

Интернет-журнал «Наукоедение» ISSN 2223-5167 <https://naukovedenie.ru/>

Том 9, №6 (2017) <https://naukovedenie.ru/vol9-6.php>

URL статьи: <https://naukovedenie.ru/PDF/20TVN617.pdf>

Статья опубликована 28.11.2017

Ссылка для цитирования этой статьи:

Колесникова Д.С., Рудниченко А.К., Верещагина Е.А., Фоминова Е.Р. Применение современных технологий распознавания речи при создании лингвистического тренажера для повышения уровня языковой компетенции в сфере межкультурной коммуникации // Интернет-журнал «НАУКОВЕДЕНИЕ» Том 9, №6 (2017) <https://naukovedenie.ru/PDF/20TVN617.pdf> (доступ свободный). Загл. с экрана. Яз. рус., англ.

УДК 004.934.2

Колесникова Дарья Сергеевна

ФГАУО ВО «Дальневосточный федеральный университет», Россия, Владивосток¹

Магистрант

E-mail: kolesnikova_dse@students.dvfu.ru

РИНЦ: http://elibrary.ru/author_profile.asp?id=936048

Рудниченко Алексей Константинович

ФГАУО ВО «Дальневосточный федеральный университет», Россия, Владивосток

Магистрант

E-mail: rudnichenko_ak@students.dvfu.ru

РИНЦ: http://elibrary.ru/author_profile.asp?id=936050

Верещагина Елена Александровна

ФГАУО ВО «Дальневосточный федеральный университет», Россия, Владивосток

Доцент кафедры «Информационной безопасности»

Кандидат технических наук

E-mail: everesh@mail.ru

РИНЦ: http://elibrary.ru/author_profile.asp?id=287436

Фоминова Екатерина Романовна

ООО «Информационный центр», Россия, Владивосток

Начальник отдела защиты информации

E-mail: katenski.work.94@mail.ru

**Применение современных технологий
распознавания речи при создании лингвистического
тренажера для повышения уровня языковой компетенции
в сфере межкультурной коммуникации**

Аннотация. Тема, представленная в данной статье, является довольно актуальной в современном мире. Повсеместное внедрение технологий распознавания позволяет улучшать навыки людей в различных областях знаний, среди которых особое место занимает обучение иностранным языкам. В данной статье представлен краткий обзор технологии распознавания образов, определение важности распознавания речи в современных условиях, а также

¹ 690922, Россия, Приморский край, г. Владивосток, Аякс-10 корпус D

актуальности создания лингвистического тренажера для обучения иностранным языкам с учетом особенностей говорящего.

В работе авторами были рассмотрены особенности распознавания голоса вне зависимости от используемых технологий, а также существующие математические алгоритмы, лежащие в основе речевой биометрии, а в частности более подробно изучена структура скрытых марковских моделей и их применение на практике для синтеза и распознавания речи.

Было обнаружено, что основной проблемой при распознавании речевых сигналов является наличие акцента и различных речевых особенностей говорящего. В связи с этим был разработан проект языкового тренажера, который бы сводил к минимуму данную проблему и помогал обучающимся улучшать произношение. В основе работы такого тренажера лежит комплексный подход к изучению иностранного языка, так как объединяет в себе четыре важные составляющие: письмо, аудирование, чтение, говорение.

Вклад авторов. Колесникова Дарья Сергеевна – автор осуществил написание статьи. Изучил математические аспекты распознавания звуковых сигналов, произвел анализ особенностей распознавания речи, описывал работу лингвистического тренажера. Рудниченко Алексей Константинович – автор осуществил написание статьи. Проанализировал текущее состояние в области развития современных распознающих систем, в том числе систем по распознаванию речи. Фоминова Екатерина Романовна – автор оказывал участие в описании работы лингвистического тренажера, собрал и проанализировал сведения об особенностях человеческой речи. Верещагина Елена Александровна – автор оказывал участие в написании статьи, формировал структуру материала; одобрил окончательную версию статьи перед её подачей для публикации.

