

ВЫЗВАННЫЙ АРТИКУЛЯТОРНЫЙ ОТВЕТ ПРИ ВНУТРЕННЕМ И ВНЕШНЕМ ПРОГОВАРИВАНИИ¹

Шевченко Андрей Олегович

Аспирант, Московский Государственный Университет
имени М.В. Ломоносова
andreyshvchenkomsu@gmail.com

Вартанов Александр Валентинович

кандидат психологических наук, доцент, старший
научный сотрудник Московский Государственный
Университет имени М.В. Ломоносова
a_v_vartanov@mail.ru

EVOKED ARTICULATORY RESPONSE DURING INTERNAL AND EXTERNAL SPEECH²

**A. Shevchenko
A. Vartanov**

Summary: This study is aimed at identifying the time intervals of dynamic speech for the subsequent analysis of the evoked EEG potentials. Understanding of exact time intervals will allow to make better use of pattern classification possibilities for BCI (brain-computer interface). The study involved 7 subjects with the following stimuli: 7 phonemes (A - [a], B - [b], F - [f], G - [g], M - [m], R - [p], U - [u]) presented in random order, EMG envelope was used for recording: the subjects were to mentally repeat the phoneme or syllable with the duration and intonation that correspond to these units in external speech. As a result of the study, average potentials were plotted for all subjects during internal phonemic utterance, external utterance (vocalization), and perception of phonemes without their repetition. The results show a close connection between phonemic perception and articulation: phonemic perception reveals articulatory activity, while articulatory commands act as anticipatory discharges due to a connection between motor and auditory centers; the amplitude of internal articulation and phonemic enunciation are similar.

Keywords: evoked potentials, inner speech, neurolinguistics, subvocalization, phonemes, phonemic perception.

Аннотация: Данное исследование направлено на выделение на основе миограммы временных интервалов активного проговаривания, для последующего анализа вызванных потенциалов ЭЭГ. Понимание точных временных рамок позволит в более эффективно использовать возможности классификации паттернов для ИМК (интерфейс мозг-компьютер). В исследовании принимали участие 7 испытуемых, которым предъявлялись следующие стимулы: 7 фонем (А – [а], Б – [б], Ф – [ф], Г – [г], М – [м], Р – [р], У – [у]), представленные в случайном порядке, для записи использовалась огибающая ЭМГ верхней губы и мышцы гортани: испытуемым предстояло мысленно повторить фонему с длительностью и интонацией, которые соответствуют данным единицам во внешней речи. В результате исследования были построены усредненные потенциалы огибающей миограммы по всем испытуемым в процессе внутреннего проговаривания фонем, внешнего проговаривания (вокализация) и восприятия фонем без их повторения. Полученные результаты демонстрируют существование тесной связи фонематического восприятия с артикуляцией: при фонематическом восприятии проявляется артикуляторная активность, а артикуляторные команды, в свою очередь, выступают упреждающим разрядом в связи с наличием связи между моторными и слуховыми центрами; обнаруживается сходство внутреннего проговаривания с фонематическим проговариванием по амплитуде.

Ключевые слова: вызванные потенциалы, внутренняя речь, нейролингвистика, субвокализация, фонемы, фонематическое восприятие.

Введение

В рамках проведенных исследований по построению классификатора фонем [3,7] возникает интерес к определению точного интервала начала внутреннего проговаривания. Помимо этого, существует вопрос о том, выражены ли при внутреннем проговаривании и фонематическом восприятии артикуляционные компоненты и некие артикуляционные автоматизмы. В исследовании Lopez-Larrazetal (2010) был представлен интерфейс беззвучной речи на основе электромиографических (ЭМГ) сигналов, регистрируемых в лицевых мышцах. Целью этого интерфейса стало распознавание

изолированных прямых слогов испанского языка. Все слоги были разделены на пять групп в соответствии с анатомическим происхождением артикуляции согласного. Сигналы электромиографии отражали электрическую активность мышц во время артикуляции. Электроды ЭМГ были размещены на мышцах только одной стороны лица. Предварительно место крепления обрабатывалось спиртом, на датчики наносился проводящий электродный гель. Биполярные электроды размещались в направлении волокон лицевой мышцы с периодичностью в один сантиметр. Полученные данные использовались для обучения классификатора для этих слогов. Сигналы от каждого артикулированного слога были пре-

1 Работа выполнена при финансовой поддержке Российского научного фонда, проект No 20-18-00067.

