

УДК 004.855.5

DOI: 10.25559/SITITO.15.201901.164-171

## Обзор методов распознавания внутреннего проговаривания на основе данных электроэнцефалограммы

Ю. Ю. Гавриленко<sup>1</sup>, Д. Ф. Саада<sup>1</sup>, А. О. Шевченко<sup>2</sup>, Е. А. Ильюшин<sup>2\*</sup>

Московский государственный университет имени М.В. Ломоносова, г. Севастополь, Россия

Московский государственный университет имени М.В. Ломоносова, г. Москва, Россия

\* eugene.ilyushin@gmail.com

### Аннотация

В статье представлен обзор исследований о разработке интерфейсов мозг-компьютер, распознающих внутреннее проговаривание на основании данных электроэнцефалограммы (ЭЭГ). Интерфейс мозг-компьютер - система, которая поддерживает взаимодействие мозга и электронного устройства, а в частности, интерфейс безмолвного доступа - система, позволяющая распознавать внутреннее проговаривание, может помочь людям, имеющим проблемы с воспроизведением речи. Эти проблемы могут быть вызваны особенностями нейродегенеративных заболеваний. Нейроинтерфейсы, способные моделировать речь, уже существуют, например, Speller P300, однако, не смотря на свою эффективность, подобное устройство не позволяет достичь высокой скорости генерации речи. Интересной является разработка ИМК, основывающегося на внутреннем проговаривании - процессе, в котором человек представляет звучание слов или фонем без самого произнесения звуков. Известно, что восстановление внутреннего проговаривания на основе данных биоэлектрической активности мозга возможно, так как в процессе внутреннего проговаривания генерируется активность артикуляторов, нервных путей между ними, а также самого головного мозга. ЭЭГ - метод электромагнитного исследования мозга, который регистрирует сигналы, возникающие при биоэлектрической активности нейронов. Технология ЭЭГ является удобной для исследования мозговых сигналов в процессе внутреннего проговаривания, поскольку она не является инвазивной и предоставляет отличное временное разрешение вплоть до одной тысячной секунды, хоть данные ЭЭГ и обладают большим количеством артефактов и не предоставляет высокого пространственного разрешения. Рассмотренные в статье исследования были разделены по типу языковой единицы, распознаваемой в течение эксперимента - слова, фонемы и гласные буквы. Исходя из результатов, полученных авторами обзораемых статей, классификация слов практически не превосходит тех точностей, которые можно было бы получить случайно, однако классификация фонем и гласных обладает достаточно высокой точностью, что говорит о перспективности исследования распознавания внутреннего проговаривания фонем и гласных.

**Ключевые слова:** интерфейс мозг-компьютер, ЭЭГ, внутреннее проговаривание, классификация.

**Для цитирования:** Гавриленко Ю. Ю., Саада Д. Ф., Шевченко А. О., Ильюшин Е. А. Обзор методов распознавания внутреннего проговаривания на основе данных электроэнцефалограммы // Современные информационные технологии и ИТ-образование. 2019. Т. 15, № 1. С. 164-171. DOI: 10.25559/SITITO.15.201901.164-171

© Гавриленко Ю.Ю., Саада Д.Ф., Шевченко А.О., Ильюшин Е.А., 2019



BY

Контент доступен под лицензией Creative Commons Attribution 4.0 License.  
The content is available under Creative Commons Attribution 4.0 License.



## A Review on Internal Pronouncing Recognition Methods Based on Electroencephalogram Data

Yu. Yu. Gavrilenko<sup>1</sup>, D. F. Saada<sup>1</sup>, A. O. Shevchenko<sup>2</sup>, E. A. Ilyushin<sup>2\*</sup>

<sup>1</sup> Lomonosov MSU Branch in Sevastopol, Sevastopol, Russia

<sup>2</sup> Lomonosov Moscow State University, Moscow, Russia

\* eugene.ilyushin@gmail.com

### Abstract

This article presents a review of recent research on brain-computer interfaces (BCI) development regarding silent speech recognition based on electroencephalogram data (EEG). BCI is a system, which supports the interaction of a brain and an electronic device. Silent speech interface is a system, which recognizes internal pronouncing. This device can help people, who have problems with speech production, which may be caused by neurodegenerative disorders. There have already been developed BCIs that can simulate speech process, e.g., a Speller P300, but, despite this, the typing speed of this method is still very slow. The development of the BCI, which is based on internal pronouncing (the process of imagining one's pronouncement of phonemes or words without intending to actually pronounce them), would be very useful. It is known that the internal pronouncement can be recovered from EEG data, because the neuro-activity from articulators, neural paths among them and the brain itself is generated during the processes of internal pronouncement. EEG is a method of electromagnetic brain analysis, which registers the signals that appear during the bioelectrical neural activity. It is one of the most convenient and easy technologies for brain signals analysis including ones that are produced during the silent speech, because it is not invasive and grants good temporal resolution down to one millisecond, but still it produces a big amount of artefacts. The articles presented in this review were divided by the type of morphological unit used in the experiment: words, phonemes and vowels recognition.