Ключевые слова: распознавание; распознающие системы; речь; распознавание речи; биометрия; речевая биометрия; скрытые марковские модели; pattern recognition; искусственный интеллект; лингвистический тренажер; обучение; иностранный язык; акцент

Введение

Исследования в области искусственного интеллекта (ИИ) представляют в настоящее время огромный интерес для человечества. В частности, одно из направлений ИИ – машинное обучение – довольно широко применяется на практике и решает класс задач по распознаванию образов (Pattern Recognition).

Распознавание образов – это научная дисциплина, целью которой является классификация объектов по нескольким категориям или классам. Объекты называются образами².

За последние несколько десятилетий распознающие системы стали доступны только благодаря значительным достижениям в области компьютерной обработки. Однако многие из этих новых автоматизированных методов основаны на идеях, которые изначально были задуманы сотни и даже тысячи лет назад. Один из старейших и наиболее основных примеров характеристики, которая использовалась для распознавания людьми, – это лицо. Данный способ идентификации известных (знакомых) личностей был удобен для малых групп населения, но с течением времени появлялись новые способы путешествий и количество людей (а, значит, и объектов для распознавания) стремительно увеличивалось благодаря таким

² НОУ ИНТУИТ | Лекция | Задача распознавания образов
<http://www.intuit.ru/studies/courses/2265/243/lecture/6241%3Fpage%3D2>.

миграциям. В связи с этим возникла необходимость в новых систематизированных способах распознавания, в том числе основанных на поведенческих характеристиках человека (например, речь или походка)³. На сегодняшний день в качестве объекта распознавания могут выступать:

- лицо, отпечаток пальца и другие биометрические параметры человека;
- звуковой сигнал (в том числе речевой);
- символы (например, в виде текста);
- различные изображения;
- другие объекты.

Несмотря на то, что создание универсальных алгоритмов с точностью распознавания до 100 % является крайне трудоемким процессом и на данный момент представляется практически невозможным, уже существуют распознающие системы, показывающие высокий процент распознавания и активно используемые в разных сферах деятельности человека (рис. 1).



Рисунок 1. Применение теории распознавания образов на практике (разработано авторами)

Среди представленных выше примеров распознавание речевых сигналов является активно развивающейся областью науки. Речевая биометрия является одним из способов идентификации личности для обеспечения информационной безопасности (голосовая верификация) [3], используется для автоматического заполнения документов (например, в медицинских организациях при заполнении амбулаторных карт или в судах для автоматизации протоколирования), а также применяется правоохранительными органами и службой безопасности для детектирования опасных переговоров в целях предотвращения террористических актов⁴. В данной статье рассмотрен еще один из частных примеров применения этой технологии на практике, а именно при создании обучающего языкового тренажера.

³ History of Biometrics | BiometricUpdate <http://www.biometricupdate.com/201501/history-of-biometrics>.

⁴ Искусственный интеллект (ИИ, Artificial Intelligence, AI) [http://www.tadviser.ru/index.php/Продукт:Искусственный_интеллект_\(ИИ,_Artificial_intelligence,_AI\)](http://www.tadviser.ru/index.php/Продукт:Искусственный_интеллект_(ИИ,_Artificial_intelligence,_AI)).

1. Постановка задачи распознавания речи

Существует множество различных информационных технологий, позволяющих усовершенствовать навыки аудирования, чтения и письма при изучении иностранных языков, но направлению улучшения навыков разговорной речи уделяется недостаточно внимания. Но в то же время для обучающегося важно научиться правильно произносить звуки, буквы и целые слова, так как это является основой для коммуникации с носителем изучаемого языка. На сегодняшний день наиболее перспективным направлением при решении данной проблемы является внедрение технологий распознавания речи, в основе которых лежат особенности голоса обучающегося [4, 9].

Каждый человек обладает уникальным голосом, но с фонетической точки зрения речь состоит из множества различных звуков, имеющих артикуляционные различия. Эти звуки в общем случае принято называть *фонемами*. Но в разных словах одни и те же фонемы могут видоизменяться, поэтому в практику также вводятся *аллофоны* – варианты фонем⁵ (рис. 2).