2 The research is financially supported by the Russian Science Foundation, Project No20-18-00067

образованы в вектор признаков, компоненты которого представляли различные глобальные характеристики. Все отобранные признаки были инвариантны к сдвигу времени, чтобы процесс классификации не был чувствителен к времени произношения. Полученные результаты показали успешность и эффективность системы при распознавании новых артикулированных слогов [5].

Авторами Khan и Jahan (2018) было проведено исследование по распознаванию фонем беззвучной речи. Электроды ЭМГ располагались на шеи и под подбородком испытуемого. Полученные результаты были проанализированы с помощью трех классификаторов, средняя точность классификации фонем находилась в диапазоне 75-80%. Авторы сделали вывод, что признаки на основе wavelet дают более высокую точность в распознавании и, вероятно, играют важную роль в дифференциации фонемного рисунка беззвучной речи на основе ЭМГ [4].

Классиком же изучения процессов внутренней речи при помощи ЭМГ является А.Н. Соколов (1967). Результаты его исследований показали, что при чтении более сложных грамматических фраз речедвигательные потенциалы проявляются наиболее сильно. Это было ярко продемонстрировано при сравнении чтения на родном языке и на недостаточно изученном иностранном языке: при чтении русский текст читался без сильных речедвигательных напряжений, в то время как текст иностранного образца вызывал заметное речедвигательное напряжение в органах артикуляции. При анализе речедвигательных напряжений во время слушания было замечено, что речедвигательные напряжения проявляются при концентрации и фиксации на речи говорящего и при затруднениях в понимании речи другого [2].

Целью данного исследования является определение временных промежутков активации артикуляционного аппарата.

Методика

В исследовании принимали участие 7 испытуемых: четыре девушки и три юноши в возрасте от 18 до 23 лет. Был проведен следующий эксперимент. Запуск внутреннего проговаривания происходил на основе аудиального стимула (фонемы). Предъявлялись следующие стимулы:

- 7 фонем (А – [а], Б – [б], Ф – [ф], Г – [г], М – [м], Р – [р], У – [у]).

Стимулы предъявлялись в случайном порядке. Начало проговаривания задавалось специальным стимулом, сообщающим проговариваемую фонему. После появления задающего стимула создавалась пауза в 500 мс. После паузы на экране появлялся фиксационный крест и звуковой сигнал, который являлся стартовой командой

внутреннего проговаривания. Испытуемых во время команды на проговаривание просили повторить про себя фонему или слог с той же длительностью и интонацией, что и при внешней речи.

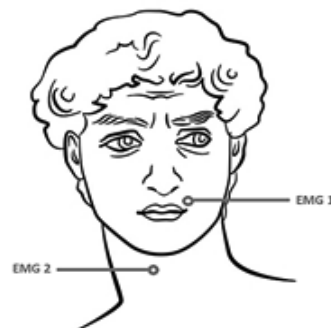


Рис. 1. Схема расположения электродов

В качестве программы для записи и просмотра ЭМГ, а также для предварительной очистки вызванных потенциалов использовалась Brain Sys (Brain Win). Регистрация электрической активности мозга проводилась монополярно, системы «Нейро-КМ» (компания «Статокин», Россия). Для предъявления стимулов использовалась программа Presentation (версия 18.0 фирмы NeurobehavioralSystems, Inc.).

Результаты

В результате проведенного исследования были построены усредненные потенциалы по всем испытуемым в процессе внутреннего проговаривания фонем, внешнего проговаривания (вокализация) и восприятия фонем без их повторения. Для усреднения и построения графиков использовалась программа Statistica 8, длительность усредненного потенциала составляла 200 мс. до стимула и 800 мс. после стимула.

Так, описывая полученные компоненты потенциала, можно наблюдать, что внешнее проговаривание фонем обусловлено резким сдвигом электрической активности от 150 мс. до 700 мс., при этом, в отличие внутренней речи, представлена обратная форма ЭМГ ВПна двух электродах (рис 2.), так, EMG 2 представляет плавную негативную активацию с компонента n200. EMG 1 представлен позитивной волной, однако, в отличие от EMG2, представлен он более поздними компонентами, схожими с P300.