**Keywords:** BCI, EEG, internal pronouncing, classification.

**For citation:** Gavrilenko Yu.Yu., Saada D.F., Shevchenko A.O., Ilyushin E.A. A Review on Internal Pronouncing Recognition Methods Based on Electroencephalogram Data // *Sovremennye informacionnye tehnologii i IT-obrazovanie* = Modern Information Technologies and IT-Education. 2019; 15(1):164-171. DOI: 10.25559/SITITO.15.201901.164-171



## Введение

Речь - стандартный способ взаимодействия людей в повседневной жизни. Однако, существуют ситуации, когда человек не может воспроизводить речь из-за особенностей заболевания, либо же общение посредством воспроизведения звуков является нежелательным, например, в опасной военной деятельности. В таких ситуациях необходимы нейроинтерфейсы, поддерживающие механизм общения без воспроизведения каких-либо звуков.

Нейроинтерфейс (интерфейс мозг-компьютер) - технология, целью которой является преобразование мысли или намерения пользователя в сигнал, с помощью которого можно управлять компьютером, креслом-каталкой или же протезом. Так называемые интерфейсы безмолвного доступа (silent speech interface) - системы, позволяющие распознавать речь без ее произнесения. Известно, что регистрируя данные, поступающие с элементов производства речи в человеке - артикуляторов (губ, языка, гортани и неба), нервных путей между ними, самого головного мозга - интерфейс безмолвного доступа может восстанавливать саму речь [14].

В данном исследовании рассматриваются результаты, полученные в работах по разработке интерфейсов безмолвного доступа, основанных на внутреннем проговаривании. Внутреннее проговаривание - возможность "слышать" свою собственную речь у себя в голове без произнесения звуков. Известно, что расшифровка внутренней речи возможна, поскольку при превосхождении предполагаемой речи активизируются нейроны и генерируются определенные сигналы, которые можно зарегистрировать с помощью таких технологий, как электроэнцефалограмма, электрокортикограмма, функциональная магнитно-резонансная томография. Наибольший интерес вызывает ЭЭГ как неинвазивный метод отведения потенциалов. Процесс анализа внутреннего проговаривания субъекта состоит из регистрации сигналов головного мозга и дальнейшей обработки полученных сигналов с помощью программного обеспечения для того, чтобы "расшифровать" полученные сигналы. В данной статье рассмотрены исследования эффективности использования различных методов для интерпретации этих сигналов.

## Применение

Разработка интерфейсов безмолвного доступа все еще в экспериментальной стадии и множество возможных приложений этих устройств делает разработку всё более перспективной.

В настоящее время люди, страдающие от болезней, приводящих к параличу (например, такой болезнью может быть боковой амиотрофический склероз), для общения пользуются интерфейсами, способными считывать по одной букве в слове. Например, существуют спеллеры - нейроинтерфейсы, основанные на событийном потенциале P300 (электро-физиологической реакции человека на стимул, которая возникает через 300 миллисекунд после предъявления стимула). Нейроинтерфейс позволяет писать слово по буквам, фокусируя внимание на определенном символе, показываемом на дисплее в матрице, в то время как столбцы и колонки случайным образом подсвечиваются. Когда подсвечивается столбец или строка, содержащие необходимый символ, вызывается событийный потенциал и, регистрируя этот потенциал, нейроинтерфейс

способен распознать символ, который задумал испытуемый [18].

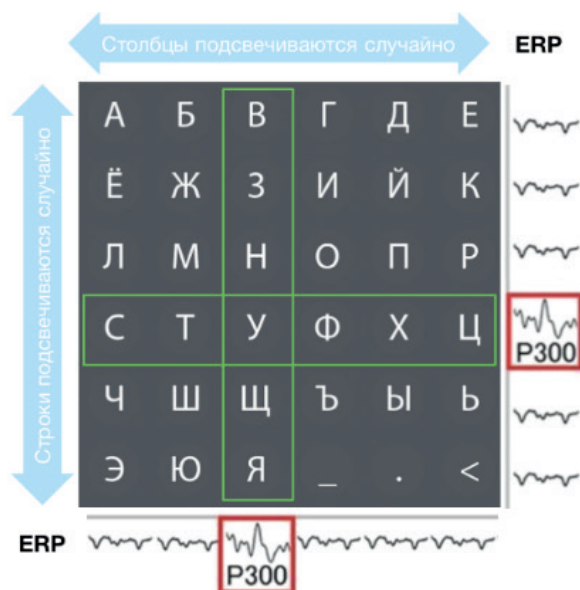


Рис. 1. Speller P300

Fig. 1. Speller P300

Такие нейроинтерфейсы показали свою эффективность, однако, они достаточно медленные и не могут быть эффективно и реалистично применены к повседневной речи. Ко всему прочему, для использования таких устройств необходимо дополнительное изучение механизма работы таких устройств: пользователю нужно учиться пользоваться такими интерфейсами. Интерфейс, основанный на внутреннем проговаривании, не требовал бы от пользователя приобретения новых навыков. Все, что нужно было бы сделать для использования устройства - подумать о нужном слове.

