

АНАЛИЗ ВЛИЯНИЯ РАЗЛИЧНЫХ ФАКТОРОВ НА КАЧЕСТВО РАСПОЗНАВАНИЯ УСТНОЙ РЕЧИ С ГУБ ГОВОРЯЩЕГО

Мясоедова М.А.¹, Мясоедова З.П.²

*(Учреждение Российской академии наук
Институт проблем управления РАН, Москва)*

В статье рассматривается влияние определённых внешних факторов на процесс восприятия устной речи зрительным способом, т.е. с губ говорящего. Дана оценка этому влиянию на качество распознавания речи по результатам проведения ряда экспериментов. Описана обучающая программа «Читаем с губ».

Ключевые слова: артикуляционный рисунок, чтение с губ, проверка статистических гипотез, парный t-критерий Стьюдента.

1. Введение

Глухой человек в сравнении со слышащими людьми обладает уникальной способностью ориентироваться в окружающем его мире в основном с помощью зрения. Испытывая при общении с окружающими большие труд-

¹ Мясоедова Мария Александровна, ведущий инженер по АСУП (mariatarfi@mail.ru).

² Мясоедова Зинаида Павловна, научный сотрудник (mzinap@mail.ru, Москва, ул. Профсоюзная, д. 65, тел. (495) 334-90-60).

ности, глухой человек вынужден внимательно следить за выражением лица и движением губ говорящего, что даёт ему возможность в некоторой степени «услышать» и правильно понять речь посредством чтения с губ.

Чтение с губ подразумевает не всегда полное распознавание сообщения. Чаще всего происходит определение лишь нескольких фонов речевых элементов, а остальная работа заключается уже в догадке с учётом контекста и всего многообразия выразительных средств (взгляд, мимика, жесты и т.д.), используемых собеседником, что позволяет человеку с определённой вероятностью понять смысл переданного сообщения [1].

Процесс чтения с губ проходит довольно сложно и зависит от многих внешних факторов. В статье приведены результаты экспериментов, позволяющих оценить влияние различных внешних факторов на качество распознавания речи при считывании её с губ, и описана реализация проекта «Читаем с губ», предназначенного для обучения пользователей с нарушениями слух, а навыкам визуального восприятия устной речи.

В рамках данного проекта проведена работа по двум направлениям, соответствующим проведению экспериментов при непосредственном контакте их участников с диктором и в режиме on-line на основе созданной собственной базы, содержащей около двух тысяч слов русского языка, отсортированных по определённым признакам.

2. Факторы, влияющие на чтение с губ

Залог успешного овладения чтением с губ определён рядом факторов как внутреннего характера, связанных исключительно с личными качествами человека, воспринимающего речь (уровень владения им языком, включая

грамотность и объём словарного запаса; умственные способности; характер; возраст и т.д.), так и факторов внешнего характера, к которым относятся: *диктор* (артикуляция, дефекты речи, характер и т.п.), и *окружающее пространство* (освещённость, посторонний шум, расстояние и т.д.) и *сообщение* (стиль языка, структура, контекст и т.п.).

С целью выявления воздействия различных факторов их на качество восприятия речи зрительным способом был проведён ряд экспериментов (табл. 1), для чего была создана экспериментальная группа из 10 человек разного пола, возраста и состояния слуха. Проведение экспериментов на основе созданной группы и при участии одного постоянного диктора, произносящего в каждом из них индивидуальный набор слов одинакового числа, проходило в два этапа. На первом этапе распознавание проходило в обычных условиях, т.е. при отсутствии посторонних факторов, а на втором – с их учётом.

Таблица 1. Факторы, учтённые в экспериментах.

Экспер.	Факторы
1	F_1 – повторное произнесение слов
2	F_2 – дикторы (артикуляция)
3	F_3 – подсказка (контекст)
4	F_4 – посторонние предметы на лице диктора
5	F_5 – посторонний шум
6	F_6 – поворот головы диктора в сторону
7	F_7 – движение головы диктора
8	F_8 – увеличение расстояния до диктора
9	F_9 – ухудшение освещения лица диктора
10	$F_1 + F_3$ – повторное произнесение слов + подсказка
11	$F_1 + F_9$ – повторное произнесение слов + плохое освещение лица диктора
12	$F_4 + F_6$ – посторонние предметы на лице диктора + поворот его головы в сторону

3. Анализ и оценка полученных результатов

На основании результатов, полученных при проведении ряда экспериментов до и после воздействия соответствующих им факторов, дана оценка степени их влияния на качество восприятия устной речи с губ говорящего.

