

# Распознавание жестовой речи с учетом комбинаторных изменений жестов

М. Г. Гриф, email: grifmg@mail.ru <sup>1</sup>

О. О. Королькова, email: ookorol@mail.ru <sup>2</sup>

А. Л. Приходько, email: alexeyayu@yandex.ru <sup>1</sup>

<sup>1</sup> Новосибирский государственный технический университет

<sup>2</sup> Новосибирский государственный педагогический университет

**Аннотация.** *Рассмотрена текущая ситуация в России и в мире в области разработки систем компьютерного перевода жестовых языков глухих. Предложены два подхода к распознаванию непрерывной жестовой речи, связанных с последовательным перебором событий-кандидатов на начало исполнения последующего жеста и с полным выявлением элементов речи – собственно жестов, состояния покоя говорящего, комбинаторных изменений параметров жестов и эпентезы.*

**Ключевые слова:** *русский жестовый язык, грамматическая система, машинный перевод, комбинаторные изменения параметров жестов, эпентеза.*

## Введение

Разработка методов двустороннего машинного перевода (русского) текста (речи) на жестовые языки глухих (русский жестовый язык, РЖЯ) в любой стране мира является важной социальной задачей, способствующей поддержке коммуникаций между глухими и слышащими. Актуальность разработки систем машинного перевода жестовых языков (от слышащего к глухому и наоборот) состоит как в недостаточном количестве переводчиков жестовых языков, так и не всегда желательном посредничестве (медицина, личностные отношения и т.п.) при коммуникациях глухих и слышащих граждан [1, 2].

В настоящее время данная задача не может считаться решенной, так как отсутствуют известные варианты машинного перевода русского текста на РЖЯ и наоборот, удовлетворяющие требованиям глухих граждан России. К основным требованиям глухих относятся безошибочность и качество визуализации перевода (компьютерного персонажа). Проведенные исследования среди глухих позволяют установить требования к безошибочности перевода не ниже 90% и к качеству визуализации на уровне перевода человека-сурдопереводчика

РЖЯ. Указанные обстоятельства затрудняют коммуникации между слышащими и глухими гражданами, что является серьезной социально-экономической проблемой. В настоящее время при переводе русского текста на РЖЯ и наоборот, используют как методы машинного обучения, так и связь грамматических систем русского звучащего языка и РЖЯ. Однако достичь необходимого уровня перевода препятствуют следующие обстоятельства:

неполнота описания грамматической системы РЖЯ;  
отсутствие «плавности» показа жестов аватаром-сурдопереводчиком, система управления которого использует систему нотаций языков глухих;

перевод осуществляется преимущественно на калькирующую жестовую речь, а не на РЖЯ, обладающий своими выразительными возможностями;

высокий процент ошибок при переводе многозначных слов и омонимов на жесты РЖЯ (более 20%);

отсутствие значительных по объему размеченных корпусов РЖЯ (Датасетов), необходимых для реализации методов машинного перевода на основе машинного обучения;

отсутствие надежных методов распознавания, как отдельных жестов глухих, так и РЖЯ в целом.

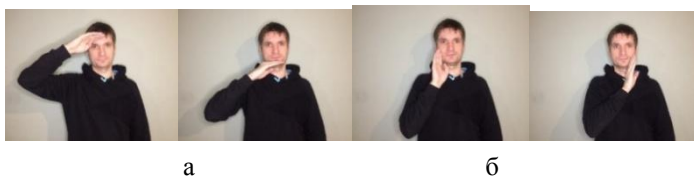
Особую сложность представляет собой задача распознавания непрерывной жестовой речи (РЖЯ). Для успешного распознавания недостаточно выделить отдельные жесты. Необходимо их уверенное выделение с учетом комбинаторных изменений параметров жестов, включая эпентезу [3]. В настоящей работе остановимся на данной задаче более подробно.

### **1. Виды комбинаторных изменений параметров жестов**

Анализ непрерывной речи на русском жестовом языке позволил выявить, что в речевом потоке на невербальных языках встречаются процессы комбинаторных изменений параметров жестов, сопоставимые с комбинаторными изменениями звуков в речевом потоке на вербальном языке. Так, под влиянием звуков друг на друга в звуковых языках могут наблюдаться процессы аккомодации (частичного приспособления артикуляций смежных звуков), ассимиляции (уподобления артикуляции звуков), гаплоггии (выпадения звуков вследствие их диссимиляции), диссимиляции (расподобления артикуляции звуков, утраты ими общих фонетических признаков), диэрезы (диэрезы) (выпадения звуков), метатезы (взаимной перестановки звуков или слогов) и эпентезы (вставки звуков) [4].

Согласно исследованиям русского жестового языка [5], для непрерывной русской жестовой речи характерны следующие комбинаторные изменения параметров жеста.

Выпадение движения происходит при использовании сочетания жестов или повторяющегося движения. Например, сравним исполнение сложного жеста (компаунда) ПАПА-МАМА ‘родители’ в словарной форме и в быстрой речи. В словарной форме эти жесты исполняются так, как показано на рис.1.



