

РАЗРАБОТКА АППАРАТНО-ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ ВОССТАНОВЛЕНИЯ МИМИЧЕСКОЙ АКТИВНОСТИ ПРИ АФАЗИИ

© 2017 И. А. Аполлонова, А. А. Полищук

Московский государственный технический университет имени Н. Э. Баумана

Разрабатываемый аппаратно-программный комплекс предназначен для обучения и восстановления речи при афазии. В данной статье рассматривается обзор методов восстановления речи, в том числе логопедических, обоснование необходимости отслеживания мимических движений, разработка медико-технических требований к аппаратно-программному комплексу (АПК) восстановления мимической активности и логопедического обучения. Также приводится обоснование необходимости создания базы данных по артикуляции здоровых людей с разметкой контрольных точек. Предложена программа исследований, позволяющая запланировать экспериментальные исследования и апробировать методику восстановления мимической активности при афазии.

Ключевые слова: афазия, артикуляция, контрольные точки, база данных, АПК (аппаратно-программный комплекс), фонемы, реабилитация инсульта, восстановление мимической активности.

По данным РОССТАТА за 2015 год по России заболевания нервной системы составляют 2257 из 113927 тысяч человек, травмы 13235 случаев на 113927 тысяч человек [1]. В США более 700 тысяч человек страдают от инсульта ежегодно, из них 85 % страдают афазией. При этом к 2020 году, по подсчетам ученых, количество заболевших удвоится [2]. Таким образом, можно утверждать, что афазия во всех своих проявлениях является достаточно распространенной болезнью. Сама по себе афазия – это системное нарушение речи, которое возникает при органических поражениях мозга, охватывает разные уровни организации речи, влияет на ее связи с другими психическими процессами и приводит к нарушению коммуникативной функции речи: нарушение собственно речи и вербального общения, нарушение других психических процессов, изменение личности и личностную реакцию на болезнь [3, 2]. Причинами развития такого вида нарушений чаще всего являются инсульт, лейкоэнцефалит, энцефалит, онкология, нейрохирургическое вмешательство, очаговая форма болезни Альцгеймера, черепно-мозговая травма, болезнь Пика, хроническая дисфункция центральной нервной системы и др. Основные симптомы прояв-

ляются в виде: хаотической перестановки слогов и звуков, пропуска букв и/или слогов в устной речи, нарушения запоминания названий предметов и т. д.

Внедрение АПК восстановления мимической активности при афазии сможет значительно снизить нагрузку на врачей, увеличить пропускную способность больниц и стационаров, даст пациенту возможность самостоятельно выполнять те или иные упражнения для восстановления мимической активности.

Методика лечения афазии

Из множества форм афазии, таких как афферентная моторная афазия, эфферентная моторная афазия, динамическая афазия, сенсорная афазия, акустико-мнестическая афазия, семантическая афазия и амнестическая афазия, выберем первую и рассмотрим ее методику лечения. Этот вид афазии вызывается поражением нижних отделов постцентральной области. Нарушение речевых кинестезий приводит к утрате тонких двигательных дифференцировок, необходимых для точной реализации артикуляторной программы [4, 194-204].

Таким образом, снижается четкость произношения звуков.

Обычно лечение афазии состоит из двух частей: медикаментозной и логопедической. В более тяжелых формах требуется оперативное вмешательство. Логопедические методики включают в себя:

- 1) тренировку лингвистических навыков
- 2) интонационную терапию

Аполлонова Ирина Анатольевна – МГТУ им. Н. Э. Баумана, зам. зав. кафедрой по учебной работе, к. т. н., доцент, apolloнова-i@yandex.ru.
Полищук Анастасия Александровна – МГТУ им. Н. Э. Баумана, студентка, gotikgit@mail.ru.

3) коррекцию речи с использованием современного компьютерного оборудования [5].