Помимо применения в качестве помощи больным, такие устройства также могут быть полезны в военной индустрии, когда любые другие виды коммуникации могут быть опасными и необходимо передавать сообщения на расстоянии без использования визуальных или звуковых методов. Например, в 2008 году Управление перспективных исследовательских проектов Министерства обороны США DARPA выделило 4 миллиона долларов на разработку интерфейса для синтетической телепатии. Согласно этому проекту, необходимо было разработать устройство, которое позволяло бы совершать общение на поле боя без использования вокализованной речи используя методы распознавания паттернов ЭЭГ [15].

Также, нейроинтерфейсы распознавания внутренней речи могут быть применены в индустрии развлечений, например, в устройствах виртуальной реальности.

## Методы регистрации сигналов головного мозга

В исследовании мозговых сигналов применяются различные технологии, как инвазивные, так и неинвазивные, и каждая из этих технологий обладает своими преимуществами.

В некоторых исследованиях [11, 19] используется метод элек-



трокортикографии - это инвазивный метод отведения потенциалов, в котором электроды накладываются на кору головного мозга. Данный метод позволяет получать точную информацию с хорошим временным и пространственным разрешением и с наименьшим отношением сигнал-шум, что является очень удобным условием для обработки данных без необходимости усреднять полученные мозговые сигналы. Однако инвазивность означает необходимость проведения хирургической операции, что является большим препятствием для свободного проведения исследований. В связи с этим предпочтение отдают, конечно, неинвазивным методам, хоть и эти методы обладают некоторыми недостатками.

Функциональная магнитно-резонансная томография, которая заключается в измерении уровня изменений в кровотоке в эпизоды нейронной активности, является неинвазивным методом. Это связано с тем, что во время активности определенной зоны мозга, приток крови к ней увеличивается. В исследовании [10] с использованием фМРТ пытались распознать бинарные ответы (Да/Нет) на вопрос "Есть ли у вас братья или сестры?". Результаты показали, что 90% ответов были распознаны корректно, что свидетельствует о возможности эффективного использования фМРТ в схожих исследованиях. Недостатки фМРТ заключаются в громоздкости самого томографа а также в медленной скорости ответа. Однако хорошее пространственное разрешение позволяет проводить эксперименты с целью понять механизмы, связанные с внутренней речью.

Среди неинвазивных устройств наиболее оптимальным по простоте использования, размерам и разрешению считается ЭЭГ (электроэнцефалограмма). Устройства для регистрации сигналов ЭЭГ достаточно просты в использовании, а электроды устройства накладываются на поверхность скальпа. Данный метод обладает очень хорошим временным разрешением. В разделе 6 будут подробнее рассмотрены проведенные исследования об интерфейсах распознавания внутреннего проговаривания, основанные на данных ЭЭГ.

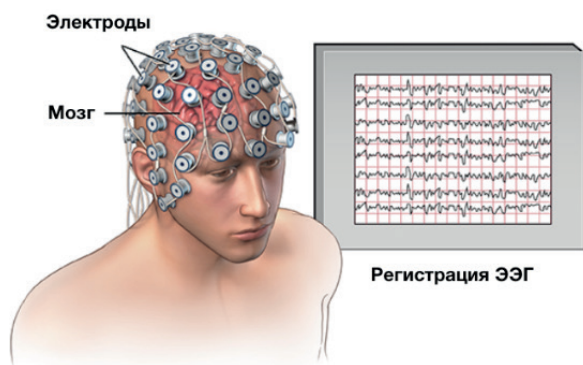


Рис. 2. Электроэнцефалограмма<sup>1</sup>  
Fig. 2. Electroencephalogram

## Внутреннее проговаривание

Внутреннее проговаривание – процесс представления какого-либо звука или слова без его произнесения и без сопутствующего

движения губ или языка [12]. В исследовании [14] впервые было продемонстрировано, что существует возможность распознавать изолированные слова на основании записей ЭЭГ и МЭГ (магнитной энцефалографии).

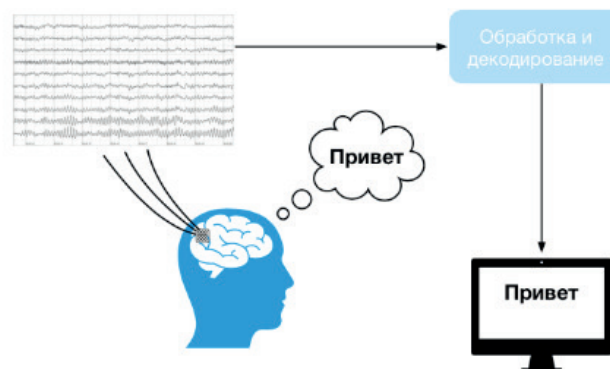


Рис. 3. Схема интерфейса распознавания внутренней речи  
Fig. 3. Diagram of the internal speech recognition interface

В головном мозге человека существуют зоны, ответственные за понимание языка и воспроизведение самой речи - зона Брока и зона Вернике. Эти зоны также ассоциируются и с внутренним проговариванием - достаточно зарегистрировать данные только с этих участков, а не со всего головного мозга [19].