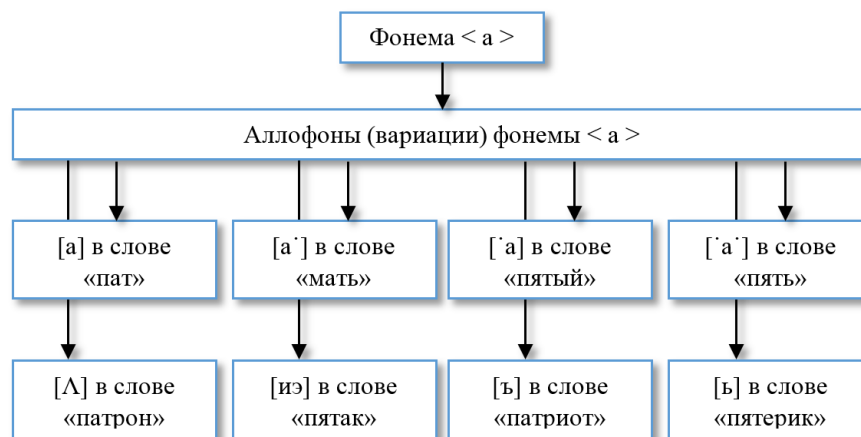


Рисунок 2. Представление аллофонов фонемы (разработано авторами)

Для успешного распознавания речи обычно рассматривают участки звукового сигнала в несколько десятков миллисекунд, называемые *фреймами* (рис. 3) [7]. Выполнить эту задачу довольно сложно ввиду того, что некоторые фонемы достаточно похожи друг на друга, но можно решить ее в терминах «вероятностей». Одни фонемы более вероятны для данного сигнала, другие – менее. Строится акустическая модель, которая является функцией, принимающей на вход участок небольшого звукового сигнала (фрейм) и выдающей распределение вероятностей различных фонем на этом фрейме. На основе акустической модели можно с определенной степенью уверенности восстановить то, что было произнесено⁶.

⁵ Фонема и ее основные функции. Варианты и вариации фонемы <http://linguistics-konspect.org/?content=5776>.

⁶ Распознавание речи от Яндекса. Под капотом у Yandex.SpeechKit / Блог компании Яндекс / Хабрахабр <https://habrahabr.ru/company/yandex/blog/198556/>.

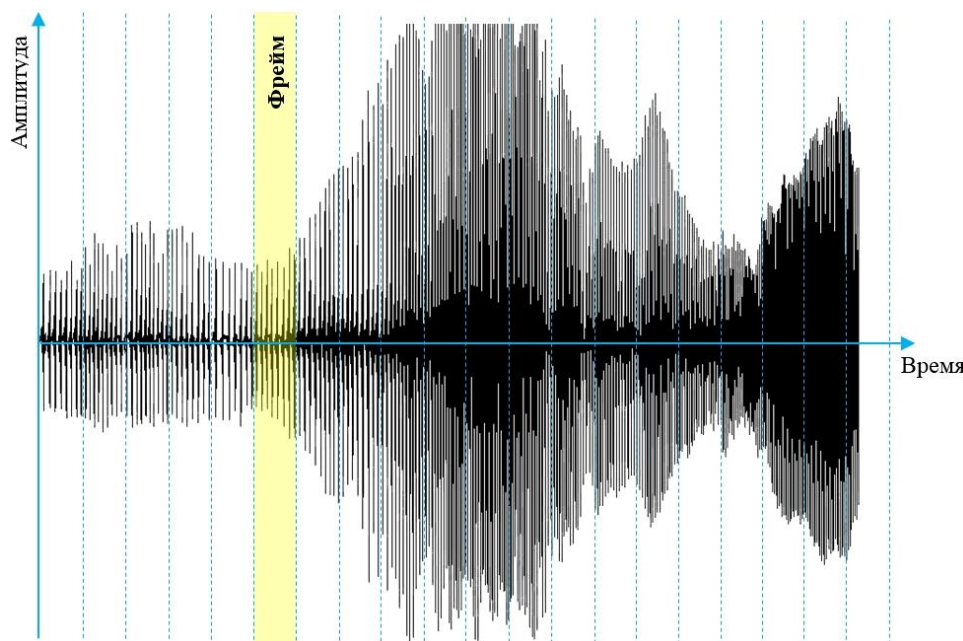


Рисунок 3. Разбиение речевого сигнала на фреймы (разработано авторами)

Рассматривая задачу в данных терминах, можно прийти к общей модели распознавания речи, которая довольно часто используется в наши дни – скрытая марковская модель (СММ) [1].