Непосредственно лечение проходит в несколько этапов. На раннем этапе главной задачей является активация процесса восстановления речи. Для этого существует комплекс методик для растормаживания и стимулирования нарушенных речевых функций, включающий в себя методику стимулирования устного высказывания при той или иной форме моторной афазии, методику растормаживания у больного непроизвольно протекающих речевых процессов (привычные речевые стереотипы, привычные ответы на вопросы, пение и т. д.). Для первичного растормаживания речи фразеологический материал подбирается логопедом по смысловой и эмоциональной значимости для больного.

На втором этапе у больного речевой синдром становится яснее, то есть проходят функциональные и нейродинамические компоненты, усложняющие картину речевого расстройства. Основной задачей является оживление и обогащение смысловой связи. Больной, овладевающий звуками: «а», «у», «м», «ы», «т», «и», «в», «с», «э», «н», «к», «и», «п», «л», «е», «ш», преодолевает апраксию артикуляционного аппарата. Оптико-тактильный метод, который основывается на сознательном выполнении пациентом речевых движений, организует осознанную речь больного: его внимание направляется на осознание и применение движений губ и языка в процессе общения.

На третьем этапе лечения основная задача – ритмическое выделение слоговой структуры, например, «ма-ма», «дя-дя», «ры-ба», «хо-чу».

На заключительных этапах лечения восстанавливают чтение и письмо пациента [6, 101-106].

Зарубежные и отечественные разработки АПК для восстановления речи

Существует несколько разработок, использующих изменение артикуляции или контура лица для определения речи с помощью ЭВМ. Например, в немецком технологическом институте разработали систему, которая по движению губ синтезирует речь. В основе этой технологии – электромиография, то есть отслеживание сигналов мимической активности при помощи электродов. Зарегистрированный мышечный импульс усиливается и отправляется на ЭВМ – направляется на обработку СПО на ПК, где эти сигналы преобразуются в текстовый файл [7, 1].

В России методики визуализации речи практически не представлены, из-за того, что в русском языке многие фонемы имеют одинаковые виземы. Некоторые согласные (глотательные или заднешелевые) воспроизводятся во рту и в горле, и не обнаруживаются визуально. Поэтому детекция производится по гласным фонемам, так как их виземы можно легко распознать и различить, и каждая визема уникальна.

Сравнение существующих АПК для распознавания речи и разработка требований для восстановления мимической активности при афазии

Один из вариантов распознавания визем, соответствующий гласным буквам, это обработка совокупности значений дескрипторов. Для этого применяется математический аппарат нейронных сетей или нечеткой логики. Нечеткий логический вывод формируется в несколько этапов:

- Введение нечеткости: функции принадлежности применяются к фактическим значениям входных переменных.
- Логический вывод: вычисляется значение истинности для предпосылок каждого правила и применяется к заключениям каждого правила.
- Композиция: нечеткие подмножества, назначенные каждой переменной вывода, объединяют в одно множество для всех переменных вывода.
- Приведение к четкости: используется в случаях, когда необходимо преобразовать нечеткий набор выводов в четкое число.

Преимуществом нейронных сетей является способность к обобщению после обучения на примерах, она может классифицировать ситуации, никогда ранее не наблюдавшиеся, связывая их со схожими ситуациями из примеров для распознавания визем.

К недостатку относится ограничение памяти АПК, так как распознавание виземы можно сравнить с системой набора текста для мобильных телефонов при помощи функции T9, и вероятность того, что АПК правильно распознает произнесенную визему при множестве вариантов одинакового расположения гласных в словах, низкая.

Существует метод для представления контуров губ в виде векторов признаков: контур губ выделяется на цветном изображении, после он проходит через оценочную функцию, то есть, определяется приблизительное положение центра губ на изображении, затем из его центра строится эллипс, который описывает и захватывает ротовую полость. Затем

уточняется контур, при помощи метода радиального расширения. Для распознавания движений губ выделяются векторы признаков из полученных данных. При поиске контура губ находятся n точек, которые нумеруют от p_1 до p_n по часовой стрелке.

