

DOI: <https://doi.org/10.17816/gc623394>

Отслеживание нейронной активности при прослушивании естественной речи у детей с применением функции временного отклика

А.О. Рогачёв^{1*}, О.В. Сыроева^{1,2}¹ Научно-технологический университет «Сириус», Краснодарский край, Российская Федерация;² Институт высшей нервной деятельности и нейрофизиологии Российской академии наук, Москва, Российская Федерация

АННОТАЦИЯ

Развитие речи играет ключевую роль в психическом развитии ребёнка. При помощи речи ребёнок не только взаимодействует с внешним миром, но и учится осознавать свои собственные состояния и управлять своим поведением. Также, несомненно, от развития речи зависят последующие успехи ребёнка в обучении и, как следствие, в профессиональной жизни. Поэтому крайне важно исследовать механизмы нарушений развития речи и разрабатывать методы их диагностики и коррекции.

В настоящее время проводится большое количество когнитивных и нейрофизиологических исследований речи и механизмов, лежащих в основе её нарушений у детей. Исследования речи с применением ЭЭГ показывают устойчивые вызванные ответы головного мозга на слуховые и зрительные стимулы, связанные с речью (отдельные фонемы, слоги, буквы и т.д.), а также изменения этих ответов у детей с диагностированными нарушениями речи. Тем не менее, вопрос о нейрофизиологических предикторах и коррелятах тех или иных нарушений речевого развития остается дискуссионным. Вероятная причина этого может заключаться в особенностях экспериментальных методик: например, метод вызванных потенциалов требует использования изолированных «идеальных» стимулов для получения устойчивого ответа, а также многократного повторения одного стимула. Экологическая валидность таких исследований может быть сниженной: реакции мозга на продолжительные, «натуральные» стимулы могут отличаться от реакций на изолированные стимулы.

В последние годы в исследованиях речи набирает популярность метод, который называется «функция временного отклика» (temporal response function). Этот метод позволяет оценивать нейрофизиологические реакции на продолжительные, естественные, экологически валидные стимулы [1–3]. Применительно к исследованиям речи данный метод позволяет изучить ответ мозга на изменение акустических, лингвистических и семантических характеристик, которые присутствуют в любом потоке естественной нарративной речи [1].

Математическую основу функции временного отклика (ФВО) составляет решение уравнения:

$$w=(S^T S+\lambda E)^{-1} \cdot S^T R,$$

где S — матрица, содержащая рассматриваемые стимульные характеристики; R — матрица, содержащая нейрофизиологический сигнал, соответствующий стимулу; w — функция временного отклика — матрица коэффициентов линейного преобразования из пространства стимулов в пространство реакций.

Таким образом, ФВО является своеобразным «мостом» между стимулом и нейрофизиологическим ответом на него и отображает те нейронные операции, которые стоят между стимулом и реакцией [1]. Матрицы S и R являются матрицами временных лагов, что позволяет оценить ответ мозга на предъявленный стимул в заданном временном окне. ФВО в последнее время широко используется в исследованиях речи [2, 3]. Однако применений данного метода в исследованиях с участием детей сравнительно мало [4, 5]. Использование экологически валидных речевых стимулов в исследованиях с детьми значительно упрощает выполнение ими экспериментальных парадигм, а также позволяет оценить мозговой ответ на ту речь, которую можно слышать и в реальной жизни, а не только в экспериментально смоделированных условиях. Разнообразие применений ФВО как к акустическим, так и к лингвистическим особенностям речи представляет особый интерес при изучении психофизиологических механизмов развития речи у детей с различными траекториями развития. В нашей работе представлено применение данного подхода к исследованию развития речи у детей от 3 до 8 лет.

В исследовании приняли участие 56 детей (33 мальчика, 23 девочки) в возрасте от 3 до 8 лет ($M=5,64$; $SD=1,33$ лет). В ходе выполнения экспериментального задания участники прослушивали три заранее записанных женским голосом аудиоистории: детский рассказ про ежю, адаптированные варианты сказок «Кирпич и воск» и «Золотая утка».

Рукопись получена: 31.03.2023

Рукопись одобрена: 26.11.2023

Опубликована online: 20.01.2024

Все аудиостимулы сопровождалась видеорядом для поддержания внимания детей. Общая длительность стимулов — 15 минут. Стимулы предъявлялись посредством программного обеспечения Presentation® (Neurobehavioral Systems, Inc., Berkeley, CA). После прослушивания каждого из рассказов детям задавалось 8 вопросов типа «да/нет», направленных на оценку понимания содержания прослушанного аудиоматериала. В отдельный день исследования изучался актуальный уровень развития рецептивной и экспрессивной речи ребёнка при помощи игровой методики Preschool Language Scales версии 5 (PLS-5).