2. Скрытые марковские модели как основа системы распознавания речи

На основе таких методов и алгоритмов как нейронные сети, модель гауссовых смесей, динамическое программирование и др. разрабатываются современные системы распознавания речи. Также на практике широко применяются скрытые марковские модели, первые упоминания о которых появились еще в 60-х годах прошлого столетия.

Математическое описание скрытых марковских моделей

Для того, чтобы дать определение скрытой марковской модели, необходимо разобрать понятие марковского процесса. *Марковский процесс* – это стохастический процесс (последовательность событий, в которых исход на любом этапе зависит от некоторой вероятности) со следующими особенностями:

- число возможных исходов или состояний известно;
- исход в любой момент времени зависит только от исхода на предыдущем этапе;
- вероятности являются постоянными в течение всего периода времени.

СММ – это статистическая модель, в основе которой лежит марковский процесс, чьи внутренние состояния скрыты от наблюдателя. Считается, что количество состояний системы и вероятности переходов известны.

Скрытая марковская модель определяется кортежем: $\langle S, \Omega, P, \Phi, \pi \rangle$, где:

- $S = \{s_1, s_2, \dots, s_N\}$ – конечный набор N состояний системы (текущее состояние системы в момент времени t обозначается как q_t);

- $\Omega = \{o_1, o_2, \dots, o_M\}$ – конечный набор M возможных символов в наблюдаемой последовательности (размер алфавита наблюдаемой последовательности);
- $P = \{p_{ij}\}$ – матрица вероятностей переходов, где p_{ij} – вероятность перехода системы из состояния s_i в состояние s_j (находится по формуле условной вероятности):

$$p_{ij} = P [q_{t+1} = s_j | q_t = s_i], \text{ где } 1 \leq i, j \leq N;$$

- $\Phi = \{\phi_i(o_k)\}$ – распределение вероятностей символов в наблюдаемой последовательности, где $\phi_i(o_k)$ – вероятность того, что символ o_k будет наблюдаться в системе, находящейся в состоянии s_j ;
- $\pi = \{\pi_i\}, 1 \leq i \leq N$ – распределение вероятностей начального состояния, где π_i – вероятность того, что s_i – начальное состояние системы.

Сокращенно СММ принято обозначать следующим образом:

$$\lambda = (P, \Phi, \pi).$$

На рис. 4 представлен графический пример скрытой марковской модели с пятью состояниями.

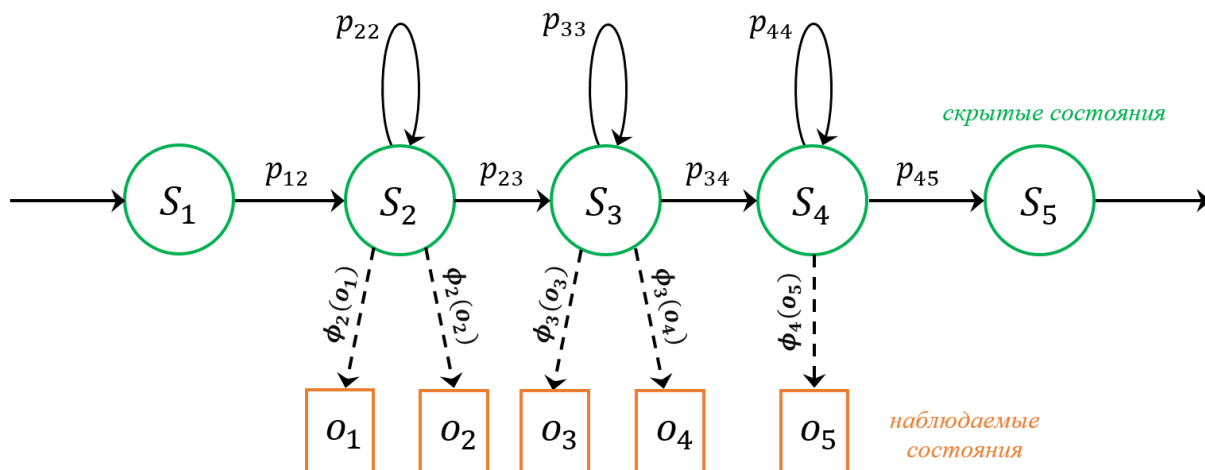


Рисунок 4. Графическое представление СММ (разработано авторами)