Недостатком этого метода является преобразование изображения в цифровое пространство $r/g, b/g$. Кроме того, виземы по изображению не распознаются на должном уровне, из-за того, что они для различных фонем схожи [8, 2].

Контурный метод подразумевает такой анализ обработки изображений, при котором обработке подвергается лишь контур изображения, в то время как анализ отдельных точек не нужен. Задача распознавания изображений состоит из двух этапов. На первом этапе производится формирование модели контура изображения губ, а на втором этапе – распознавание. Распознавание изображения происходит следующим образом: на устройство распознавания подается реализация контура изображения губ $\chi(l)$, принадлежащая к одному из классов. Задача распознавания состоит в отнесении $\chi(l)$ к одному из классов букв («А», «Б» и т. д.). Для принятия решения формируется безусловное отношение правдоподобия, и оно сравнивается с пороговым значением.

Достоинством этого метода является достаточно высокая вероятность распознавания произвольного контура групп отнесения его к нужному классу ($P = 84\%$) [9, 1-6].

Метод активного контура позволяет анализировать движения губ, для которого главным источником информации являются характерные точки или области, или активный контур (МАК) [10, 1]. По геометрической модели лица человека выделяют дескрипторы, однозначно определяющие изменение мышц рта при разговоре. Недостатком данного метода является то, что программа недостаточно хорошо распознает изображение вне размеров 336 x 448 мм, поэтому необходимо корректировать размеры вручную. Также имеются проблемы с распознаванием людей негроидной расы.

Анализ требований для АПК по восстановлению мимической активности при афазии.

Область применения аппаратно-программного комплекса для восстановления мимической активности при афазии: коррекционная работа при афазии, возможная замена существующего метода электронейромиографии. В результате должен вос-

становиться артикуляционный аппарат и, как следствие, четкость произношения звуков у пациента. Непосредственное функциональное назначение изделия в лечебно-диагностическом процессе: это обучающая программа. То есть имеется заранее подготовленная программа, которая в своей базе данных хранит артикуляцию нескольких здоровых дикторов. Задача пациента – повторить увиденные и услышанные виземы. Тем самым будет разрабатываться мимическая активность пациента, так как его собственная артикуляция будет фиксироваться на веб-камеру и сравниваться с имеющейся базой данных. Система должна обеспечивать выполнение своих функциональных задач, а именно измерение положения контрольных точек, их обработку и анализ, сохранение результатов в базу данных, а также давать заключение о правильности произношения фонемы. Принцип действия основан на определении на лице пациента контрольных точек, распознавая положение которых, программа может утверждать о правильности или неправильности произнесения фонемы пациентом, исходя из имеющейся базы данных. Для корректного распознавания лица пациента и последующей расстановки контрольных точек пациент должен сидеть строго напротив веб-камеры, на расстоянии от нее приблизительно 50-70 см, так чтобы на экране четко отображалась зона артикуляции. Так как средством регистрации является веб-камера, то ее разрешение должно быть, как минимум, 1 280 x 720 для четкого выявления контуров губ на изображении. Продолжительность непрерывной работы за компьютером не должна превышать двух часов для 1 пациента. Средняя освещенность медицинской палаты, где возможно проведение восстановительной работы с пациентом, должна быть не менее 100 лк согласно СНиП 23-05-2010, СП 52.13330.2011 и Сан-ПиН 2.21/2.1.1.1278-03. Основываясь на пяти здоровых дикторах, выдвигаем следующие требования к погрешности для фонемы. Для корректного распознавания контура губ на лице относительная погрешность программной расстановки не должна превышать 5 %. При условии дальнейшего использования данных о положении контрольных точек автоматическая программная погрешность должна составлять не более 2 % для контрольных точек 1, 4, 7, 9, 11, 13, 15, 17. Не допускается ронять ПК и царапать объектив веб-камеры. Для самостоятельной работы пациента с обучающей программой доста-

точно провести вводный инструктаж и технику безопасности при работе с персональным компьютером.