Проводилась регистрация 32-канальной ЭЭГ с использованием электроэнцефалографа Brain Products actiChamp (Brain Products GmbH, Gilching, Germany). Референт был установлен на позиции FCz. Предобработка ЭЭГ осуществлялась при помощи библиотеки MNE для Python и включала в себя фильтрацию данных в диапазоне от 1 до 15 Гц, визуальную инспекцию записи на наличие зашумленных каналов, интерполяцию плохих каналов (при необходимости), удаление глазодвигательных артефактов при помощи ICA, перереферирование записи на усредненный электрод. Синхронизация ЭЭГ и стимула производилась по метке на начало стимула. Далее и ЭЭГ, и стимул были сопоставлены в рамках выбранных эпох. Дальнейшая обработка выполнялась в среде MATLAB (версии 2021b) при помощи пакета mTRF Toolbox [1]. При помощи функций пакета выполнялось вычисление огибающей речевого стимула, которая затем подавалась в качестве стимула на вход ФВО. Производилось снижение частоты дискретизации стимула и ЭЭГ до 128 Гц. Для изучения было выбрано временное окно от -200 до 800 мс. Для анализа был выбран коэффициент прогнозирования ФВО, который представляет собой коэффициент корреляции между реальными данными и данными, которые предсказывает модель после кросс-валидации и обучения.

Среднее значение коэффициентов прогнозирования по всей выборке составило $0,041$ (разброс от $-0,002$ до $0,106$) и статистически значимо отличается от нуля ($t(55)=13,1$; $p < 0,001$). Выявлена статистически значимая положительная корреляция между значениями коэффициента прогнозирования, усреднёнными интраиндивидуально по всем каналам ЭЭГ, и возрастом участников ($r=0,379$; $p=0,004$). Способность линейной модели, лежащей в основе ФВО, предсказывать ЭЭГ-сигнал увеличивается с возрастом ребёнка.

Выявлена положительная статистически значимая корреляция между значениями коэффициента прогнозирования и значениями по шкале рецептивной речи методики PLS-5 ($r=0,33$; $p=0,026$). Значения PLS-5 сильно положительно коррелируют с возрастом участников ($r=0,596$; $p < 0,001$).

Также выявлена положительная корреляция между коэффициентом прогнозирования модели и результатами опросника на понимание прослушанного аудиоматериала ($r=0,39$; $p=0,012$). Кроме того, баллы опросника коррелируют со значениями по шкале рецептивной речи PLS-5 ($r=0,82$; $p < 0,001$) и с возрастом участников ($r=0,51$; $p=0,001$). Содержательно коэффициент прогнозирования функции временного отклика демонстрирует процесс кортикального отслеживания стимула, на которое в данный момент обращено внимание, и значимо коррелирует с пониманием услышанной речи [2, 3]. В нашей работе показано, что коэффициент прогнозирования статистически значимо положительно связан с возрастом детей и с их способностью к пониманию речи, измеренной при помощи методики PLS-5, а также с результатами опросника на понимание прослушанного, полученными непосредственно после выполнения экспериментальной задачи. Таким образом, применение функции временного отклика позволяет оценивать способность коры головного мозга к отслеживанию речевого акустического сигнала у детей. Также применение метода предоставляет нейрофизиологические корреляты рецептивной речи и процессов, связанных с пониманием услышанного. Возможно применение экспериментальной парадигмы для диагностики нейрофизиологических коррелятов рецептивной речи в различных возрастных группах у участников с различными уровнями развития языковых и речевых способностей. Представленная экспериментальная парадигма является частью исследования, проводимого направлением «Нейробиология устной и письменной речи при расстройствах развития» Научного центра когнитивных исследований университета «Сириус».

Ключевые слова: восприятие речи; развитие речи; экологически валидные стимулы; функция временного отклика; ЭЭГ; preschool language scales.

Как цитировать:

Рогачёв А.О., Сысоева О.В. Отслеживание нейронной активности при прослушивании естественной речи у детей с применением функции временного отклика // Гены и клетки. 2023. Т. 18, № 4. С. 640–644. DOI: <https://doi.org/10.17816/gc623394>

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Благодарности. Авторы благодарят всех участников исследования и команду проекта.

Источник финансирования. Финансирование проекта осуществлялось Министерством науки и высшего образования Российской Федерации (Соглашение № 075-10-2021-093; Проект COG-RND-2262).