Задача распознавания аудио событий в терминах СММ будет выглядеть следующим образом: на вход детектора аудио событий поступает звуковой сигнал, представленный последовательностью:

$$\Omega = \{o_1, o_2, \dots, o_M\},$$

где: o_i – значение параметра звукового сигнала (одно из M), принимаемое детектором в i -тый момент времени. Отрезки времени, в которые детектор снимает эти параметры, являются состояниями $S = \{s_1, s_2, \dots, s_N\}$ модели $\lambda = (P, \Phi, \pi)$. Каждая из этих моделей соответствует различным типам аудио событий, например, определенным словам. Для того, чтобы система смогла выбрать то аудио событие, которое в наибольшей степени соответствует исходному отрезку звукового сигнала (другими словами, распознать слово), необходимо отыскать вероятности появления последовательности $\Omega = \{o_1, o_2, \dots, o_M\}$ для каждой из имеющихся моделей $\lambda = (P, \Phi, \pi)$. Таким образом, имеется набор наблюдаемых состояний (речевой сигнал) и вероятностная модель, сопоставляющая скрытые состояния (фонемы) и наблюдаемые величины. Дальнейшее восстановление наиболее вероятной последовательности скрытых

состояний позволяет выполнить *алгоритм Витерби* (алгоритм поиска наиболее подходящего списка состояний).

Структура системы распознавания речи на СММ

В общем случае структура системы распознавания речи, основанной на скрытых марковских моделях, состоит из следующих элементов (рис. 5):

- акустическая модель – сравнивает принимаемый речевой сигнал с множеством заранее построенных звуковых моделей (СММ), описывающих определенные звуки (буквы, слова) в речи⁷;
- языковая модель – определяет наиболее вероятные словесные последовательности; для разных языков данная модель выглядит по-разному: например, для русского языка необходимо учитывать многочисленные формы одного слова, что усложняет систему, а в английском языке достаточно использовать статистические модели (N-граммы);
- модуль шумоочистки и отделение полезного сигнала;
- декодер – сопоставляет данные процесса распознавания от акустической и языковой моделей и выдает результирующую последовательность распознанного речевого сигнала⁸.

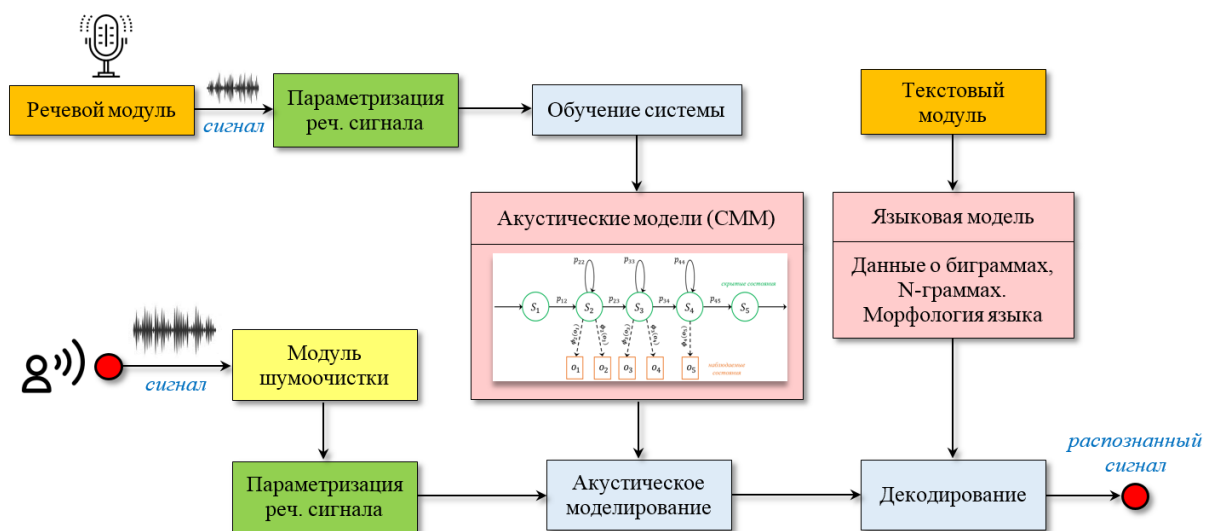


Рисунок 5. Блок-схема системы распознавания речи на основе СММ (разработано авторами)

Стоит отметить, что скрытые марковские модели хоть и являются довольно популярными в области распознавания речи, сейчас многие ученые большие надежды возлагают на нейронные сети [10]. Тем не менее, алгоритм СММ может применяться и в совокупности с другими (в том числе и с нейронными сетями), что позволяет улучшать качество распознавания.

