

Обзор систем автоматического детектирования речевых нарушений у пациентов с боковым амиотрофическим склерозом (БАС)

Вашкевич Максим Иосифович,

кандидат технических наук, доцент кафедры электронных вычислительных средств Белорусского государственного университета информатики и радиоэлектроники

Рушкевич Юлия Николаевна,

кандидат медицинских наук, доцент, ведущий научный сотрудник неврологического отдела Республиканского научно-практического центра неврологии и нейрохирургии (Республика Беларусь, г. Минск)

Аннотация

В работе выполнен аналитический обзор систем детектирования речевых нарушений у пациентов с боковым амиотрофическим склерозом (БАС). Рассмотрены вопросы применения акустических и артикуляционных параметров в системах детектирования речевых нарушений. Описаны особенности речи и голоса пациентов с БАС, а также современные методики оценки голоса и речи пациента. Уделено внимание особенностям процесса классификации и принятия решений в системах детектирования речевых нарушений. Выполнено обобщающее сравнение различных систем детектирования речевых нарушений при БАС, разработанных за последние пять лет.

Ключевые слова: боковой амиотрофический склероз (БАС), речевой сигнал, классификация.

ВВЕДЕНИЕ

Боковой амиотрофический склероз (БАС) является быстропротекающим нейродегенеративным заболеванием, не поддающимся лечению. Следствием БАС является поражение верхних и нижних мотонейронов головного и спинного мозга [1]. Гибель мотонейронов приводит к постепенной атрофии мышц, отвечающих за движение (ходьбу, дыхание, глотание). Страдает функция речи, проявляется дизартрия, расстройство произносительной организации речи, связанная с нарушением движения органов артикуляции (губ, языка, челюсти, надгортанника).

Выделяют две формы БАС: бульбарную (первые симптомы связаны с прогрессирующим параличом мышц головы и шеи) и спинномозговую (проявляется как ухудшение координации рук или ноги т.д.). В речевом плане, при бульбарной форме БАС происходит постепенная потеря разборчивости речи пациента. Бульбарные симптомы

(т.е. проблемы с глотанием, слюноотделением, речью) наблюдаются в 30% случаев выявления БАС. В остальных случаях они возникают на более поздних стадиях заболевания [2].

В настоящее время не выявлено биомаркеров БАС, что значительно затрудняет процесс диагностирования. В среднем на постановку правильного диагноза уходит 12 месяцев [3]. Данное обстоятельство побуждает многих исследователей к разработке методов анализа речи для раннего диагностирования БАС. Такой подход обоснован описанными в литературе случаями [4], когда признаки расстройства речи возникали за три года до постановки диагноза.

В настоящее время выявление речевых нарушений у пациентов с БАС выполняется врачом диагностом на основе слухового восприятия. К преимуществам метода оценки «на слух» относят его неинвазивный характер и потенциальную чувствительность к продромальной стадии бульбарной формы БАС [5]. К недостаткам метода относят то, что он требует специального обучения и, как любой субъективный метод, подвержен ошибкам. Кроме того, некоторые исследования показывают, что бульбарные моторные отклонения могут проявляться до ощутимых изменений в артикуляции [4, 6].

Второй важной задачей после выявления речевых нарушений, связанных с БАС, является объективная оценка этих нарушений. Объективная оценка важна: 1) как объективизация и подкрепление оценки врача [7]; 2) как средство для отслеживания (мониторинга) развития болезни. В последнем случае понимается не только пассивное отслеживание, но и с целью оценивания эффективности различных подходов к лечению БАС.

На текущем этапе для отслеживания прогрессирования БАС применяется показатель ALSFRS-R (ALS Functional Rating Scale- Revised) [8]. ALSFRS-R представляет собой опросник, его часть, отвечающая за оценку бульбарной подсистемы, основана на определении степени тяжести симптомов в речи, глотании и слюноотделении. Однако данный показатель не отражает специфических особенностей дизартрии и имеет низкую гранулярность для отображения различных аспектов речи пациентов с БАС [9].