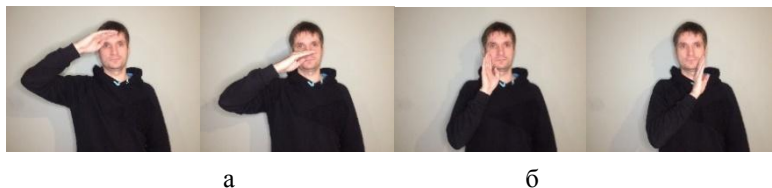
*а – компонент ПАПА, б – компонент МАМА*

*Рис. 1. Исполнение компаунда ПАПА-МАМА в словарной форме*

В быстрой речи эти же жесты имеют несколько иное исполнение (рис. 2).

Сопоставительный анализ изображений на рис. 1 и рис. 2 позволяет выявить выпадение движения при исполнении компонента ПАПА.

Для высказываний на РЖЯ характерны ассимиляция конфигурации и ассимиляция локализации, при которых конфигурация и локализация становятся похожими на соответствующие параметры одного из соседних жестов. Приведем пример ассимиляции конфигурации – исполнение компаунда БРАТ-СЕСТРА в словарной форме (рис. 3) и в быстрой речи (рис. 4).

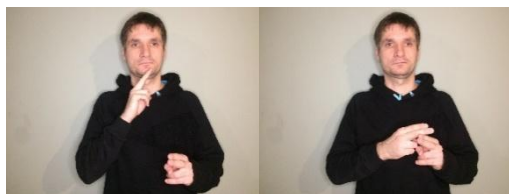


*а – компонент ПАПА, б – компонент МАМА*

*Рис. 2. Исполнение компаунда ПАПА-МАМА в быстрой речи*

Высказываниям на РЖЯ свойственна perseverация, т.е. удержание по инерции на месте пассивной руки вместо возвращения ее в

положение покоя, которая наблюдается при исполнении одnorучного жеста после показа двуручного жеста. Примером этого вида комбинаторных изменений может служить исполнение жеста БРИГАДИР (рис. 5).

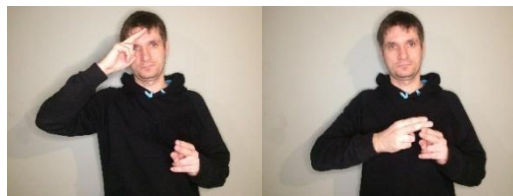


а

б

*а – компонент БРАТ, б – компонент СЕСТРА*

*Рис. 3. Исполнение компаунда БРАТ-СЕСТРА в словарной форме*

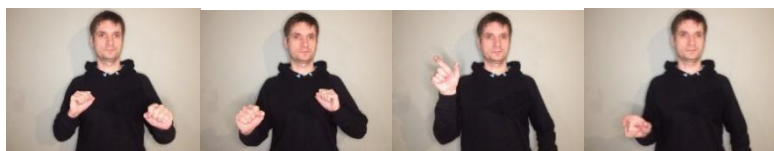


а

б

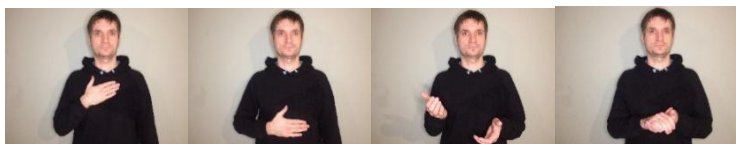
*а – компонент БРАТ, б – компонент СЕСТРА*

*Рис. 4. Исполнение компаунда БРАТ-СЕСТРА в быстрой речи*



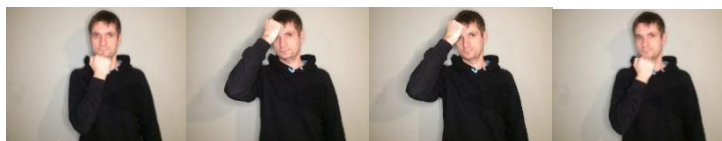
*Рис. 5. Жест БРИГАДИР*

Антиципация, т.е. заблаговременная подготовка пассивной руки для исполнения двуручного жеста, следующего за одnorучным жестом, может быть проиллюстрирована исполнением на РЖЯ высказывания «Приятно познакомиться» (рис. 6).



*Рис. 6. ПРИЯТНО ПОЗНАКОМИТЬСЯ*

Для РЖЯ характерна метатеза, т.е. перестановка начальной и конечной локализаций при исполнении жестов с двойной локализацией. Например, может меняться начальная и конечная локализация при исполнении жеста СПАСИБО (рис. 7).



*а*

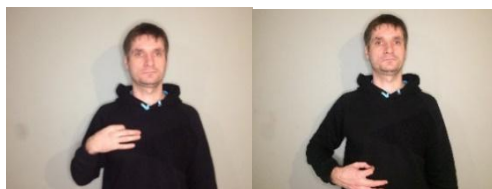
*б*

*а – жест СПАСИБО, б – жест СПАСИБО с измененной начальной и конечной локализацией*

*Рис. 7. Исполнение жеста СПАСИБО*

Высказываниям на РЖЯ свойственна эпентеза, т.е. добавление движения между соседними жестами. Этот вид комбинаторных изменений труден для распознавания.