⁷ Распознавание речи. Типичная структура системы распознавания речи / Geektimes <https://geektimes.ru/post/64594/>.

⁸ Распознавание речи – Википедия https://ru.wikipedia.org/wiki/Распознавание_речи.

3. Проект тренажера для обучения иностранным языкам

Классическое учебное пособие по изучению иностранного языка содержит следующие виды учебного материала:

1. Тексты:
 - классические;
 - подготовленные автором пособия;
 - диалоги;
 - составленные студентом.
2. Словари.
3. Тесты:
 - с выбором одного ответа;
 - с выбором нескольких ответов;
 - с выбором слова (термина, ответа на задачу) из списка;
 - тесты упорядочения.
4. Упражнения (разновидности):
 - ответить на вопрос;
 - найти в тексте ответ на задание;
 - выбрать слова по определенному признаку;
 - составить фразу из слов.

На практике печатное учебное пособие развивает только навыки чтения и письма, тогда как для успешной коммуникации с носителями изучаемого языка необходимо также развивать навыки говорения и аудирования, и для этих целей, как правило, обращаются к преподавателям. Но в современном мире, где ценится каждая сэкономленная минута, появляется потребность в обучении из любого места: в общественном транспорте, дома, за городом и т. д. К тому же, не все люди имеют возможность оплачивать платные занятия с репетитором, а тяга к качественному и полноценному изучению языка есть. В таких условиях все чаще прибегают к использованию различных компьютерных программ, мобильных и веб-приложений (в том числе особой популярностью пользуются: онлайн-сервис Duolingo⁹, образовательная платформа Lingualeo¹⁰, программный продукт Rosetta Stone¹¹ и другие). И хотя речевым навыкам в последнее время начинают уделять все больше внимания, систем, способных указать пользователю на ошибки в произношении и дать конкретные рекомендации по его улучшению, практически не существует. В связи с чем возникла идея о создании такого тренажера, который будет охватывать все стороны изучения иностранного языка, включая чтение, аудирование, письмо и говорение, а также позволит выработать программу обучения под конкретного человека с его особенностями речи.

⁹ URL: <https://www.duolingo.com>.

¹⁰ URL: <https://lingualeo.com>.

¹¹ URL: <http://www.rosettastone.eu/>.

Лингвистический тренажер должен поддерживать как набор классических упражнений, развивающих навыки чтения и письма, так и упражнения на правильное звучание речи и аудирование. Виды учебного материала классического типа, предусмотренные в тренажере: работа с текстами, упражнения, выполнение тестов, работа со словарями. Дополнительно вводятся:

- работа с текстом, который набирает учащийся – его воспроизведение, анализ текста на грамматику и синтаксис;
- работа с речью, воспроизводимой учащимся;
- аудиальные тесты и упражнения (табл.).

Таблица

Функционал тренажера (разработано авторами)

Интерфейс пользователя	Функция	Примечание
Речь	Запись в файл	
	Работа онлайн	Работа с речью происходит в анализаторе речи
Текст	Чтение текста	
	Запись текста	Звуковое воспроизведение записанного текста
Упражнения	Классические текстовые упражнения	
	Аудиальные упражнения	
Тесты	Выполнение тестов с последующей проверкой	Отметить правильный вариант ответа
		Ввести слово из списка
		Найти соответствие

В соответствии с этим, в состав тренажера должен входить лингвистический процессор, который будет выполнять функции синтеза и анализа речи, а также процессор, распознающий произносимый пользователем текст.

Аудиальная составляющая языкового тренажера

Часто распознаванию аудиосигналов мешают различные шумы: от аппаратуры и разнообразные окружающие шумы [6]. Также, сложности в распознавании возникают из-за особенностей говорящего, среди которых особое внимание стоит уделять диалекту, дефектам речи и временным изменениям характеристик голоса (как, например, хрипота при болезни). Кроме того, немаловажным в данном вопросе является акцент обучающегося.