Золотым стандартом оценивания бульбарной функции при БАС является измерение скорости произнесения слов (измеряется количеством слов в минуту – сл/мин) [10]. Известно, что скорость речи изменяется раньше, чем происходят изменения в разборчивости речи [2, 4]. При чтении связанных предложений нормой считается скорость речи, превышающая 160 сл/мин. Скорость в 120-160 сл/мин относят к ранним симптомам, при этом могут происходить начальные изменения в разборчивости речи. Скорость меньшая 120 сл/мин относят к поздним симптомам, характерна для периода упадка бульбарной функции [Rong-2015]. Несмотря на то что в целом данный метод позволяет правильно оценить динамику ослабления бульбарной функции, он не в силах точно оценить небольших

изменений в лучшую или худшую сторону, иначе говоря – его шкала является грубой [5].

Таким образом, в настоящее время все больше осознается необходимость в разработке новых методов раннего детектирования признаков бульбарных нарушений у пациентов с БАС, а также методов объективной оценки бульбарной функции. По мнению многих исследователей перспективным направлением в данной области является разработка систем автоматического анализа речи [9-10, 12]. В данной обзорной статье рассматриваются различные подходы к задаче автоматического анализа речи пациентов с БАС, предложенные в последнее время.

1. Бульбарные речевые нарушения при БАС

1.1. Общая характеристика голоса и речи пациентов с БАС

Описывая речь пациентов с БАС, различные исследователи выделяют следующие общие признаки: нарушение временной (темпоральной) структуры речи, наличие неуместных пауз, ухудшение артикуляции (особенно согласных звуков) [13]. Голос пациентов с БАС становится хриплым, жестким (грубым), появляются гиперназальность, моно-тонность, слабость [9].

Такие характеристики голоса, как напряжение, грубость или иногда сдавленность, связаны с поражением верхних моторных нейронов (англ. UMN – upper motor neuron). В то время как хриплость или слабость в голосе, гиперназальность, неточность в артикуляции при отсутствии изменений в скорости речи являются классическими признаками дисфункции нижних моторных нейронов (англ. LMN – lower motor neuron) [5].

Однако указанные признаки характерны для поздней стадии развития БАС и в общем случае не подходят для описания речи пациентов с БАС на ранней стадии.

1.2. Дизартрия

Задача выявления признаков речевых нарушений при БАС часто возникает в контексте более широкой задачи – классификации типа дизартрии [14, 15]. Дизартрия – общее название речевых расстройств, связанных с работой артикуляционных органов, происходящих вследствие неврологических заболеваний. Выделяют пять основных типов дизартрии: вялая (англ. *flaccid*), спастическая (англ. *spastic*), атаксическая (англ. *ataxic*), гипокинетическая (англ. *hypokinetic*) и гиперкинетическая (англ. *hyperkinetic*) [16]. Известно, что БАС приводит к появлению 1) вялой дизартрии, если поражены нижние моторные нейроны; 2) спастической дизартрии, если поражены верхние моторные нейроны и 3) смешанного спастически-вялого типа дизартрии, если поражены как верхние, так и нижние моторные нейроны. Смешанный тип является наиболее характерным для БАС, при нем наблюдаются неточности в артикуляции, гиперназальность, замедленность речи и нарушения просодики [5].

Таким образом, задача детектирования речевых нарушений при БАС осложняется тем, что в речи и голосе могут появляться различные признаки в зависимости от того, какая группа моторных нейронов поражена.

2. Разновидности систем детектирования бульбарных нарушений речи при БАС

В настоящее время проводится широкая исследовательская работа по определению оптимального подхода к детектированию бульбарных нарушений по речи. Исследуются вопросы выбора речевого теста, определения набора речевых признаков, решается вопрос об использовании (или неиспользовании) артикуляционных данных.

2.1. Подсистемный подход (subsystem approach)

В [2] был предложен систематический подход к оценке бульбарных нарушений на основе анализа отдельных подсистем речеобразования (отсюда и название – подсистемный подход). Известно, что в системе речеобразования можно выделить четыре подсистемы: респираторную (дыхательную), фонаторную, артикуляционную и резонаторную (рис. 1).

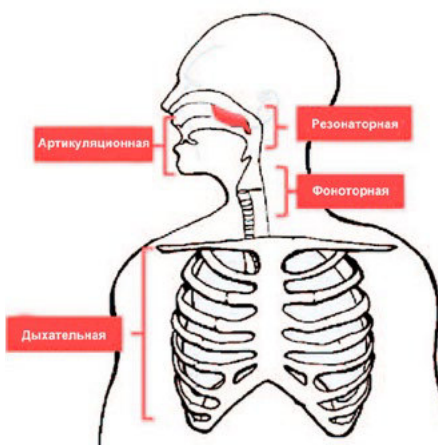


Рис. 1. Четыре подсистемы речеобразования [2]