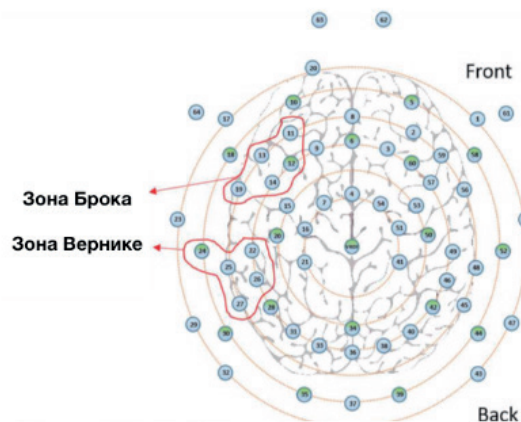


Рис. 4. Зоны Брока и Вернике  
Fig. 4. Broca and Wernicke Zones

## Технология проведения эксперимента

В большинстве работ прослеживается примерно одинаковый ход исследования. Вначале проводится сам эксперимент с испытуемыми, во время которого происходит процесс регистрации ЭЭГ, затем проводится предварительная обработка сигналов и декодирование сигналов: извлечение признаков и классификация.

### Эксперимент и запись

Для проведения эксперимента отбирается некоторое количество испытуемых, (от 4 до 20 человек). Испытуемые не должны иметь психических заболеваний. Для чистоты эксперимен-

<sup>1</sup> EEG Measures of Cognition // HVMN® Biohacker Guide. November 21, 2016 [Электронный ресурс]. URL: <https://hvmn.com/biohacker-guide/cognition/eeg-measures-of-cognition> (дата обращения: 11.12.2018).





та испытуемые должны быть правшами, а также одного возраста (как правило, 23 - 27 лет) и пола чтобы обеспечить наиболее сильное сходство в устройстве мозга.

На скальп испытуемого надевается устройство для регистрации ЭЭГ. Во время проведения эксперимента испытуемые должны быть сосредоточены на задании и не должны испытывать дискомфорта, поэтому кресло, в котором сидит человек, должно быть удобным и помещением, в котором проводится эксперимент, должно быть изолировано от громких звуков. Для уменьшения числа шумов испытуемых просят не двигаться, стараться меньше моргать и смотреть в одну точку на экране компьютера. Проблему на этом этапе может составить выгодный выбор расположения электродов [13]. Поскольку мозг каждого человека уникален, нельзя гарантировать, что результаты для каждого будут одинаковыми. Помимо скальпа, электроды могут накладываться на области около глаз, чтобы идентифицировать сигналы, вызванные морганием, и избавиться от них.

В зависимости от эксперимента могут предъявляться как визуальные, так и звуковые стимулы. Визуальные стимулы представляют собой сообщение-инструкцию на экране компьютера, например, слог, который необходимо представлять; звуковые - гудок, символизирующий начало эпохи эксперимента. Как правило, эксперимент протекает следующим образом: в начале звучит сигнал, означающий начало эпохи, вместе с этим предъявляется стимул-подсказка определенного вида, в течении которой испытуемый должен представлять произнесение одной из языковых единиц, затем следует отдых. Эксперименты проводятся в несколько сессий и для каждой языковой единицы проводится несколько попыток. В результате эксперимента будет получен набор сигналов, соответствующих представлению языковых единиц.

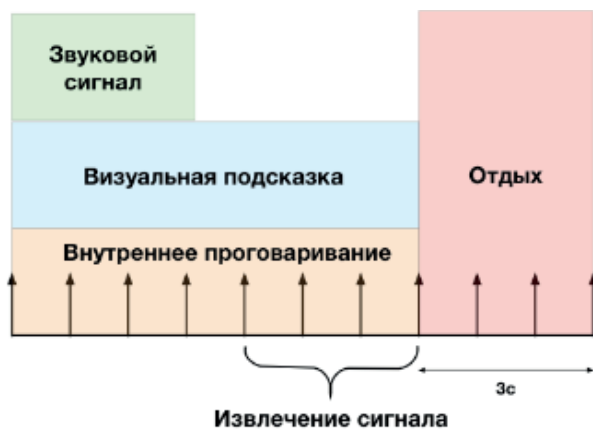


Рис. 5. Схема процедуры эксперимента [17]

Fig. 5. Scheme of the experiment procedure [17]

### Предварительная обработка данных

После сбора данных нужно провести их обработку. Типичными процедурами в этом процессе является применение полосового фильтра, который отсекает определенный диапазон частот. Обычно отсекаются высокие частоты, поскольку сигналы на этих частотах относятся либо к мышечной активности, либо к внешней среде. Предварительная обработка может проводиться с помощью таких программных пакетов, как Fieldtrip или EEGLab, совместимых с программой MATLAB. После этого

записанные данные разбиваются на эпохи, длительность которых соответствует длительности одной эпохи эксперимента. Полученные эпохи также разделяются на участки, соответствующие самому стимулу (внутреннему проговариванию) и процессу, предшествующему стимулу.