Акцент – способ произношения слов, который возникает среди людей в определенном регионе или стране. В то время как язык и произношение определяются различными стандартами, акцент может относиться к отклонению от стандарта. Зачастую понятие акцента путают с *диалектом*. Диалект характеризуется принадлежностью человека к определенной местности или социальному уровню и проявляется, прежде всего, в используемой им лексике и грамматике. Понятие акцента связано только с особенностями произношения.

Для определения акцентированной речи могут быть получены различные характеристики акцента речевого сигнала. Эти характеристики можно разделить на *временные особенности* (длительность озвучивания, средняя длительность слова), *частотные характеристики* и *образцы интонации*, и все они могут быть полезны для обнаружения акцентированной речи. Извлекая информацию об акценте из речевого сигнала и используя эту информацию для независимой системы распознавания, можно улучшить автоматическое распознавание речи.

Однако, для распознавания акцента необходимо иметь обширную базу надиктованных текстов или слов от носителей языка, чтобы в дальнейшем программа могла оценить прогресс обучающегося и высказать предположения о наличии либо отсутствии у него акцента. К сожалению, математическое описание акцента отсутствует, что является большой проблемой в рассматриваемой области.

Таким образом, необходимо разработать систему, учитывающую технологические особенности распознавания речевых сигналов, и которая в дальнейшем будет применяться в качестве обучающего тренажера. При разработке такого тренажера необходимо учесть следующие особенности:

- должен быть произведен выбор наиболее подходящего математического описания работы тренажера (например, системы, основанные на технологии СММ и нейронных сетях, в наши дни довольно популярны; в частности, компания Яндекс использует их в своих разработках в области речевых технологий¹²). При этом рассматриваемые алгоритмы должны упрощать распознавание акцентов и дефектов речи (или, по крайней мере, стремиться к этому);
- необходимо создать обширную базу данных эталонных моделей произношения звуков, букв и слов на иностранном языке (то есть, произвести обучение системы для последующего распознавания);
- результатом работы тренажера должно быть приложение (платформа), посредством которого обучающийся сможет взаимодействовать с обучающей системой. Данный тренажер должен выдавать пользователю информацию о его совпадении/несовпадении с акустическими моделями с рекомендациями по улучшению звучания.

Предполагается, что процесс обучения пользователя навыкам говорения на таком тренажере будет включать в себя следующие стадии:

1. Первичная запись голоса обучаемого с целью выявления уникальных характеристик голоса, дефектов речи, акцента и сопоставления с эталонной моделью произношения.
2. В соответствии с выявленными дефектами и акцентами речи обучаемому будет предложено пройти курс упражнений для постановки речи. На данной стадии система выдает информацию о совпадении или несовпадении у обучающегося произнесенного с услышанным. В случае несовпадения система предлагает пройти упражнение еще раз для достижения правильной постановки звучания иностранных слов. Если система выдала положительный результат о совпадении, считается, что обучающийся успешно выполнил задание.
3. Промежуточное тестирование с целью сопоставления первичного тестирования с полученным результатом. Если система выявила повышение процента совпадения с эталонной моделью произношения, следовательно, курс упражнений остается неизменным, в противном случае система предлагает другие упражнения для постановки речи.

¹² SpeechKit – комплекс речевых технологий Яндекса, который включает распознавание и синтез речи, голосовую активацию и выделение смысловых объектов в произносимом тексте. – Технологии Яндекса <https://tech.yandex.ru/speechkit/>.

4. Итоговое тестирование обучающегося, которое предполагает ознакомление с результатами и дальнейшее закрепление навыков [2].

Отсюда следует, что по итогу обучения на данном тренажере обучающийся улучшит навыки правильного произношения и разговорной речи на иностранном языке, и присутствие учителя уже не будет обязательным. Но, конечно же, это не должно исключать живые беседы с носителями языка, так как владение свободной речью на изучаемом языке является важным аспектом [5, 8].

Заключение

Таким образом, представленный лингвистический тренажер при его реализации позволит подойти к проблеме изучения иностранного языка комплексно, охватывая все стороны обучения, что в дальнейшем послужит прочной основой для успешной межкультурной коммуникации.