В статье дано описание экспериментов с учётом таких факторов, как смена дикторов и характер сообщения.

3.1. ЭКСПЕРИМЕНТ А

Данный эксперимент был направлен на выявление воздействия чёткости артикуляции говорящего на качество распознавания участниками слов, произнесённых в произвольном порядке вначале основным диктором, участвующим во всех экспериментах, а затем новым диктором.

Построенная согласно полученным результатам гистограмма (рис.1) наглядно демонстрирует разницу в распознавании участниками эксперимента (R_i) слов, произнесённых разными дикторами.

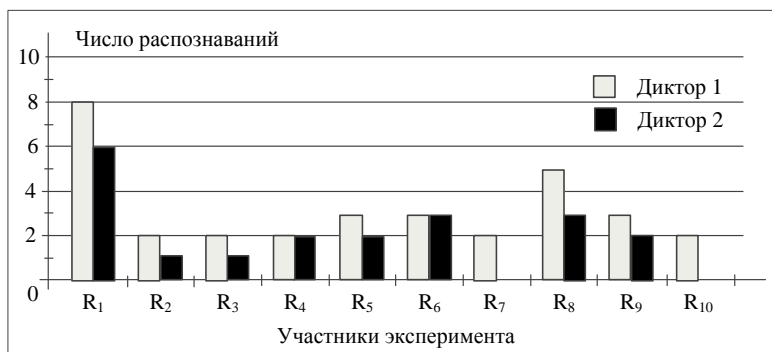


Рис. 1. Распознавание слов по всем участникам эксперимента

Графики, приведённые на рис. 2, отражают разное восприятие предложенных к распознаванию слов, зависящее от артикуляции дикторов.

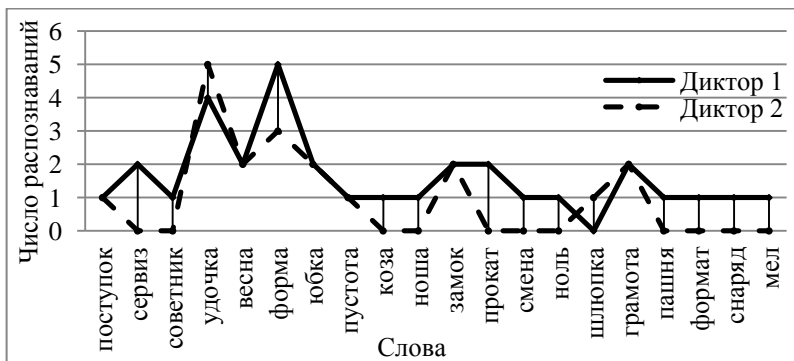


Рис. 2. Распознавание слов в зависимости от артикуляционных характеристик дикторов

Сравнение полученных результатов по двум этапам эксперимента, рассмотренное как измерение парных разностей, проведено с помощью парного t- критерия Стьюдента. Проверка статистических гипотез, представляющих собой предположение о случайном или значимом сходстве или различии характеристик выборки, позволила признать влияние фактора существенным и на этой основе определить показатели генеральной совокупности [2].

Для выявления степени воздействия на качество распознавания речи при чтении с губ указанных в табл. 1 факторов, рассмотрены общие результаты по этапам всех экспериментов, проведённых в обычных условиях и с учётом соответствующих им факторов (рис. 3).