В [2] для оценивания дыхательной подсистемы использовался процент пауз при чтении специально разработанного предложения. Фонаторная подсистема оценивается параметром отношения энергии шума к энергии гармоник при протяжном произнесении гласной /a/. Работа резонаторной подсистемы оценивалась параметром нозального баланса (англ. nasalalance), т.е. отношением нозальной и оральной компонент энергии в речевом сигнале. Артикуляционная подсистема оценивалась путем измерения пиковой скорости движения губ, языка и челюсти при произнесении тестовых фраз. Развитие данного подхода можно найти в [3, 11]. Несмотря на то что в целом результаты применения подсистемного подхода является многообещающими, трудности в проведении тестов всех подсистем речеобразований затрудняют его применение на практике [17]. Исследователи, работающие в данном направлении, задаются вопросом: можно ли обойтись исключительно анализом речи для выявления признаков бульбарных нарушений? Стремление это объясняется тем, что речевой сигнал легко записать как в больничных, так и в домашних условиях

и вполне возможно, что информации, которая в нем содержится, достаточно для автоматического детектирования БАС [17].

2.2. Речевые задания (выбор речевого теста)

В текущее время ведется активная работа по исследованию различных речевых тестов (заданий) в задаче раннего выявления бульбарных нарушений. Наиболее распространены следующие два типа речевых заданий:

- чтение специально разработанных предложений [2,9-10,12,];
- диадохокинетический (ДДК) тест [17-18].

Для правильного произнесения специально разработанных предложений от тестируемого требуется слаженная работа артикуляционных органов. Например, в [10] использовалось предложение: «Byu Bobby a purru». В некоторых работах использовались предложения общего характера [12, 15, 17, 19-20]. Например, в [12] использовались предложения с простым содержанием, такие как «Мне нужна помощь» или «Позвони мне, когда у тебя получится».

ДДК тест заключается в быстром произнесении слогов (например, «па-/та/-ка» или «бу-/тер/-кап») с максимальной быстротой и точностью на одном дыхании. Данный тест широко применяется в дифференциальной диагностике и для выявления нарушений в мышечном аппарате речи [21].

Иногда в качестве тестового задания используют протяжное произнесение гласных звуков. Данный тест может использоваться как элемент более сложного составного теста [2-3, 14], так и выступать в качестве единственного источника информации для классификации [22]. В литературе встречаются и другие разновидности речевых заданий. Так, в [23] для классификации использовались записи повторения прослушанной фразы и записи спонтанной речи.

2.3. Акустический анализ речи в системах детектирования бульбарных нарушений при БАС

Для описания акустических свойств речи могут быть использованы различные параметры, главной задачей при этом является поиск такой системы параметров, которые бы обеспечивали максимальную точность детектирования бульбарных нарушений.

Одним из самых распространенных способов параметризации речевого сигнала является представление его в виде мел-частотных кепстральных коэффициентов (МЧКК) [24]. Это представление пришло из области систем распознавания речи, но впоследствии хорошо зарекомендовало себя и при построении систем детектирования патологий в речи [25-26]. Вычисление МЧКК учитывает особенности восприятия звука слуховой системой человека. В контексте рассматриваемой задачи МЧКК применялись в работах [18, 23]. Однако в работе [12] указывается, что при расчете МЧКК выполняется декоррелирующее преобразование, в результате которого может потеряться важная информация, необходимая для различения здоровой речи от речи пациента с БАС. Поэтому в [12] применялось представление речевого сигнала в виде энергетического спектра.

В [27] для анализа выбирались только фрагменты речевого сигнала, содержащие гласные звуки /а/ и /и/ для которых анализировались формы спектральных огибающих, а также положение формантных частот. В работе [22] в качестве речевого задания выступало протяжное произнесение звука /а/, а в качестве параметров брались джиттер (кратковременное произвольное отклонение частоты колебания голосовых складок) и шиммер (кратковременные отклонения в амплитуде звуковой волны), а также параметр паталогичности вибрато (т.е. наличие колебаний в частоте основного тона с частотой от 9 до 14 Гц, для здоровых людей частота вибрато находится в диапазоне от 5 до 8 Гц).