Местоположение контрольных точек и их связь с мимической активностью

На губах можно выделить контрольные точки, положение которых, по отдельности или в сумме, однозначно определяет изменение мышц рта при разговоре. Чем точнее контрольные точки смещаются в заданных пределах, тем четче произносится слово или фонема. Программа WpfApplication1 позво-

ляет проводить все исследования. Расстановка этих контрольных точек представлена на рисунке 1.

Точки 1-2-3-4-5-6-7-8-9-10 характеризуют круговую мышцу рта; 1-2-3-4 и 4-5-6-7 – малую и большую скуловую мышцу; 2-3-4 – мышцу, поднимающую верхнюю губу; 8-9-10 характеризует мышцу, опускающую нижнюю губу и подбородочную мышцу; 6-7-8 и 2-1-10 – щечную мышцу; 1-10 и 7-8 – мышцу, опускающую угол рта [11].

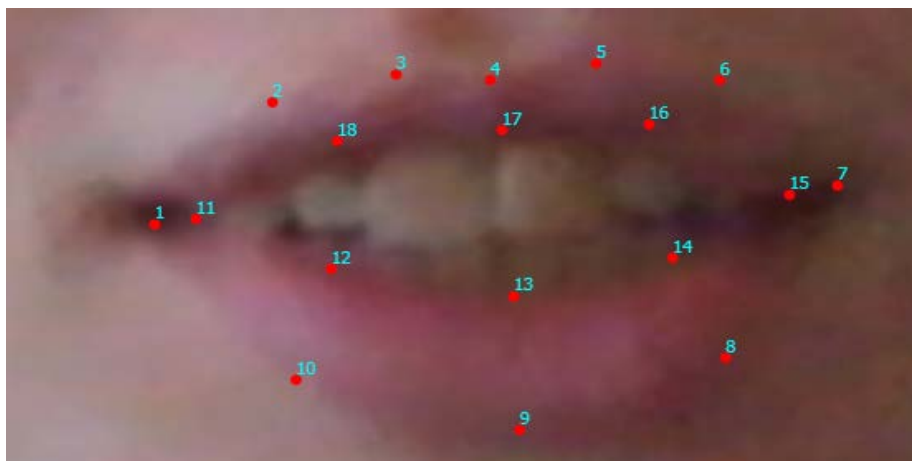


Рис. 1. Расстановка контрольных точек в программе WpfApplication1

Положение контрольных точек описывается в таблице 1.

Таблица 1

Контрольные точки и их положение в программе WpfApplication1

Номер в программе	Описание контрольной точки
1	Правый угол внешнего контура губ
2	Средняя точка между 9 и 1 во внешнем верхнем контуре губ
3	Правая верхняя точка дуги Купидона
4	Средняя точка верхнего внешнего контура губ
5	Левая верхняя точка дуги Купидона
6	Средняя точка между 4 и 7 во внешнем верхнем контуре губ
7	Левый угол внешнего контура губ
8	Средняя точка между 9 и 7 во внешнем нижнем контуре губ
9	Средняя точка нижнего контура губ
10	Средняя точка между 1 и 9 во внешнем нижнем контуре губ
11	Правый угол внутреннего контура губ
12	Средняя точка между 11 и 13 во внутреннем нижнем контуре губ
13	Средняя точка внутреннего нижнего контура губ
14	Средняя точка между 13 и 15 во внутреннем нижнем контуре губ
15	Левый угол внутреннего контура губ
16	Средняя точка между 17 и 15 во внутреннем верхнем контуре губ
17	Средняя точка внутреннего верхнего контура губ
18	Средняя точка между 17 и 11 во внутреннем верхнем контуре губ

Анализ экспериментальных данных по оценке артикуляции у здоровых людей и у больных афазией.