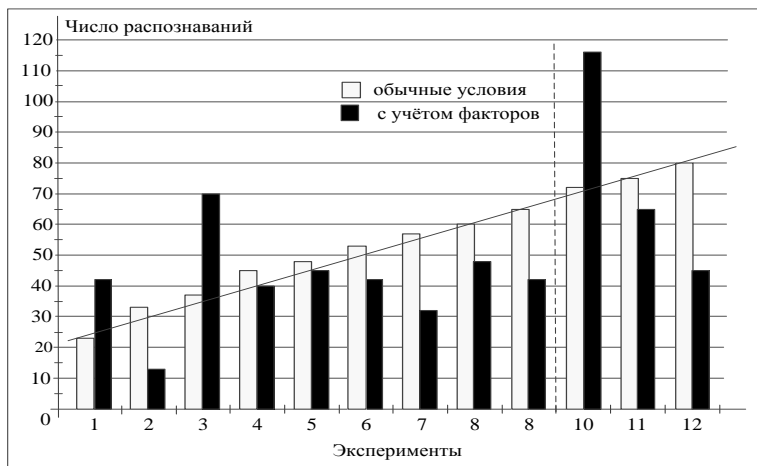


Рис. 3. Распознавание слов по всем экспериментам

Из рис. 3 следует, что повторное произнесение слов и подсказка (контекст), а также совместное влияние этих факторов существенно повышают качество распознавания. Наибольшее влияние негативного характера оказывают плохая артикуляция дикторов, движение головы диктора и плохая освещённость его лица, а также совместное воздействие таких факторов, как посторонние предметы на лице диктора и поворот его головы в сторону.

Из данного рисунка также виден рост общего числа распознаваний всеми участниками в обычных условиях с каждым следующим экспериментом, что свидетельствует о значении полученного ими опыта, дающего возможность повысить свой навык в чтении речи с губ.

3.2. ЭКСПЕРИМЕНТ Б

Серьёзные проблемы при распознавании речи с губ говорящего возникают при наличии в сообщении слов-паронимов, имеющих при произнесении схожие артикуля-

ционные рисунки. Для определения степени зависимости качества распознавания речи от характера сообщения был проведён ряд экспериментов на основе использования групп слов-паронимов различной структуры и состава.

Описание экспериментов дано на примере использования группы слов-паронимов (табл. 2), отличающихся между собой числом и составом фонем в первом слоге. Второй слог для всех слов является общим. В данной таблице жирным шрифтом выделены пары фонем, не отличающиеся между собой артикуляционным рисунком («г–к»), либо имеют слабо выраженное отличие («з–с»), либо имеют между собой некоторые отличительные особенности.

Таблица 2. Структура двуслоговых слов-паронимов

Слова	г–к	л–р	а–я	з–с	ный
разный	–	р	а	з	ный
рясный	–	р	я	с	
грязный	г	р	я	з	
красный	к	р	а	с	
гласный	г	л	а	с	
классный	к	л	а	с (сс)	

Схема образования слов-паронимов, использованных в данном эксперименте, приведена на рис. 4. Двигаясь от одного из входов (жирные стрелки) по определённой цепочке, получаем каждое указанное в табл. 2 слово.

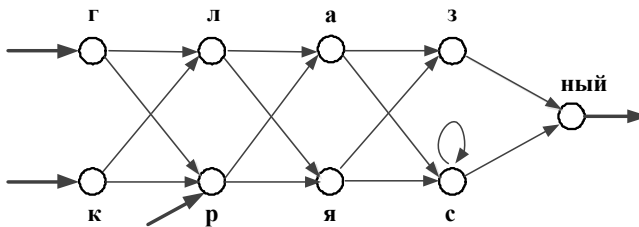


Рис. 4. Схема образования двуслоговых слов-паронимов

Для определения правильного распознавания каждого из слов был проведён эксперимент, результаты которого позволили выявить частоту распознавания каждого из слов и замены его сходным по артикуляционному рисунку другим словом данной группы.

Взаимозаменяемость слов внутри одной группы продемонстрирована на построенной в соответствии с полученными результатами схеме (рис. 5).