В работах [9, 17] предлагается автоматизировать поиск подходящих речевых параметров, используя свободно распространяемый программный пакет OpenSmile [28], позволяющий извлекать из речи порядка 6 тысяч параметров. Для отбора релевантных параметров в [9] использовался корреляционный анализ каждого отдельного параметра x с меткой класса $y \in \{\text{здоровый, больной}\}$. Для последующего рассмотрения отбирались только признаки, имевшие уровень ложных срабатываний (англ. false discovery rate) $q < 0.05$. Среди самых информативных признаков в [9] оказались МЧКК, а также параметры модуляционного спектра речи RASTA (англ. RelAtive SPectrA) [24, 29]. Недостатком подхода, предложенного в [9], является то, что каждый из параметров речи анализировался независимо от остальных. Известно, что два отдельных параметра x_1 и x_2 могут не иметь сильной корреляцией с меткой класса y , однако их объединение может порождать признак с сильной корреляционной связью [30].

2.4. Анализ артикуляционных параметров речи в системах детектирования бульбарных нарушений

Часто наряду с акустическим анализом речи выполняется и анализ артикуляционной активности. Для сбора данных о движении артикуляционных органов на них крепится система датчиков.

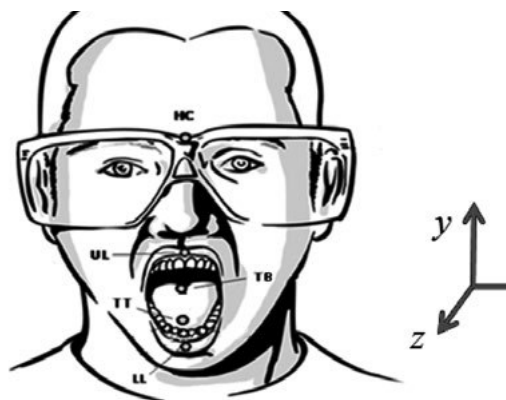


Рис. 2. Положение артикуляционных датчиков [17]. Обозначения: TT (tongue tip) – кончик языка, TB (tongue body) – середина языка, TD (tongue dorsum) – спинка языка (на рисунке не показан), LL (lower lip) – нижняя губа, UL (upper lip) – верхняя губа

Чаще датчики крепятся на язык, губы (рис. 2) и иногда челюсть [2]. В большинстве работ артикуляционные данные используются в качестве дополнения к акустическим признакам [17-18, 23]. Как правило, исследователи не указывают, какая именно информация в артикуляционных данных является наиболее существенной. Исключением является [23], где было показано, что наиболее важным является параметр ускорения движения кончика языка. Причем исключительная важность данного параметра не зависела от типа речевого задания.

3. Особенности построения систем классификации бульбарных нарушений

3.1. Подготовка информационных признаков

Известно, что успешное решение задачи классификации зависит от качества используемых информационных признаков в рассматриваемых нами системах. Можно выделить следующие способы получения информационных признаков (рис. 3).