### Извлечение признаков

Одна из главных задач в разработке модели классификации данных ЭЭГ - этап извлечения признаков. В этой задаче необходимо преобразовать данные к виду понятных для классификатора векторов признаков: необходимо выделить из речевого сигнала участки, которые могут быть использованы для распознавания внутреннего проговаривания. Каждой из эпох, полученных на этапе предварительной обработки сигналов, применяется определенный метод извлечения признаков. Это может быть дискретное вейвлет-преобразование (при помощи свертки сигнала со специальной вейвлет-функцией преобразует сигнал из временного представления в частотно-временное), линейный дискриминантный анализ (применяется для нахождения линейных комбинаций признаков, наилучшим образом разделяющих классы объектов или событий). Также, рассчитываются средние величины, дисперсия, которые могут как-то охарактеризовать каждый из признаков.

### Классификация

Последним этапом в эксперименте является процесс классификации. В существующих исследованиях применяются современные алгоритмы, такие как метод опорных векторов, метод k-ближайших соседей, скрытая марковская модель, нейронные сети, экстремальное машинное обучение.

## Обзор проведенных экспериментов

Исследования, связанные с распознаванием внутреннего проговаривания, можно разделить на три типа: распознавание фонов, распознавание слогов, распознавание определенных слов. Последняя задача является достаточно сложной из-за множества нюансов, в том числе из-за того, что каждый человек может наделять определенное слово особым смыслом, в связи с чем универсальная классификация слов может быть сложной задачей.

Рассмотрим исследования, проведенные по каждому из вышеперечисленных типов экспериментов.

### Представление слов

В одном из исследований, основанных на распознавании внутреннего проговаривания слов, Ямагучи и соавторы [1] разработали систему распознавания внутреннего проговаривания с использованием ЭЭГ и ЭМГ (электромиографии). Система была разделена на две фазы: фаза обучения и фаза декодирования. Эти две фазы были применены в двух экспериментах, где испытуемые говорили фразы вслух на первом этапе и представляли их во второй фазе. Первый эксперимент состоял в том, чтобы идентифицировать слова из японского эквивалента игры «Камень, ножницы, бумага», второй - идентифицировать четыре слова, обозначающих времена года на японском языке. На этапе обучения области мозга, связанные с речью, были идентифицированы путем применения независимого анализа компонентов (ICA). В задаче с распознаванием времен года сигналы были преобразованы в последовательности гласных и согласных букв и эти взаимоотношения были приняты на вход Скрытой Марковской Модели. В результате ис-



следования была получена достаточно высокая точность распознавания - около 65 процентов.

В исследовании Серешке [20] было разработано два ИМК. Первый интерфейс был создан так, чтобы он смог различить 10-секундные отрезки внутреннего проговаривания слова «нет» и эквивалентные по времени отрезки тишины. Вторым ИМК был способен различить 10-секундные отрезки повторения слова «да» и эквивалентные по времени отрезки повторения слова «нет». Двенадцать испытуемых прошли тренировочные оффлайновые сессии и две онлайн-овые, тестирующие каждый из разработанных интерфейсов. Признаки были извлечены с помощью фильтра, основанного на быстрой корреляции (FCBF). Полученные признаки были использованы для тренировки SVM. Результаты работы первого ИМК показали точность в 75.9%± 11.4, второй ИМК показал немного меньшую точность - примерно 69%.

В исследовании Порбадника, Уэстера и Каллисса [2], помимо задачи распознавания, также была поставлена цель доказать, что внутреннее проговаривание может быть эффективно распознано, если произносимые слова состоят в блоках (последовательностях). Это исследование показало, что существует соотношение между порядком слов и степенью распознавания. Запись внутреннего проговаривания в блоках слов позволяло испытуемым больше концентрироваться на задании, что повлекло за собой уменьшение количества шумов.

В исследовании [16] для мультиклассовой классификации были выбраны слова "go", "back", "left", "right", "stop". Данные слова имеют четкое, вполне определенное значение, тем самым минимизируется вероятность того, что испытуемые будут представлять себе разные по смыслу вещи, думая об этих словах. В качестве классификатора в данной работе было выбрано экстремальное машинное обучение, основанное на сигмоидальной функции активации, поскольку такой метод имеет малую вычислительную сложность. Для каждой языковой единицы также была проведена бинарная классификация. Результаты показали, что мультиклассовая классификация не дает результатов, отличных от случайных - точность оставалась около 40%. При бинарной классификации максимальная достигнутая точность составила 87.9%.

#### Представление фоном

В исследовании Чиа и соавторов [3] испытуемым было предложено представлять 5 различных фонем, включающих в себя движения челюсти (-aa, ae), языка (-l, -r), небной занавески (-m, -n) и губ (-uu, -ow), а также фрикативные согласные (-s, -z). Были применены методы наивного байесовского классификатора и линейного дискриминантного анализа. Результаты показали, что сигналы подобного рода могут быть отличимы друг от друга. Также было показано, что для онлайн-овой классификации подходят простые линейные классификаторы.