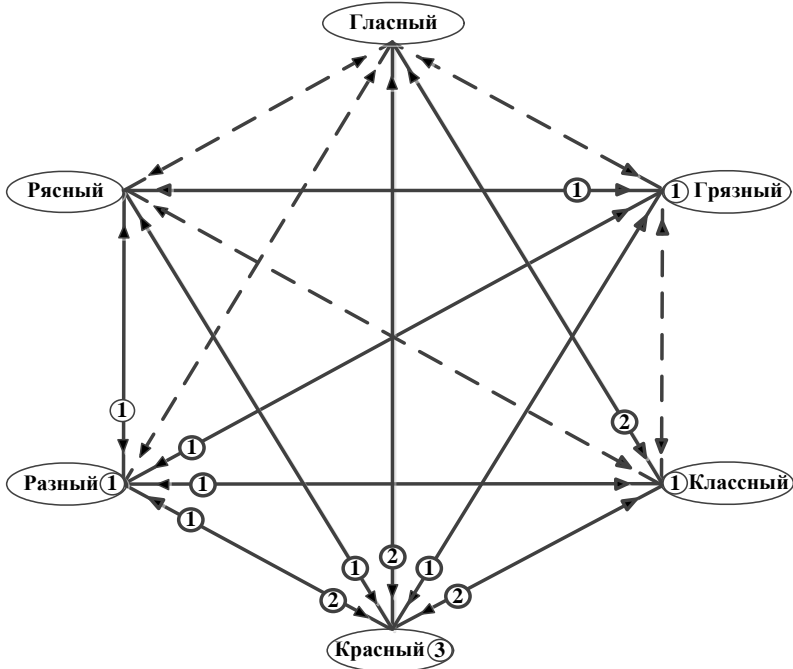


Рис. 5. Распознавание слов-паронимов всеми участниками эксперимента

Число правильных распознаваний по каждому из слов указано рядом с его названием, а отсутствие стоящей рядом с ним цифры означает, что данное слово в процессе

проведения эксперимента не было распознано ни разу ни одним из его участников. Цифры, стоящие перед стрелкой одного из слов, соответствуют числу замен им другого слова, расположенного по другую сторону линии. Штриховая линия соединяет слова, по которым не было взаимозаменяемости.

Результаты, отсортированные по частоте указания каждого из слов в ответах участников эксперимента, позволили выявить из них наиболее «популярных», что отражено на диаграмме (рис. 6).

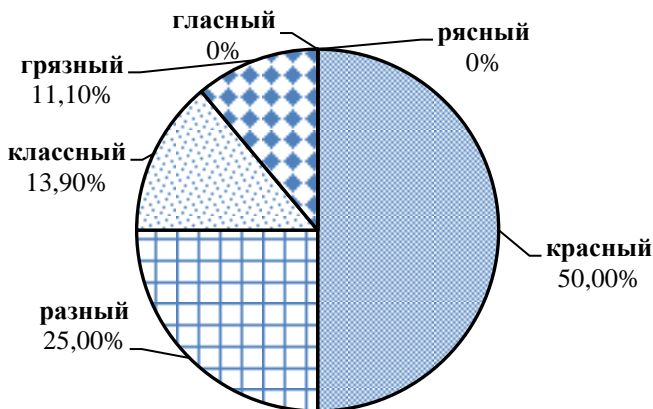


Рис. 6. «Популярность» сходных по артикуляционному рисунку слов при распознавании их всеми участниками

Из приведённого рис. 6 видно, что слова «красный» и «разный» являются самыми «популярными» словами, содержащимися в ответах участников.

Подводя итог, можно сказать, что распознавание слов-паронимов со схожими артикуляционными рисунками при их автономном произнесении не позволяет получить хорошие результаты при чтении с губ даже человеку, неплохо владеющему этим мастерством.

4. Программа «Читаем с губ»

Обучающая on-line программа «Читаем с губ» предназначена для людей с потерей слуха, желающих получить определённый навык по восприятию устной речи собеседника путём считывания её с губ говорящего. Эта программа также может быть полезна для общения с людьми, лишёнными в силу определённых причин возможности полноценно использовать свой речевой аппарат, которых можно понять лишь по движению губ и мимике [3].

В основе работы программы заложен принцип сравнения схожих по артикуляционному рисунку речевых элементов, хранящихся в базе данных с числом элементов около 1800. Наблюдение за произнесением диктором разных элементов позволяет более внимательно следить за движением его губ и выявить тонкости при восприятии слов, схожих по артикуляционному рисунку.