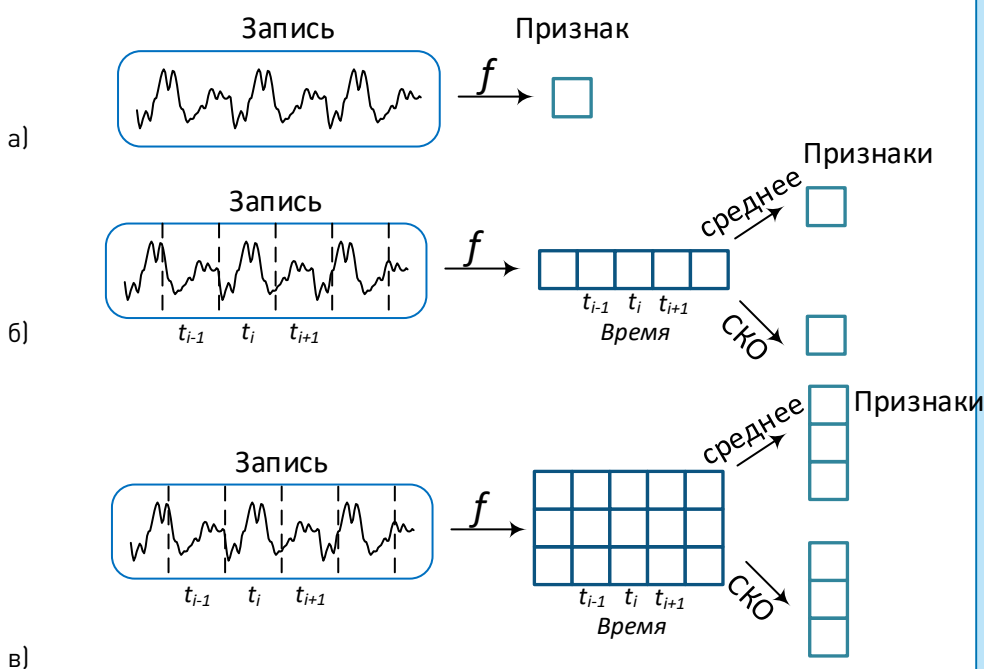


Рис. 3. Способы формирования информационных признаков

В первом случае (рис. 3, а) каждой записи речевого сигнала ставится в соответствие один признак (например, значение джиттера, шиммера или индекса патологичности вибрато). Во втором случае (рис. 3, б) сигнал разбивается на кратковременные сегменты, измеряемый параметр вычисляется для каждого сегмента, после чего выполняется усреднение, расчет среднеквадратического отклонения (СКО) либо любой другой статистики. Примером признаков второго типа может служить вычисление статистики частоты основного тона. Третий вариант подготовки информационных признаков похож на второй, только для каждого сегмента сигнала рассчитывается вектор параметров, который затем усредняется. Примером

в данном случае может служить МЧКК, которые вычисляются для сегментов длительности 70 мс, а затем усредняются [17].

В некоторых случаях длинная запись разбивается на сегменты заданной длины (например, 0,8 с), и параметры вычисляются и усредняются для этих сегментов [23]. В этом случае полученные признаки называются надсегментными (англ. supra-segmental). В случае использования надсегментных признаков из одной записи можно получить несколько характеристических векторов. Использование надсегментных признаков мотивируется тем, что особенности речи пациентов с БАС относятся к категории пара-лингвистической информации, которая содержится в сегментах значительной длины [23].

Следует отметить, что в случае использования в качестве классификатора свёрточную нейронную сеть (англ. CNN – convolutional neural network) операция усреднения признаков может не выполняться, а на вход сети подается весь набор полученных параметров без усреднения (рис. 3, в) [23].

3.2. Организация процесса классификации

Чаще всего в качестве классификатора в системах детектирования бульбарных нарушений используется метод опорных векторов (англ. SVM – support vector machine) [9-10, 17, 23]. Это объясняется тем, что SVM имеет хорошее математическое обоснование, а также для него разработаны свободно распространяемые программные пакеты с его реализацией. Вторым популярным методом является использование глубоких нейронных сетей (англ. DNN – deep neuron network) [17, 23]. Данный выбор объясняется прежде всего успехом DNN в смежных областях – распознавании и синтезе речи. Реже используется подход на основе линейного дискриминантного анализа (англ. LDA – linear discriminant analysis) [19, 31] и метода k ближайших соседей (англ. kNN – k nearest neighbors) [22].

Чаще всего для оценки точности классификатора используют следующие характеристики: точность (Acc), чувствительность (Sens), специфичность (Spec), F-мера и средняя полнота (R_{ave}). Ниже приводятся выражения для их вычисления [30]:

$$\text{Acc} = \frac{TP + TN}{TP + FP + TN + FN}, \quad \text{Sens} = \frac{TP}{TP + FN},$$
$$\text{Spec} = \frac{TN}{TN + FP}$$
$$R_{ave} = \frac{1}{2}(\text{Sens} + \text{Spec}), \quad F_{\text{мера}} = 2 \frac{\text{Sens} \cdot \text{Spec}}{\text{Sens} + \text{Spec}}$$

где TP , TN , FP , FN – истинно положительные, истинно отрицательные, ложно положительных, ложно отрицательные результаты классификации (под положительным результатом понимают верную постановку диагноза

БАС). Точность характеризует частоту правильных решений, выносимых классификатором. Однако данный параметр не даёт полного представления о качестве работы классификатора, особенно если в речевой базе классы здоровых голосов и голосов с патологией не сбалансированы. Параметр чувствительности показывает способность классификатора детектировать патологию, если она есть. Специфичность характеризует способность классификатора определять отсутствие патологии, когда она действительно отсутствует. Средняя полнота – показывает общую способность классификатора относить голос к правильному классу. F-мера показывает среднее гармоническое значение точности и специфичности.