Для проверки работоспособности метода проводился эксперимент на макете АПК

восстановления мимической активности при афазии. Были отобраны 5 здоровых людей и 5 актеров, имитирующих, нарушения речи, связанные с афазией. Каждый из них произносил на камеру при заданном уровне освеще-

ценности 6 гласных фонем: «А», «О», «У», «Э», «И», «Ы». При работе использовалась камера «Fujifilm FinePix JZ300». Максимальное разрешение видеозаписи: 1 280 x 720. Максимальная частота кадров видеоролика: 30 кадров/с. Число кадров в секунду при разрешении видеозаписи 1 280 x 720 равно 24. Далее все полученные 60 видео формата .mp4. при помощи программы Format Factory конвертировались в .avi. для

удобства в дальнейшей работе и обработке. После чего каждое видео было раскодировано программой VirtualDub (видеопоток преобразуется в набор отдельных файлов, которые соответствуют статическим видеокадрам). Каждый кадр назывался в виде sp1phAfr0.jpg, алгоритм кодирования названия представлен в таблице 2. Для того, чтобы далее используемые программы могли идентифицировать эти кадры.

Таблица 2

Алгоритм составления названия файлов

sp (от. англ. speaker)	Переводится как «диктор»
0, 1, 2,.....	Порядковый номер диктора
ph (phoneme)	Переводится как «фонема»
A, N, P,.....	Название фонемы
fr (frame)	Переводится как «кадр»
1, 2, 3,.....	Порядковый номер кадра

Для того, чтобы расставить контрольные точки, необходимо загрузить две программы ASM Photo и WpfApplication1. Первая создает файл в формате .xml. (содержание файла представлено на рисунке 2) с положением контрольных точек в системе отсчета XY, а вторая расставляет эти точки непосредственно на кадры sp1phAfr0.jpg, sp1phAfr1.jpg и т. д.

Реализуемая программами ASM Photo и WpfApplication1 автоматическая расстановка контрольных точек представлена на рисунке 3 при произнесении диктором фонемы «И». Как можно заметить, расстановка некоторых контрольных точек не является полностью корректной, что в дальнейшем может привести к ошибкам в обработке данных и, соответственно, к ошибкам в АПК для восстановления мимической активности при афазии.

```

<!--Outer lip contour points-->
<Point y="991" x="461" n="1"/>
<Point y="959" x="495" n="2"/>
<Point y="945" x="529" n="3"/>
<Point y="949" x="556" n="4"/>
<Point y="944" x="584" n="5"/>
<Point y="954" x="615" n="6"/>
<Point y="982" x="641" n="7"/>
<Point y="1017" x="606" n="8"/>
<Point y="1033" x="555" n="9"/>
<Point y="1022" x="501" n="10"/>
<!--Inner lip contour points-->
<Point y="991" x="466" n="11"/>
<Point y="996" x="513" n="12"/>
<Point y="998" x="555" n="13"/>
<Point y="993" x="596" n="14"/>
<Point y="982" x="636" n="15"/>
<Point y="974" x="595" n="16"/>
<Point y="972" x="555" n="17"/>
<Point y="976" x="514" n="18"/>