Анализируя усредненные показатели проговаривания про себя фонем (рис. 1), среди компонентов можно выделить негативную волну на 200 мс., которая смещается в позитивный компонент в интервале 300-400 мс. При этом существует очевидное сходство компонентов миографии при восприятии и внутреннем проговаривании (рис 3)

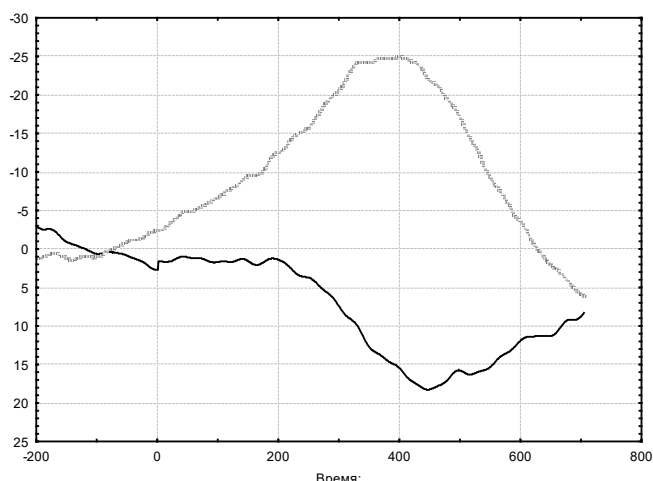


Рис. 2. Усредненный моторный ответ во время произнесения фонем вслух. EMG 1- сплошная линия, EMG 2 – пунктир.

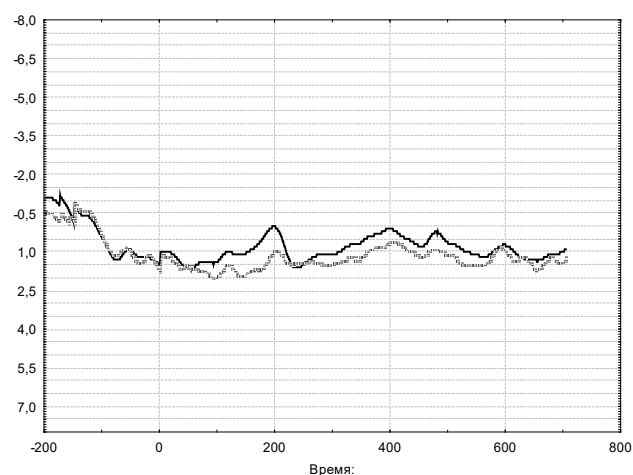


Рис. 3. Усредненный моторный ответ во время произнесения фонем про-себя. EMG 1- сплошная линия, EMG 2 – пунктир.

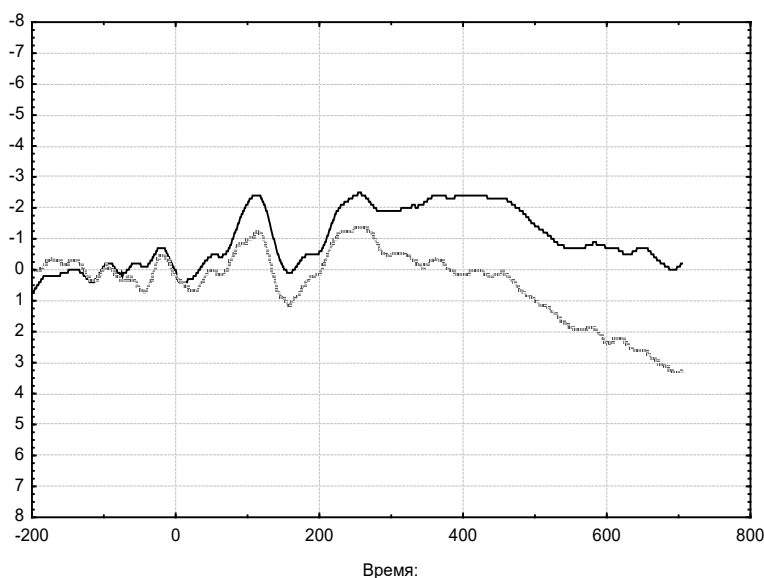


Рис. 4 Усредненный моторный ответ во время восприятия фонем. EMG 1- сплошная линия, EMG 2 – пунктир

Анализируя компоненты артикуляционного ответа, при восприятии фонем можно обратить внимание на то, что они сдвинуты в более ранние компоненты относительно процесса внутреннего проговаривания и повторяют волну, но уже с 150 мс.