Кроме того, исследователи РЖЯ относят к комбинаторным изменениям параметров жеста редукцию движения, под которой понимают исполнение движения с меньшей траекторией [5]. Пример исполнения жеста СЫН в словарной форме представлен на рис. 8, а исполнение этого же жеста с редукцией движения – на рис. 9:



*Рис. 8. Исполнение жеста СЫН в словарной форме*

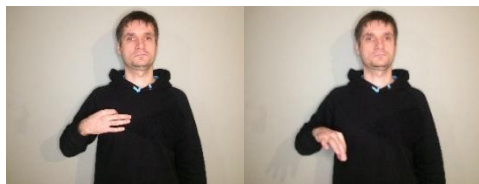


Рис. 9. Исполнение жеста СЫН в быстрой речи

Анализ рис. 8 и рис. 9 позволяет сделать вывод о том, что в быстрой речи произошла редукция движения.

Учет перечисленных комбинаторных изменений жестов важен для того, чтобы в процессе перевода непрерывной русской жестовой речи на русский вербальный язык не принять жесты с изменениями за новые языковые единицы, чтобы сделать перевод более точным и адекватным, разработать модель обучения нейронных систем переводу с РЖЯ на русский звучащий язык.

## **2. Подходы к распознаванию жестовой речи с учетом комбинаторных изменений параметров жестов**

В [3] приведен обзор работ, связанных с выделением эпентезы в мировых жестовых языках. Основная идея заключается в выделении признаков начала следующего жеста или паузы в речи. Это могут быть состояние покоя говорящего, положения рук, взгляда и т.п. [6]. Применяют также приемы вычленения «быстрых» и «медленных» движений в видео потоке. Если процесс жестовой речи свести к анализу последовательности выявленных компонент (конфигурация-ориентация ладони, локализация, траектории рук и характер движения, немануальные компоненты), то каждое такое событие является «кандидатом» на начало распознавания следующего жеста. При этом может оказаться, что процесс нового поиска окажется запущенным, пока не окончен предыдущий. Подобный подход к распознаванию жестовой речи является достаточно трудоемким. Представляется, что его можно модернизировать с введением функции принадлежности (или некоторой вероятности) каждого события к началу нового жеста. В этом случае, алгоритм распознавания нового жеста можно запускать только после некоторых, наиболее вероятных событий. Другой подход состоит в выделении всех элементов жестовой речи – собственно жестов, состояния покоя говорящего, комбинаторных изменений жестов и эпентезы. В этом случае необходимо расширить Датасет для распознавания жестов и другими элементами, связанными с комбинаторными изменениями жестов. На данный момент сложно оценить необходимый объем дополнений. С учетом большого числа

типов комбинаторных изменений жестов, он может оказаться значительным.

### **Заключение**

Сформулированы основные проблемы в области систем компьютерного сурдоперевода русского жестового языка. Приведены примеры исполнения языковых единиц, иллюстрирующих такие типы комбинаторных изменений параметров жестов, как выпадение движения, ассимиляция конфигурации, персеверация, антиципация и метатеза. Предложены два подхода к распознаванию непрерывной жестовой речи, связанных с последовательным перебором событий-кандидатов на начало исполнения последующего жеста и с полным выявлением элементов речи – собственно жестов, состояния покоя говорящего, комбинаторных изменений параметров жестов и эпентезы.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (Грант РФФИ ИНД\_а № 19-57-45006).

### **Список литературы**

1. Гриф, М. Г. Лексические и грамматические аспекты разработки компьютерного сурдопереводчика русского языка : монография / М. Г. Гриф [и др.]. – Новосибирск: Изд-во НГТУ, 2013. – 292 с.
2. Мануева, Ю. С. Построение системы компьютерного сурдоперевода русского языка / Ю. С. Мануева, М. Г. Гриф, А. Н. Козлов // Труды СПИИРАН. 2014. – Спб., 2014. – № 6(37). – С. 170-187.
3. Elakkiya, R. Recognition of Russian and Indian sign languages used by the deaf people / R. Elakkiya [и др.] // Научный вестник НГТУ. – Новосибирск: Изд-во НГТУ, 2020. – № 2–3 (79). – С. 57–76.
4. Лингвистический энциклопедический словарь / Гл. ред. В. Н. Ярцева. – М.: Советская энциклопедия, 1990. – 683 с.
5. Введение в лингвистику жестовых языков. Русский жестовый язык: учебник / Коллектив авторов: ред. С. И. Буркова и В. И. Киммельман. – Новосибирск: Изд-во НГТУ, 2019. – 356 с.
6. Кебец, П. Л. Маркеры перехода хода в дискурсе русского жестового языка / Материалы Седьмой Конференции по типологии и грамматике для молодых исследователей (г. Санкт-Петербург, 4-6 ноября 2010 г.). – СПб.: Наука, 2010. – С. 75-80.