Важным аспектом процесса классификации является этап принятия решения. В простейшем случае решение совпадает с той меткой, которую записи присвоил классификатор. Однако в случае использования надсегментных признаков для одной записи речи имеется несколько характеристических векторов, каждый из которых подается на классификатор, а решение принимается на основе мажоритарного голосования. Если для одного лица имеется несколько записей его речи, то классифицируется каждая из них, а окончательное решение затем принимается мажоритарным голосованием [12, 23]. В следующей таблице приведены сводные данные по системам классификации, которые были опубликованы в последние годы.

Таблица 1

Сравнение систем детектирование бульбарных нарушений при БАС

Источник	Тип классификатора	Число признаков	Речевая база (число больных БАС/ число здоровых)	Точность (Acc), %	Чувствительность (Sens), %	Специфичность (Spec), %	Средняя полнота (R_{avg}), %	F-мера, %
Системы не использующие артикуляционные данные								
[18]	XGBoost	–	34/49	85	92	80	86	85
[17]	SVM	–	11/11	65	58	68	63	61
[17]	DNN	–	11/11	92*	82*	96*	89*	88*
[23]	SVM	39	8/8	–	–	–	–	89**
[23]	DNN	39	8/8	–	–	–	–	90**
[9]	SVM	1	мужчины 26/40	79	78	76	77	77
[9]	SVM	–	женщины 30/27	83	86	78	82	82
[12]	CNN	120	13/13	76	72	81	76	76
[22]	LDA	4	39/15	91	87	92	90	89
[22]	kNN	3	39/15	92	76	98	87	86

Продолжение табл. 1

Источник	Тип классификатора	Число признаков	Речевая база (число больных БАС/ число здоровых)	Точность (Acc), %	Чувствительность (Sens), %	Специфичность (Spec), %	Средняя полнота (R_{over}), %	F-мера, %
Системы использующие артикуляционные данные								
[18]	XGBoost	-	34/49	90	94	85	90	89
[17]	SVM	-	11/11	81	82	81	81	81
[17]	DNN	-	11/11	97*	89*	100*	94*	94*
[23]	SVM	6	8/8	-	-	-	-	100**
[23]	DNN	6	8/8	-	-	-	-	95**

* В исследовании использовалась не сбалансированная по возрасту и полу речевая база.

** Все больные БАС имели существенные бульбарные нарушения.

Формально лучшим результатом является классификатор на основе DNN [17] с применением артикуляционных данных, однако, как указывают авторы, они использовали несбалансированную по возрасту и полу базу, поэтому оценки точности являются смещенными. Результаты, полученные в [23], также весьма высоки, однако используемая авторами речевая база содержала мало голосов, кроме того, все пациента с БАС на момент исследования имели существенные бульбарные нарушения. Во всех остальных случаях средняя полнота результатов классификации не превышает 90%, что оставляет большой задел для совершенствования подходов к детектированию бульбарных нарушений при БАС.

Заключение

В работе выполнен аналитический обзор подходов к построению систем детектирования и мониторинга бульбарных нарушений у пациентов с БАС. Показано, что в настоящее время наиболее активно ведутся исследования в направлении использования акустического анализа речевого сигнала для детектирования бульбарных нарушений у пациентов с БАС. Анализ существующих систем показал, что в среднем их точность составляет 85-90%, что препятствует применять их на практике. Можно сделать вывод о том, что в данной области сохраняется ряд открытых и нерешенных вопросов, а также о необходимости продолжать работу в данном направлении.