В работе Бригама и Кумара [4] были использованы те же данные, что и в предыдущей работе, но подход к классификации был другой - в этом исследовании в качестве признаков были использованы коэффициенты авторегрессии, а в качестве алгоритма классификации - метод k-ближайших соседей. Основной целью этого эксперимента было использование представляемой речи для идентификации испытуемого в задаче аутентификации. Исследователи выяснили, что использование слогов вместо полных слов устраняет влияние смысла этого слова на мозговые сигналы. Их метод обработки сигналов показал высокий уровень точности идентификации субъекта,

однако, точность понизилась при дальнейших сессиях. Авторы предполагают, что на это повлияло самочувствие участников эксперимента. В итоге, максимальная достигнутая точность работы классификатора составила 44%.

#### Представление гласных

Один из экспериментов по распознаванию слогов был проведен в статье Да Салла [5]. В этом эксперименте во время внутреннего проговаривания представлялась артикуляция при произнесении звуков /a/ и /u/. Результаты показали, что активация двигательной коры головного мозга вместе с воображаемой гласной может быть классифицирована с помощью общих пространственных шаблонов и нелинейного классификатора. Однако, результаты также показали, что на точность классификации больше влияют мышечные движения, чем само внутреннее проговаривание. Итоговая точность - от 68 до 78 процентов.

В исследовании Мацумото и Хори [6] рассматривалось представление произнесения звуков a, i, u, e, o. Была исследована разница в использовании метода релевантных векторов и метода опорных векторов с Гауссовым ядром. Целью эксперимента было снизить вычислительную стоимость метода используя 19 каналов, общие пространственные фильтры и адаптивный сбор данных. Результаты показали, что разница в точности классификации небольшая (77% и 79%), более того, стоимость использования первого метода выше и требует больше данных для хороших результатов.

Классификация пар гласных также была произведена в работе за авторством Мин, Ким, Парк и Ли [8]. В качестве признаков были использованы средние значения, дисперсия, стандартное отклонение и асимметрия. Классификация производилась при помощи метода опорных векторов с функцией радиального базиса, а также с помощью экстремального машинного обучения с разными ядрами. Максимальная точность классификации была достигнута с помощью метода экстремального машинного обучения - пары гласных /a/ и /i/ были классифицированы с точностью 99.3%.

В работе Сармиенто и соавторов [7] была предпринята попытка разделить ментальные состояния при представлении открытых и закрытых гласных без представления артикуляции. Около областей Вернике и Брока был расположен 21 электрод. Анализ спектра мощности был применен для распознавания шумных сигналов. Нелинейный метод опорных векторов показал точность выполнения между 84% и 94%.

Завершая, следует отметить исследование авторов Камалакканна и Раджжумар [9]. В экспериментах были использованы также гласные буквы /a/, /i/, /u/, /e/, /o/. Для извлечения признаков были оценены параметры, характеризующие сигнал: оценка средних значений, дисперсии и стандартного отклонения, оценка средней мощности. Для классификации была использована нейронная сеть с 2 скрытыми слоями, которая была обучена с помощью метода обратного распространения ошибки. Входной слой состоял из 4 узлов, которые соответствовали 4 признакам, выходной слой состоял из 5 узлов, соответствующих 5 распознаваемым гласным. Средняя точность классификации, который был получен - 44%.

#### Проблемы в существующих исследованиях

Подводя итоги, стоит заметить, что, несмотря на то, что было произведено достаточно много исследований по теме класси-



фикации внутреннего проговаривания, разработка интерфейсов безмолвного доступа все еще находится на начальной стадии и существует много проблем, которые еще предстоит решить. Проблемой является выгодный выбор числа и места расположения электродов. При малом числе электродов на устройстве для регистрации ЭЭГ, произведенная классификация будет обладать очень низкой точностью. В связи с тем, что анатомические особенности строения головы и мозга каждого человека уникальны, расположение электродов на голове одинаковым образом не может гарантировать одинаковую точность для двух разных людей. Также, актуальной остается проблема онлайн-классификации. Для того, чтобы устройство безмолвного доступа на основе распознавания внутреннего проговаривания было удобным в реальной жизни, важно достичь большой скорости классификации, что может быть препятствием при разработке интерфейсов на основе ЭЭГ, поскольку они требуют проведения нескольких попыток для распознавания языковой единицы. Также очевидным является то, что все проведенные исследования ограничены очень небольшим числом речевых стимулов, поэтому сложно говорить о положительных результатах классификации для большего набора звуков. Также, последнее, о чем стоит сказать - существует проблема отсутствия единой точки зрения о конкретной методологии эксперимента, которая могла бы гарантировать постоянную точность в результатах. Существующие исследования различаются по виду предъявляемого стимула, по инструкциям, предъявляемым испытуемым, а также по методикам классификации, и у каждого из экспериментов есть свои преимущества и недостатки.

## Заключение

В данной работе были рассмотрены исследования, касающиеся изучения распознавания внутреннего проговаривания. Существующие исследования можно разделить на 3 типа по виду языковых единиц, которые представляли испытуемые. Это могут быть фонемы, гласные буквы и определенные слова. Исходя из результатов авторов, классификация слов практически не превосходит тех точностей, которые можно было бы получить случайно, однако классификация фонем и гласных обладает достаточно высокой точностью. Исследование проблемы классификации внутреннего проговаривания находится всё еще на начальной стадии и в будущем предстоит найти решения ко многим сопутствующим этому исследованию проблемам.