СПИСОК ЛИТЕРАТУРЫ

1. Crosse M.J., Di Liberto G.M., Bednar A., Lalor E.C. The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli // *Frontiers in Human Neuroscience*. 2016. Vol. 10. P. 604. doi: 10.3389/fnhum.2016.00604
2. Alday P.M. M/EEG analysis of naturalistic stories: A review from speech to language processing // *Language, Cognition and Neuroscience*. 2019. Vol. 34, N 4. P. 457–473. doi: 10.1080/23273798.2018.1546882
3. Di Liberto G.M., Hjortkjær J., Mesgarani N. Editorial: Neural Tracking: Closing the Gap Between Neurophysiology and Translational Medicine // *Frontiers in Neuroscience*. 2022. Vol. 16. P. 872600. doi: 10.3389/fnins.2022.872600
4. Kalashnikova M., Peter V., Di Liberto G.M., et al. Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech // *Scientific Reports*. 2018. Vol. 8, N 1. doi: 10.1038/s41598-018-32150-6
5. Di Liberto G.M., Peter V., Kalashnikova M., et al. Atypical cortical entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia // *NeuroImage*. 2018. Vol. 175. P. 70–79. doi: 10.1016/j.neuroimage.2018.03.072

КОНТАКТНАЯ ИНФОРМАЦИЯ

* А.О. Рогачёв; адрес: Российская Федерация, 354340, Краснодарский край, пгт Сириус, пр-т Олимпийский, д. 1;
e-mail: aorogachev@gmail.com

DOI: <https://doi.org/10.17816/gc623394>

Neural tracking of natural speech listening in children: temporal response function (TRF) approach

A.O. Rogachev^{1*}, O.V. Sysoeva^{1,2}¹ Sirius University of Science and Technology, Krasnodar Region, Russian Federation;² Institute of Higher Nervous Activity and Neurophysiology of Russian Academy of Sciences, Moscow, Russian Federation

ABSTRACT

Speech development is crucial for a child's mental growth. Moreover, speech development significantly impacts a child's educational and professional achievements. It enables the child to interact with the external environment and develop self-awareness and behavioral skills. Thus, the study of the mechanisms of speech development disorders and the development of diagnostic and remediation strategies is essential.

Numerous cognitive and neurophysiological investigations into speech and its associated disorders among children are presently being conducted. Electroencephalography (EEG) studies demonstrated constant evoked reactions in response to auditory and visual stimuli associated with speech, including individual phonemes and syllables. Moreover, alterations in these reactions were detected among children with diagnosed speech ailments. The debate surrounding the neurophysiological predictors and correlates of specific speech development disorders continues. The use of isolated "ideal" stimuli and multiple repetitions of a single stimulus, as required by the method of evoked potentials, may create peculiarities in experimental techniques. Thus, brain responses to prolonged, "natural" stimuli may differ from those obtained with isolated stimuli. This could potentially reduce the ecological validity of such studies.

In recent years, the temporal response function has become increasingly popular in speech research. This method enables estimating neurophysiological responses to continuous, natural, and ecologically valid stimuli [1–3]. When applied to speech research, this method allows for the study of the brain's response to changes in acoustic, linguistic, and semantic characteristics present in natural narrative speech [1].

The mathematical basis of the temporal response function (TRF) is the solution of the equation:

$$w=(S^T S+\lambda E)^{-1}\cdot S^T R,$$

It is calculated from the stimulus characteristics, represented by the matrix S , the neurophysiological signal corresponding to the stimulus, represented by matrix R , and the temporal response function, represented by matrix w , a matrix of linear transformation coefficients from stimulus space to response space [1]. The TRF serves as a "bridge" between the stimulus and the neurophysiological response as it reflects the neural operations that occur between the two. The S and R matrices are matrices with time lags, enabling estimation of the brain's response to the presented stimulus within a specific time period.

The TRF has been utilized extensively in speech studies [2, 3]. Nevertheless, few studies have used this approach in research that involves children [4, 5]. The use of ecologically valid speech stimuli in child studies simplifies their performance in experimental paradigms and enables the evaluation of brain responses to speech as it occurs in real-life situations, not only in experimentally created conditions. The TRF has various applications to both linguistic and acoustic features of speech, which attracts particular interest in studying the psychophysiological mechanisms of speech development in children with various developmental trajectories. This approach is applied in our study of speech development in children aged 3 to 8 years.

Fifty-six children, consisting of 33 boys and 23 girls aged between 3 and 8 years, participated in this study with a mean age of 5.64 (SD=1.33 years). Participants were required to listen to three audio stories, including a children's story about hedgehogs and adapted versions of the tales "Brick and Wax" and "The Golden Duck", all of which were recorded by a female voice. All audio stimuli were accompanied by video to maintain children's attention. The total duration of the stimuli was 15 minutes. The audio stories were presented using Presentation® software from Neurobehavioral Systems, Inc. in Berkeley, CA. The comprehension of the stories was assessed by asking children 8 "yes/no" questions after each story. Furthermore, on a different day of the study, the Preschool Language Scales Fifth Edition (PLS-5) method was used to examine the child's current level of receptive and expressive speech development.

