people with hearing problems. These languages were formed later than sounding languages, and have become the objects of scientific research only recently. First systems of machine translation to sign languages have appeared about twenty years ago; and nowadays, there is still no such system for Russian Sign Language. The purpose of this article is to generalize recent experience in constructing machine translation systems for sign languages; and to consider strategies of solving several common semantic problems intrinsic for all sign languages and therefore typical for Russian Sign Language.

Sign languages are languages in their own right; they differ from sounding languages both in grammar and in semantics. This fact makes it impossible to use directly all the algorithms and computer programs that were worked out for sounding languages, particularly for machine translation systems. Hence, although the developers of machine translation systems for sign languages use the same basic types of strategies as for sounding languages (statistic translation, direct translation, transfer, interlingual translation), the realization of these strategies for signed languages is a rather nontrivial task.

Complexities, which arise in connection with the problem of designing machine translation system for sign language, result *prima facie* from the necessity of converting one modality (sounding or graphical) into another (spatial). Several factors stipulate the inevitability of intermediate level of semantic representation, which is neutral in relation to the above mentioned modalities and realizes interlingual machine translation strategy.

Signs that are realized within sign space are much more variable than signs of sounding language. In this regard, the most complex translation problems arise from classificatory predicates. Representation of such predicate in fact consists in simulating (by hands motions) the dynamic picture described by an input text. From theoretical point of view, such signs have potentially unlimited number of variations, and one of the ways of defining them is to build a formal calculus specifying a finite number of basic primitives (elemental actions and parts of actions) together with a finite number of rules for combining these primitives into definitions of more complex actions.

One more group of problems arises from differences between mechanisms of metaphor formation in sign languages, on the one hand, and in sounding languages, on the other; besides, the sets of lexical senses expressible in these two classes of languages are not equal.

Experience accumulated during the previous history of solving these general semantic problems in interlingual machine translation systems for sign languages may be useful for creating machine translation systems for Russian Sign Language.