## References

- [1] Yamaguchi H., Yamazaki T., Yamamoto K., Ueno S., Yamaguchi A. et al. Decoding Silent Speech in Japanese from Single Trial EEGs: Preliminary Results. *Journal of Computer Science & Systems Biology*. 2015; 08:285-291. (In Eng.) DOI: 10.4172/jcsb.1000202
- [2] Porbadnigk A., Wester M., Callies J., Schultz T. EEG-based Speech Recognition - Impact of Temporal Effects. In: Proceedings of the International Conference on Bio-inspired Systems and Signal Processing - Vol. 1: BIOSIGNALS, (BIOS-TEC 2009), 2009, pp. 376-381. (In Eng.) DOI: 10.5220/0001554303760381
- [3] Chia X., Hagedorn J.B., Schoonover D., D'zmura M. EEG-based discrimination of imagined speech phonemes. *International Journal of Bioelectromagnetism*. 2011; 13(4):201-206. Available at: [https://pdfs.semanticscholar.org/b74f/c325556d1a7b5eb05fe90cde1f0c891357a3.pdf?\\_ga=2.236761025.1328281046.1558969187-1926881747.1554310837](https://pdfs.semanticscholar.org/b74f/c325556d1a7b5eb05fe90cde1f0c891357a3.pdf?_ga=2.236761025.1328281046.1558969187-1926881747.1554310837) (accessed 11.12.2018). (In Eng.)
- [4] Brigham K., Kumar B.V.K.V. Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy. Proceedings of the 4th International Conference on Bioinformatics and Biomedical Engineering, Chengdu, 2010, pp. 1-4. (In Eng.) DOI: 10.1109/ICBBE.2010.5515807
- [5] DaSalla C.S., Kambara H., Sato M., Koike Y. Single-trial classification of vowel speech imagery using common spatial patterns. *Neural Netw.* 2009; 22(9):1334-1339. (In Eng.) DOI: 10.1016/j.neunet.2009.05.008
- [6] Matsumoto M., Hori J. Classification of silent speech using support vector machine and relevance vector machine. *Applied Soft Computing*. 2014; 20:95-102. (In Eng.) DOI: 10.1016/j.asoc.2013.10.023
- [7] Sarmiento L.C., Lorenzana P., Cortes C.J., Arcos W.J., Bacca J.A., Tovar A. Brain computer interface (BCI) with EEG signals for automatic vowel recognition based on articulation mode. Proceedings of the 5th ISSNIP-IEEE Biosignals and Biorobotics Conference (2014): Biosignals and Robotics for Better and Safer Living (BRC), Salvador, 2014, pp. 1-4. (In Eng.) DOI: 10.1109/BRC.2014.6880997
- [8] Min B., Kim J., Park H.J., Lee B. Vowel Imagery Decoding toward Silent Speech BCI Using Extreme Learning Machine with Electroencephalogram. *BioMed Research International*. 2016; 2016:2618265. 11 p. (In Eng.) DOI: 10.1155/2016/2618265
- [9] Kamalakkannan R., Rajkumar R., Madan R.M., Shenbaga D.S. Imagined Speech Classification using EEG. *Advances in Biomedical Science and Engineering*. 2014; 1(2):20-32. (In Eng.)
- [10] Naci L., Cusack R., Jia V.Z., Owen A.M. The Brain's Silent Messenger: Using Selective Attention to Decode Human Thought for Brain-Based Communication. *Journal of Neuroscience*. 2013; 33(22):9385-9393. (In Eng.) DOI: 10.1523/JNEUROSCI.5577-12.2013
- [11] AlSaleh M.M., Arvaneh M., Christensen H., Moore R.K. Brain-computer interface technology for speech recognition: A review. Proceedings of the 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), Jeju, 2016, pp. 1-5. (In Eng.) DOI: 10.1109/APSIPA.2016.7820826
- [12] Hashim N., Ali A., Mohd-Isa W.N. Word-Based Classification of Imagined Speech Using EEG. In: Alfred R., Iida H., Ag Ibrahim A., Lim Y. (eds) Computational Science and Technology. ICCST 2017. Lecture Notes in Electrical Engineering. Springer, Singapore, 2018; 488:195-204. (In Eng.) DOI: 10.1007/978-981-10-8276-4\_19
- [13] Denby B., Schultz T., Honda K., Hueber T., Gilbert J.M., Brumberg, J.S. Silent Speech Interfaces. *Speech Communication*. 2009; 52(4):270-287. (In Eng.) DOI: 10.1016/j.specom.2009.08.002
- [14] Suppes P., Lu Zh.-L., Han B. Brain Wave Recognition of Words. *Proceedings of the National Academy of Sciences*. 1997; 94(26):14965-14969. (In Eng.) DOI: 10.1073/pnas.94.26.14965