Обсуждение результатов

Полученные результаты согласуются с проведенными исследованиями Соколова (1967), и подтверждают, что при фонематическом восприятии проявляется артикуляторная активность. В процессе внутреннего проговаривания электрод EMG1 расположенный над губой проявил большую выраженность, как и в проведенных исследованиях Соколова. Описывая фонематическое восприятие можно говорить об участии

автоматических артикуляторных командах, которые выступают упреждающим разрядом, и обуславливается это тем, что существует тесная связь между моторными и слуховыми центрами. При этом внутреннее проговаривание крайне схоже по амплитуде с фонематическим проговариванием. Полагаем, что актуальность использования ЭМГ для построения классификатора на основе ЭЭГ, при условии того, что ЭМГ может выступать индикатором начала внутреннего проговаривания. Таким индикатором могут выступать негативные компоненты, в нашем случае n2, меняющийся на компонент p3, может свидетельствовать о начале процесса внутреннего проговаривания, а, как следствие, о классификации.

В настоящее время уже были представлены успеш-

ные результаты построения интерфейсов на основе ЭМГ данных, также были показаны способы и подходы к моделированию предартикуляции на основе данного подхода. Так, в исследовании Schultz и Wand (2010) было показано, что моделирование, в зависимости от фонологических признаков, снижает количество ошибок при распознавании речи в интерфейсах более чем на 33% [6]. Исходя из этого, мы предполагаем, что на основе ЭМГ построение классификатора распознавания речи может быть успешным, однако для эффективного проектирования интерфейса необходимо большее количество электродов с большей частотной разверткой. Многие авторы также использовали кросскорреляционный анализ ЭМГ и подтвердили, что ЭМГ полезно в практическом плане для оценки наличия или отсутствия внутренней речи [1]. С помощью полученных данных мы планируем построить классификатор, в котором плавающим окном будут выступать зоны активационных изменений ЭМГ, а критерием начала проговаривания станут не только данные ЭЭГ, но и миография.

Заключение

На основании полученных нами результатов мы можем сделать следующие заключения:

1. Фонематическое восприятие сопряжено с артикуляцией. В данном случае можно говорить о том, что происходит упреждающий разряд на органы артикуляции при восприятии речи.
2. Фонематическое восприятие и соответствующая ему произвольная остаточная артикуляция возникают раньше (на 50 мс в среднем) по сравнению с сознательно активированным внутренним проговариванием. Фонематическое проговаривание (проговаривание, возникающее одновременно с восприятием) по огибающей миограммы и по ее амплитуде сходно с внутренним проговариванием.
3. Временной интервал, связанный с внутренним проговариванием, начинается в промежутке с 200 мс до 600 мс от начала команды на внутреннее проговаривание.

ЛИТЕРАТУРА

1. Бахтин, О.М., Кривко, Е.М., Киров, В.Н. Электромиографические компоненты, ассоциированные с внутренней речью // Журнал медико-биологических исследований. — 2020. — №2. — с. 111-120. URL: <https://cyberleninka.ru/article/n/elektromiograficheskie-komponenty-assotsirovannye-s-vnutrenney-rechyu> (дата обращения: 20.11.2022).
2. Соколов, А.Н. Внутренняя речь и мышление. — М.: «Просвещение», 1967. — 248 с.
3. Gavrilenko, Y., Saada, D., Ilyushin, E., Vartanov, A., and Shevchenko, A. The electroencephalogram based classification of internally pronounced phonemes // *Advances in Intelligent Systems and Computing*. — 2021. — Vol. 1310. — p. 335–341.
4. Khan, M., & Jahan, M. Classification of myoelectric signal for sub-vocal Hindi phoneme speech recognition // *Journal of Intelligent & Fuzzy Systems*. — 2018. — p. 1–8. doi:10.3233/jifs-161067 (<https://doi.org/10.3233/jifs-161067>)
5. Lopez-Larraz, E., Mozos, O.M., Antelis, J.M., & Minguéz, J. Syllable-based speech recognition using EMG // *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. — 2010. — p. 1-4. doi:10.1109/iembs.2010.5626426 (<https://doi.org/10.1109/iembs.2010.5626426>)
6. Schultz, T., & Wand, M. Modeling coarticulation in EMG-based continuous speech recognition // *Speech Communication*. — 2010. — Vol. 52(4). — p. 341–353. doi:10.1016/j.specom.2009.12.002 (<https://doi.org/10.1016/j.specom.2009.12.002>)
7. Suyuncheva, A., Saada, D., Gavrilenko, Y., Shevchenko, A.O., Vartanov, A.V., and Ilyushin, E. Reconstruction of words, syllables and phonemes of internal speech by EEG activity // *Advances in Intelligent Systems and Computing*. — 2021. — Vol. 1358. — p. 319-328.

© Шевченко Андрей Олегович (andreyshevchenkomsu@gmail.com), Вартанов Александр Валентинович (a_v_vartanov@mail.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»