A 32-channel EEG was obtained using a Brain Products actiCHamp (Brain Products GmbH, Gilching, Germany) with reference electrodes positioned at the FCz location. EEG pre-processing was completed with the MNE library for Python, which entailed data filtering between 1 and 15 Hz, visually examining record for any noisy channels, interpolation of deficient channels (as needed), removal of oculomotor artifacts using independent component analysis, and re-referencing the EEG recording

Received: 31.03.2023

Accepted: 26.11.2023

Published online: 20.01.2024

to an average electrode. The EEG and stimulus were synchronized by labeling at the start of the stimulus. They were subsequently aligned during specific epochs. Processing was carried out with MATLAB (version 2021b) using the mTRF Toolbox [1]. The Toolbox's functions were employed to assess the speech stimulus envelope, which was then introduced as input to the TRF. The stimulus and EEG sampling rate were reduced to 128 Hz, and the analysis used a time window ranging from -200 to 800 ms. The TRF prediction coefficient, representing the correlation coefficient between actual data and data predicted by the model post-training and cross-validation, was selected for analysis.

The mean value for prediction coefficients across the entire sample was 0.041 (range: -0.002 to 0.106). These coefficients were significantly different from zero ($t(55)=13.1$, $p < 0.001$). Additionally, a significant positive correlation was found between the prediction coefficients averaged intraindividually across all EEG channels and the age of the participants ($r=0.379$, $p=0.004$). The linear model underlying the TRF was able to predict the EEG signal better as the age of the child increased. A significant positive correlation was observed between the prediction coefficient values and the values on the receptive speech scale of the PLS-5 ($r=0.33$, $p=0.026$). In addition, PLS-5 scores were strongly correlated with age ($r=0.596$, $p < 0.001$). There was a positive correlation observed between the model prediction coefficient and the scores obtained from the listening comprehension questionnaire ($r=0.39$, $p=0.012$). Additionally, the questionnaire scores were found to be significantly associated with scores from the PLS-5 receptive speech scale ($r=0.82$, $p < 0.001$) as well as with the age of study participants ($r=0.51$, $p=0.001$).

Substantively, the predictive coefficient of the temporal response function illustrates the cortical tracking process of the stimulus currently receiving attention and is significantly associated with listening comprehension [2, 3]. Our research indicates a significant and positive correlation between children's age, their comprehension of speech as measured by the PLS-5 method, and the results of the listening comprehension questionnaire conducted immediately after the experimental task. The prediction coefficient supports this finding. Thus, the use of the temporal response function enables the evaluation of the cerebral cortex's capacity to follow the acoustic signal of speech in children. Additionally, this approach yields neurophysiological markers of speech reception and comprehension processes. It is feasible to apply an experimental framework to identify neurophysiological correlations of receptive speech across various age groups and participants with varying levels of language and speech skills. The experimental paradigm presented here is a component of research carried out by the Neurobiology of Oral and Written Speech in Developmental Disorders division at the Center for Cognitive Sciences, Sirius University. The authors extend their gratitude to the study participants and project team.

Keywords: speech perception; speech development; ecologically valid stimuli; temporal response function; EEG; pre-school language scales.

To cite this article:

Rogachev AO, Syssoeva OV. Neural tracking of natural speech listening in children: temporal response function (TRF) approach. *Genes & Cells*. 2023;18(4):640–644. DOI: <https://doi.org/10.17816/gc623394>

ADDITIONAL INFORMATION

Acknowledgments. The authors thank all study participants and the project team.

Funding sources. The research is supported by the Ministry of Science and Higher Education of the Russian Federation (Agreement No. 075-10-2021-093, Project No. COG-RND-2262).

REFERENCES

1. Crosse MJ, Di Liberto GM, Bednar A, Lalor EC. The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Frontiers in Human Neuroscience*. 2016;10:604. doi: 10.3389/fnhum.2016.00604
2. Alday PM. M/EEG analysis of naturalistic stories: A review from speech to language processing. *Language, Cognition and Neuroscience*. 2019;34(4):457–473. doi: 10.1080/23273798.2018.1546882
3. Di Liberto GM, Hjortkjær J, Mesgarani N. Editorial: Neural Tracking: Closing the Gap Between Neurophysiology and Translational Medicine. *Frontiers in Neuroscience*. 2022;16:872600. doi: 10.3389/fnins.2022.872600
4. Kalashnikova M, Peter V, Di Liberto GM, et al. Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Scientific Reports*. 2018;8(1). doi: 10.1038/s41598-018-32150-6
5. Di Liberto GM, Peter V, Kalashnikova M, et al. Atypical cortical entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia. *NeuroImage*. 2018;175:70–79. doi: 10.1016/j.neuroimage.2018.03.072.

AUTHORS' CONTACT INFO

* A.O. Rogachev; address: 1 Olimpiyskiy avenue, 354340 Krasnodar region, Sirius, Russian Federation; e-mail: aorogachev@gmail.com