- [15] Bogue R. Brain-computer interfaces: control by thought. *Industrial Robot: An International Journal*. 2010; 37(2):126-132. (In Eng.) DOI: 10.1108/01439911011018894
- [16] Qureshi M.N.I., Min B., Park H., Cho D., Choi W., Lee B. Multi-class Classification of Word Imagination Speech With Hybrid Connectivity Features. *IEEE Transactions on Biomedical Engineering*. 2018; 65(10):2168-2177. (In Eng.) DOI: 10.1109/TBME.2017.2786251
- [17] Nguyen C.H., Karavas G.K., Artemiadis P. Inferring imagined speech using EEG signals: A new approach using Riemannian manifold features. *Journal of Neural Engineering*. 2018; 15(1):016002. (In Eng.) DOI: 10.1088/1741-2552/aa8235
- [18] Combaz A., Manyakov N.V., Chumerin N., Suykens J.A.K., Hulle M.M.V. Feature Extraction and Classification of EEG Signals for Rapid P300 Mind Spelling. Proceedings of the 2009 International Conference on Machine Learning and Applications, Miami Beach, FL, 2009, pp. 386-391. (In Eng.) DOI: 10.1109/ICMLA.2009.27
- [19] Martin S., Brunner P., Iturrate I., Millán J.R., Schalk G., Knight R.T., Pasley, B.N. Word pair classification during imagined speech using direct brain recordings. *Scientific Reports*. 2016; 6:25803. 10 p. (In Eng.) DOI: 10.1038/srep25803
- [20] Sereshkeh A.R., Trott R., Bricout A., Chau T. Online EEG Classification of Covert Speech for Brain-Computer Interfacing. *International Journal of Neural Systems*. 2017; 27(8):1750033. (In Eng.) DOI: 10.1142/S0129065717500332

Поступила 11.12.2018; принята к публикации 10.02.2019;  
опубликована онлайн 19.04.2019.  
Submitted 11.12.2018; revised 10.02.2019;  
published online 19.04.2019.

#### Об авторах:

**Гавриленко Юлия Юрьевна**, студент, Филиал МГУ имени М.В. Ломоносова в городе Севастополе, Московский государственный университет имени М.В. Ломоносова (299000, Россия, г. Севастополь, ул. Героев Севастополя, д. 7), ORCID: <http://orcid.org/0000-0002-8704-6030>, [gavrilenko.yulia@bk.ru](mailto:gavrilenko.yulia@bk.ru)

**Саада Даниель Фирасович**, студент, Филиал МГУ имени М.В. Ломоносова в городе Севастополе, Московский государственный университет имени М.В. Ломоносова (299000, Россия, г. Севастополь, ул. Героев Севастополя, д. 7), ORCID: <http://orcid.org/0000-0003-4959-8093>, [daniel.saada@mail.ru](mailto:daniel.saada@mail.ru)

**Шевченко Андрей Олегович**, студент, факультет психологии, Московский государственный университет имени М.В. Ломоносова (119991, Россия, г. Москва, Ленинские горы, д. 1), ORCID: <http://orcid.org/0000-0002-9118-2617>, [andreyshevchenkomsu@gmail.com](mailto:andreyshevchenkomsu@gmail.com)

**Ильюшин Евгений Альбинович**, аспирант, ведущий программист, лаборатория открытых информационных технологий, факультет вычислительной математики и кибернетики, Московский государственный университет имени М.В. Ломоносова (119991, Россия, г. Москва, Ленинские горы, д. 1), ORCID: <http://orcid.org/0000-0002-9891-8658>, [eugene.ilyushin@gmail.com](mailto:eugene.ilyushin@gmail.com)

*Все авторы прочитали и одобрили окончательный вариант рукописи.*

#### About the authors:

**Yuliya Yu. Gavrilenko**, student, Lomonosov MSU Branch in Sevastopol, Lomonosov Moscow State University (7 Geroev Sevastopolya St., Sevastopol 299000, Russia), ORCID: <http://orcid.org/0000-0002-8704-6030>, [gavrilenko.yulia@bk.ru](mailto:gavrilenko.yulia@bk.ru)

**Daniel F. Saada**, student, Lomonosov MSU Branch in Sevastopol, Lomonosov Moscow State University (7 Geroev Sevastopolya St., Sevastopol 299000, Russia), ORCID: <http://orcid.org/0000-0003-4959-8093>, [daniel.saada@mail.ru](mailto:daniel.saada@mail.ru)

**Andrey O. Shevchenko**, student, Faculty of Psychology, Lomonosov Moscow State University (1, Leninskie gory, Moscow 119991, Russia), ORCID: <http://orcid.org/0000-0002-9118-2617>, [andreyshevchenkomsu@gmail.com](mailto:andreyshevchenkomsu@gmail.com)

**Eugene A. Ilyushin**, Graduate student, Senior Software Developer, Laboratory of Open Information Technologies, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University (1, Leninskie gory, Moscow 119991, Russia), ORCID: <http://orcid.org/0000-0002-9891-8658>, [eugene.ilyushin@gmail.com](mailto:eugene.ilyushin@gmail.com)

*All authors have read and approved the final manuscript.*