Список литературы

1. *Kiernan, M.C.* Amyotrophic lateral sclerosis / *M.C. Kiernan, S. Vucic, B.C. Cheah, M.R. Turner, A. Eisen, O. Hardiman, J.R. Burrell, M.C. Zoing* // *Lancet*. — 2011. — vol.377 (issue 9769). — P. 942–955.
2. *Green, J.R.* Bulbar and speech motor assessment in ALS: challenges and future directions / *J.R. Green, Y. Yunusova, M.S. Kuruvilla, J. Wang, G.L. Pattee, L. Synhorst, L. Zinman, J.D. Berry* // *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. — 2013. — № 14. — P. 494–500.
3. *Yunusova Y.* Detection of bulbar ALS using a comprehensive speech assessment battery/ *Y. Yunusova, J. S. Rosenthal, J. R. Green, S. Shellikeri, P. Rong, J. Wang, L. H. Zinman* // in *Proc. of the International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications*, 2013, P. 217–220.
4. *Yorkston, K. M.* Speech deterioration in amyotrophic lateral sclerosis: Implications for the timing of intervention / *K. M. Yorkston, E. A. Strand, R. Miller, A. Hillel, and K. Smith* // *Journal of medical speech-language pathology*. — 1993 — vol. 1, no. 1. — pp. 35–46.
5. *Yunusova, Y.* Clinical Measures of Bulbar Dysfunction in ALS / *Y. Yunusova, E.K. Plowman, J.R. Green, C. Barnett, P. Bede* // *Frontiers in Neurology*. — 2019. — vol. 10. — P. 1–11.
6. *Mefferd, A.M.* A novel fixed-target task to determine articulatory speed constraints in persons with amyotrophic lateral sclerosis/ *A.M. Mefferd, J.R. Green, G. Pattee* // *Journal of communication disorders*. — 2012. — no.1, vol. 45. — P. 35–45.
7. *Laaridh, I.* Automatic Prediction of Speech Evaluation Metrics for Dysarthric Speech/ *I. Laaridh, W.B. Kheder, C. Fredouille, C. Meunier* // *Proc. Interspeech 2017*, 1834–1838.
8. *Cedarbaum, J. M* The ALSFRS-R: a revised als functional rating scale that incorporates assessments of respiratory function/ *Jesse M Cedarbaum, Nancy Stambler, Errol Malta, Cynthia Fuller, Dana Hilt, Barbara Thurmond, Arline Nakanishi* // *Journal of the neurological sciences*. — 1999. — vol. 169, no. 1. — P. 13–21.
9. *Norel, R.* Detection of amyotrophic lateral sclerosis (ALS) via acoustic analysis / *R. Norel, M. Pietrowicz, C. Agurto, S. Rishoni, G. Cecchi* // *Proc. Of Interspeech 2018*. — P. 377–381.
10. *Bandini, A.* Classification of Bulbar ALS from Kinematic Features of the Jaw and Lips: Towards Computer-Mediated Assessment / *A. Bandini, J.R. Green, L. Zinman, Y. Yunusova* // *Proc. Interspeech 2017* — P.1819–1823.
11. *Rong, P.* Predicting early bulbar decline in amyotrophic lateral sclerosis: a speech subsystem approach/ *P. Rong, Y. Yunusova, J. Wang, J. R. Green* // *Behavioural Neurology*. — 2015. — id.183021. — P. 1–11.
12. *An, K.* Automatic early detection of amyotrophic lateral sclerosis from intelligible speech using convolutional neural networks/ *K. An, M. Kim, K. Teplansky, J. Green, T. Campbell, Y. Yunusova, D. Heitzman, J. Wang* // *Proceedings of Interspeech 2018* — P. 1913–1917.
13. *Caruso, A. J., Burton, E. K.* Temporal Acoustic Measures of Dysarthria Associated with Amyotrophic Lateral Sclerosis // *Journal of Speech, Language, and Hearing Research*. — 1987. — №1, vol.30. — P. 80–87.
14. *Guerra, C., Lovey, D.* A modern approach to dysarthria classification. // *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEMBS)*. — 2003. — vol.~3. — P. 2257–2260.
15. *Liss, J. M.* Discriminating dysarthria type from envelope modulation spectra / *J.M. Liss, S. LeGendre, A. J. Lotto* // *Journal of Speech, Language, and Hearing Research*. — 2011. — № 5 (53). — P. 1246–1255.