```

Рис. 2. Координаты контрольных точек, создаваемые программой ASM Photo

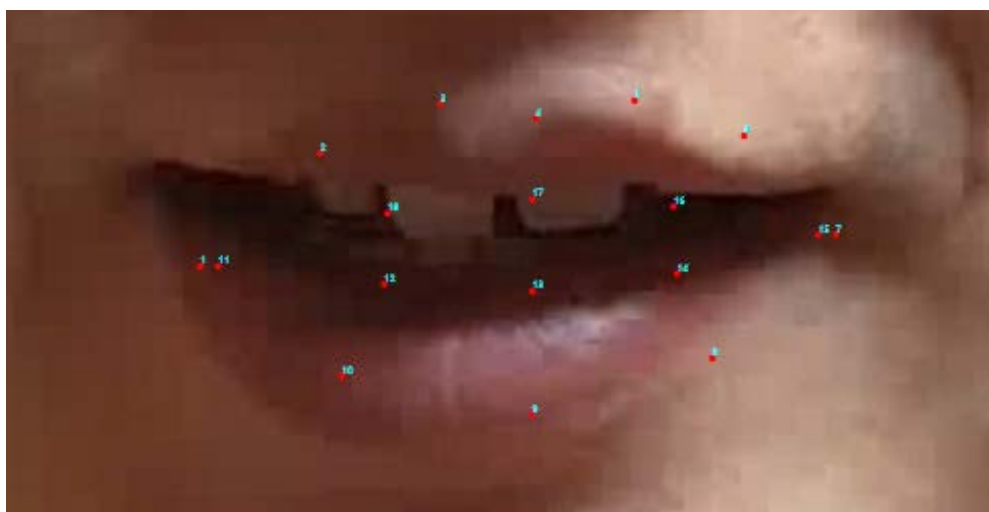


Рис. 3. Автоматическая расстановка контрольных точек программами ASM Photo и WpfApplication1

Для предотвращения ошибок в дальнейшей обработке, необходимо корректировать положение контрольных точек вручную (рис. 4).



Рис. 4. Откорректированные вручную контрольные точки для фонемы «И»

Координаты контрольных точек после автоматической и ручной расстановки занесем в таблицу 3 для выявления относительной погрешности. Выберем точки 1, 4, 7, 9, 11, 13, 15, 17.

Таблица 3

Координаты контрольных точек при автоматической и ручной расстановке

Номер контрольной точки	Координата	Автоматическая расстановка контрольных точек	Ручная расстановка контрольных точек	Относительная погрешность, %
1	X	461	448	2.90
	Y	991	964	2.80
4	X	556	540	2.96
	Y	949	950	0.11
7	X	641	639	0.31
	Y	982	971	1.13
9	X	555	536	3.54
	Y	1033	1037	0.39
11	X	466	452	3.10
	Y	991	965	2.69
13	X	555	536	3.54
	Y	998	991	0.71
15	X	636	636	0
	Y	982	973	0.92
17	X	555	545	1.83
	Y	972	969	0.31

Как видно из таблицы 3 для точек 1, 4, 9, 11 и 13 относительная погрешность выше, чем допустимое значение в 2 %, что доказывает, что автоматическая расстановка контрольных точек для АПК для восстановления мимической активности при афазии не является полностью достоверной. Поэтому в перспективе предлагается использовать дополнительный фильтр для перевода изображения в другой цветовой диапазон [8, 1] или инфракрасный датчик, отслеживающий движения артикуляторного аппарата. Предполагается, что предложенные, или иные альтернативные методы помогут повысить точность расстановки контрольных точек автоматизированным методом на 30 %. Также предлагается использовать дополнитель-

ное оборудование для стимуляции лицевых мышц, что будет способствовать восстановлению мимической активности при афазии.

Выводы

В соответствии с алгоритмом были составлены контрольные точки для шести гласных фонем. При автоматической расстановке контрольных точек программами ASM Photo и WpfApplication1 для фонемы «И» были найдены следующие значения относительной погрешности: по x-координате от 1.08 % до 5.34 %, по y-координате от 0.57 % до 1.86 %, средняя погрешность в расстановке контрольных точек составила 1.98 %.

Для фонемы «А» были найдены следующие значения относительной погрешности: по x-координате от 0.97 % до 3.84 %, по

у-координате от 0.52 % до 2.42 %, средняя погрешность в расстановке контрольных точек составила 1.80 %.

Для фонемы «О» были найдены следующие значения относительной погрешности: по х-координате от 0 % до 3.80 %, по у-координате от 0 % до 1.61 %, средняя погрешность в расстановке контрольных точек составила 1.41 %.