Использование данной программы позволяет в автоматическом режиме фиксировать число правильных ответов по каждому из пользователей, что даёт возможность по окончании сеанса провести анализ полученных каждым пользователем результатов, дать в целом оценку его способности к считыванию речи с губ говорящего, показать динамику результатов обучения.

При оформлении тестовых заданий были использованы стандартные программные средства: Flash, ActionScript, PHP, MySQL, XML, HTML.

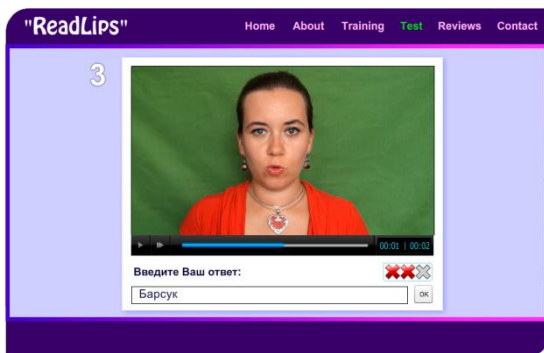
Работа программы по усмотрению пользователя может проходить в двух режимах:

- *демонстрационном*, носящем ознакомительный характер с просмотром хранящихся в базе данных отдельных видеофайлов, демонстрирующих произнесение диктором речевых элементов русского языка, сходных по

артикуляционному рисунку, в сопровождении соответствующего текста (титров);

– *рабочем*, предусматривающем на основе просмотра пользователем представленного диктором без звукового сопровождения слова, произвольно выбранного из базы видеофайлов, получение от него результата распознавания. При этом пользователю предоставляется возможность выбора режимов работы: скорость (обычный речевой темп или замедленный), наличие определённых внешних факторов, использование подсказки, которая, в соответствии с категорией представленного к распознаванию слова, выражается в выводе списка всех слов, входящих в данную категорию.

На рис. 8 приведена экранная форма рабочего режима программы.



*Рис. 8. Экранная форма рабочего режима проекта
«Читаем с губ»*

Данная программа позволяет самостоятельно овладеть навыками чтения с губ благодаря преимуществам, заключающимся в возможности работы большого числа пользователей в режиме on-line; лёгкости доступа к воспроизведению или пересмотру видеозаписи ролика неог-

раниченное число раз; гибкости обучения с возможностью работы в удобном для себя темпе; возможности самоконтроля выполнения заданий.

4. Заключение

Результативность чтения с губ достигается при наличии навыков в этом процессе, полученных в результате длительных и серьёзных тренировок.

Знание факторов, влияющих на распознавание речи путём чтения губ, и применение компьютерных технологий и обучающих программ для овладения этим навыком, существенно помогут людям с ослабленным слухом научиться лучше распознавать звучащую речь визуальным способом.

Овладение этим мастерством позволит улучшить способность коммуникации людей со слуховой патологией со слышащими, что чрезвычайно важно для общества.

Литература

1. МЕТТ А.И., НИКИТИНА Н.А. *Зрительное восприятие устной речи.* – М.: Просвещение, 1965. – С.197.
2. С. ГЛАНЦ *Медико-биологическая статистика* // М.: Практика, 1999. – С. 462.
3. ЗАХАРОВА И.Г. *Информационные технологии в образовании* // Российский фонд фундаментальных исследований: десять лет служения российской науке. – М.: Издательский центр «Академия», 2003. – С. 192.

ANALYSIS OF THE INFLUENCE OF VARIOUS FACTORS ON QUALITY RECOGNITION SKILLS WITH THE SPEAKER'S MOUTH

Maria Myasoedova, Institute of Control Sciences of RAS, Moscow, leading engineer for CAM (mariamarfi@mail.ru).

Zinaida Myasoedova, Institute of Control Sciences of RAS, Moscow, research associate (mzinap@mail.ru, Moscow, Profsoyuznaya st., 65, (495)334-90-60).

Abstract: The article examines the impact of certain external factors on the process of perception of oral speech visual way, ie with the speaker's mouth. The estimation of this influence on the quality of speech recognition on the results of a series of experiments. Described training program "Read lips."

Keywords: articulation pattern, lip-reading, statistical hypothesis testing, paired Student t-test.