16. Tomik, B. Dysarthria in amyotrophic lateral sclerosis: A review / B. Tomik, R.J. Guiloff // Amyotrophic Lateral Sclerosis – 2010. – vol. 11, № 1-2. – P. 4-15.
17. Wang, J. Towards Automatic Detection of Amyotrophic Lateral Sclerosis from Speech Acoustic and Articulatory Samples/ J. Wang, P.V. Kothalkar, B. Cao, D. Heitzman// Proc. Interspeech 2016 — P. 1195-1199.
18. Spangler, T. Fractal features for automatic detection of dysarthria / T. Spangler, N. V. Vinodchandran, A. Samal, J. R. Green // IEEE EMBS International Conference on Biomedical Health Informatics (BHI), Orlando, USA, Feb. 16-19, 2017. — Orlando, 2017 — pp. 437-440.
19. Vashkevich, M. Features extraction for the automatic detection of ALS disease from acoustic speech signals/ M. Vashkevich, E. Azarov, A. Petrovsky, Yu. Rushkevich // Proceedings of inter. conf. Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), Poznan, Poland, 19-21 Sept. 2018. — P. 321-326.
20. Гвоздович А.Д. Детектирование бульбарных нарушений при боковом амиотрофическом склерозе на основе анализа речевого сигнала./ Гвоздович А.Д., Рушкевич Ю.Н., Вашкевич М.И.// Доклады БГУИР.— 2018. — № 6 (116) — С. 52-58.
21. Wang, Y. T. Analysis of diadochokinesis in ataxic dysarthria using the motor speech profile programTM/ Y.-T. Wang, R. D. Kent, J. R. Duffy, and J. E. Thomas// Folia phoniatica et logopaedica: official organ of the International Association of Logopedics and Phoniatics (IALP). — 2009. — vol. 61, no. 1. — P. 1-11.
22. Vashkevich, M. Bulbar ALS detection based on analysis of voice perturbation and vibrato / M. Vashkevich, A. Petrovsky, Yu. Rushkevich // Proceedings of inter. conf. Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), Poznan, Poland, 18-20 Sept. 2019. — P. 267-272.
23. Illa, A. Comparison of speech tasks for automatic classification of patients with amyotrophic lateral sclerosis and healthy subjects/ Illa, A., Patel, D., Yaminiy, B., Meera, S., Shivashankar, N., Veeramaniz, P.K., Vengalilz, S., K.~Polavarapuz, S.N., Naliniz, A., Ghosh, P.K.// Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). — 2018. — P. 6014-6018.
24. Рылов А.С. Анализ речи в распознающих системах. Мн.: Бестринт, 2003. 264 с.
25. Fraile, R. Automatic detection of laryngeal pathologies in records of sustained vowels by means of mel-frequency cepstral coefficient parameters and differentiation of patients by sex / Fraile R, Sáenz-Lechón N., Godino-Llorente J., et al. // Folia Phoniatica Logopaedica. — 2009. — vol. 61. — P. 146-152.
26. Лысак, А.П. Акустический анализ голоса в норме и патологии // Речевые технологии. — 2012. — №4. — С. 68-82.
27. Gvozдович, A. Detection bulbar dysfunction in ALS patients using acoustic analysis of vowels extracted from continuous speech / A. Gvozдович, M. Vashkevich, Yu. Rushkevich, A. Petrovsky // Proceedings of inter. conf. Pattern recognition and information processing (PRIP), Minsk, Belarus, 21-23 May 2019. — P. 120-124.
28. Eyben, F. Recent developments in OpenSmile, the munich open-source multimedia feature extractor/ F. Eyben, F. Weninger, F. Gross, B. Schuller // in Proc. of the ACM International Conference on Multimedia. — 2013. — P. 835-838.

29. Козлачков С.Б. Исследование свойств модуляционных параметров речевых сигналов / С.Б. Козлачков, А.М. Бонч-Бруевич, С.В. Дворянkin // Речевые технологии. — 2013. — №3-4. — С. 135–142.
30. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / пер. с англ. А.А. Слинкина. М.: ДМК Пресс, 2015. 400 с.
31. Вашкевич М. И. Акустический анализ голоса для выявления речевых нарушений при боковом амиотрофическом склерозе / М.И. Вашкевич, А.Д. Гвоздович, Ю.Н. Рушкевич, А.А. Петровский // Доклады БГУИР. - 2018. - № 7 (117). - С. 64 - 68.

REVIEW OF THE SYSTEMS FOR AUTOMATIC DETECTION OF AMYOTROPHIC LATERAL SCLEROSIS (ALS) FROM SPEECH

Vashkevich M. I.,

Candidate of Technical Sciences, Associate Professor of the Department of Electronic Computing Facilities of the Belarusian State University of Informatics and Radioelectronics

Rushkevich Y. N.,

Candidate of Medical Sciences, Associate Professor, Leading Researcher of the Neurological Department of the Republican Scientific and Practical Center of Neurology and Neurosurgery (Republic of Belarus, Minsk)

Abstract.

An analytical review of the systems for detecting amyotrophic lateral sclerosis (ALS) from speech is presented. The issues of the use of acoustic and articulation parameters for ALS speech detection are considered. The features of speech and voice of patients with ALS, as well as current methods of assessing the ALS voice and speech are described. Attention is paid to the specifics of the classification and decision-making process in ALS speech detection systems. Comprehensive comparisons of various systems for the ALS speech detection published over the past 5 years have been made.

Key words: amyotrophic lateral sclerosis (ALS), speech signal, classification