Для фонемы «Э» были найдены следующие значения относительной погрешности: по х-координате от 0.11 % до 3.28 %, по у-координате от 0.45 % до 2.23 %, средняя погрешность в расстановке контрольных точек составила 1.32 %.

Для фонемы «У» были найдены следующие значения относительной погрешности: по х-координате от 0.58 % до 4.84 %, по у-координате от 0.02 % до 0.99 %, средняя погрешность в расстановке контрольных точек составила 1.30 %.

Для фонемы «Ы» были найдены следующие значения относительной погрешности: по х-координате от 0 % до 4.86 %, по у-координате от 0.17 % до 1.15 %, средняя погрешность в расстановке контрольных точек составила 1.59 %.

Относительная погрешность всех рассмотренных фонем соответствует установленному требованию для АПК для восстановления мимической активности при афазии (< 2 %).

Проведенные на макете АПК для восстановления мимической активности при афазии испытания, доказали работоспособность метода и его технической реализации, а также показали, что можно распознать, оценить, определить качественную и количественную разницу в артикуляции здорового человека и больного афазией с точностью до 2 %.

ЛИТЕРАТУРА

DEVELOPMENT OF HARDWARE-SOFTWARE SYSTEM FOR APHASIA'S MIMIC ACTIVITY RESTORATION

© 2017 I. A. Apollonova, A. A. Polishchuk

Bauman Moscow State Technical University

The developed hardware-software complex is intended for training and aphasia's speech restoration. In this article reviews the methods of speech restoration, including logopedic speech therapy, the rationale for monitoring the mimic movements, as well as the development of medical and technical requirements for the hardware-software complex (HSC) for restoration mimic activity and logopedic training, and the rationale for creating an articulation database of healthy people with fixation of control points. Based on the results of the research, a series of experiments will be carried out that will allow testing the technique for restoring the mimic activity during aphasia.

Keywords: aphasia, articulation, control points, database, HSC (hardware-software complex), phonemes, rehabilitation of stroke, restoration of facial activity.

1. ВОЗ. Данные Глобальной обсерватории здравоохранения. Всемирная организация здравоохранения. [В Интернете] 2014 [Цитировано: 24 Февраль 2017] <http://www.who.int/>.

2. Aphasia Statistics. Pinterest [В Интернете] [Цитировано: 24 Март 2017] <https://ru.pinterest.com/pin/430516045599398002/>

3. Рябова (Ахутина) Т. В. Механизм поражения речи по данным афазиологии / Т. В. Рябова (Ахутина) // Вопросы психолингвистики. Институт языкознания РАН. – 2011. – 2 (14).

4. Шохор-Троцкая М. К. Коррекция сложных речевых расстройств / М. К. Шохор-Троцкая. – Москва: ТЦ Сфера, 2006.

5. Завистовская Т. А. Автоматизация процесса формирования текстовых сообщений на основе последовательности цифровых изображений лица оператора при наличии сильных шумов / Т. А. Завистовская, Е. Е. Ковшов. б.м. : ФГБОУ ВПО МГТУ «СТАНКИН».

6. Soldatov. Lip Reading: Preparing Feature Vectors / Soldatov, Stanislav. – Moscow: Graphics&Media Laboratory.

7. Яранцева Т. В. Применение алгоритма распознавания изображений со случайной формой для чтения по губам / Т. В. Яранцева, Т. П. Киселева, Р. Г. Хафизов // Методы и устройства передачи и обработки информации. – 2015. – Т. 17. – С. 52-57.

8. Рогозина М. М. Автоматическое определение дескрипторов эмоционального состояния на основе системы кодирования лицевых движений / М. М. Рогозина. – МИ-ЭМ, 2012.

9. Deagostini. Мышцы лица. Тело человека снаружи внутри . Deagostini, 2008 г., 11.

11. Deagostini. Мышцы лица. Тело человека снаружи и внутри // Deagostini. 2008.