ЛИТЕРАТУРА

1. M. Gales and S. Young The Application of Hidden Markov Models in Speech Recognition. Foundations and Trends in Signal Processing, 2007, vol. 1, No. 3, pp. 195-304.
2. Верещагина Е. А., Фоминова Е. Р., Восприятие иностранного языка при помощи аудиального канала восприятия для обучения иностранному языку, 39 международная научно-методическая конференция «Актуальные вопросы качества образования», 8 декабря 2015 г.
3. Лебеденко, Ю. И. Биометрические системы безопасности: пособие / Ю. И. Лебеденко. – Тула: Издательство ТулГУ, 2012. – 159 с.
4. Материалы Региональной научно-практической конференции студентов, аспирантов и молодых учёных по естественным наукам, Владивосток, 15-30 апреля 2016 г. [Электронный ресурс]. – Электрон. дан. – Владивосток: Дальневост. федерал. ун-т, 2016. – URL: https://www.dvfu.ru/schools/school_of_natural_sciences/sciences/theconference/new-page.php.
5. Панина, Т. С. Современные способы активизации обучения: учебное пособие / Т. С. Панина, Л. Н. Вавилова; под ред. Т. С. Паниной. – 4-е изд., стер. – М.: Издательский центр «Академия», 2010. – 176 с.
6. Первушин Е. А. Обзор основных методов распознавания дикторов. Математические структуры и моделирование, 2011, № 24, с. 41-54.
7. Рабинер Л. Р., Шафер Р. В. Цифровая обработка речевых сигналов / Пер. с англ. М.: Радио и связь, 1981. 496 с.
8. Современные образовательные технологии: учебное пособие / под ред. Н. В. Бордовской. – 2-е изд., стер. – М.: КНОРУС, 2011. – 432 с
9. Сорокин В. Н., Вьюгин В. В., Тананыкин А. А. Распознавание личности по голосу: аналитический обзор. Информационные процессы, 2012, Том 12, № 1, с. 1-30.
10. Тампель И. Б. Автоматическое распознавание речи – основные этапы за 50 лет // Научно-технический вестник информационных технологий, механики и оптики. 2015. №6. URL: <http://cyberleninka.ru/article/n/avtomaticheskoe-raspoznvanie-rechi-osnovnyye-etapy-za-50-let>.

Kolesnikova Daria Sergeevna

Far Eastern federal university, Russia, Vladivostok
E-mail: kolesnikova_dse@students.dvfu.ru

Rudnichenko Aleksey Konstantinovich

Far Eastern federal university, Russia, Vladivostok
E-mail: rudnichenko_ak@students.dvfu.ru

Vereshchagina Elena Alexandrovna

Far Eastern federal university, Russia, Vladivostok
E-mail: everesh@mail.ru

Fominova Ekaterina Romanovna

Far Eastern federal university, Russia, Vladivostok
E-mail: katenski.work.94@mail.ru

The use of modern speech recognition technologies to create a linguistic simulator to improve the level of language competence in the field of intercultural communication

Abstract. The topic presented in this article is quite relevant in the modern world. The widespread introduction of recognition technologies allows to improve the skills of people in various fields of knowledge, among which a special place is occupied by teaching foreign languages. This article provides a brief overview of the technology of pattern recognition, the importance of speech recognition in modern conditions, as well as the urgency of creating a linguistic simulator for teaching foreign languages, taking into account the characteristics of the speaker.

In the work, the authors considered the features of voice recognition regardless of the technologies used, as well as existing mathematical algorithms underlying speech biometrics, and in particular, the structure of hidden Markov models and their application in practice for speech synthesis and recognition were studied in more detail.

It was found that the main problem in the recognition of speech signals is the presence of accent and various speech features of the speaker. In this connection, a project of a language simulator was developed that would minimize this problem and help students improve pronunciation. At the heart of the work of this simulator is a complex approach to learning a foreign language, as it combines four important components: writing, listening, reading, speaking.

Keywords: recognition; recognition systems; speech; speech recognition; biometrics; speech biometrics; hidden Markov models; pattern recognition; artificial intelligence; linguistic simulator; teaching; foreign language; accent