

# Новый подход к подавлению шума в кохлеарных имплантатах: Одноканальный алгоритм подавления шума<sup>1</sup>

Хуали Чжоу (Huali Zhou)<sup>1</sup>, Нинъюань Ван (Ningyuan Wang)<sup>2</sup>, Нэнхэн Чжэн (Nengheng Zheng)<sup>3</sup>, Гуанчжэн Ю (Guangzheng Yu)<sup>1\*</sup> и Цинлинь Мэн (Qinglin Meng)<sup>1\*</sup>

<sup>1</sup> Акустическая лаборатория, Институт физики и оптоэлектроники, Южно-Китайский технологический университет, Гуанчжоу, Китай;

<sup>2</sup> Компания «Ньюротрон Биотехнологджи Инк.» («Nurotron Biotechnology Inc.»), Ханчжоу, Китай;

<sup>3</sup> Гуандунская ведущая лаборатория интеллектуальной обработки информации, Колледж электроники и информационной инженерии, Шэньчжэньский университет, Шэньчжэнь, Китай.

## ОТКРЫТЫЙ ДОСТУП

Под редакцией: Юки Хаяшида (Yuki Hayashida),  
Университет Осаки, Япония

Проверено: Себастьян Осили (Sebastián Ausili),  
Университет Майами, США; Ли Сю (Li Xu), Университет  
Огайо, США

\* Автор, отвечающий за переписку: Гуанчжэн Ю  
(Guangzheng Yu), [scgzyu@scut.edu.cn](mailto:scgzyu@scut.edu.cn); Цинлинь Мэн  
(Qinglin Meng), [mengqinglin@scut.edu.cn](mailto:mengqinglin@scut.edu.cn)

Специализированный раздел:

Статья была отправлена в «Нейронные  
технологии» («Neural Technology»), раздел журнала  
«Frontiers in Neuroscience» («Неизвестное в неврологии»)

Получено: 30 ноября 2019 года

Принято: 16 марта 2020 г.

Опубликовано: 21 апреля 2020 г.

Ссылка на источник:

Чжоу Х., Ван Н., Чжэн Н., Ю Г. и Мэн Ц. (2020)

Новый подход к подавлению шума в кохлеарных  
имплантатах: Одноканальный алгоритм подавления шума.  
(Zhou H, Wang N, Zheng N, Yu G and Meng Q (2020) A New  
Approach for Noise Suppression in Cochlear Implants: A  
Single-Channel Noise Reduction Algorithm) Front. Neurosci.  
14:301. doi: 10.3389/fnins.2020.00301)

Улитка «трансформирует» вибрации акустического «языка» в воздухе в импульсы нейронного «языка», которые затем передаются в мозг для слухового понимания и/или восприятия.

Во время этого процесса «трансформации» внутри улитки высокое разрешение в областях, отвечающих за передачу времени, частоты и интенсивности, гарантирует высокое качество входной нейронной информации для мозга, что жизненно важно для имеющих у людей слуховых способностей. При этом кохлеарные имплантаты (КИ) характеризуются грубым искусственным кодированием и интерфейсами, и пользователи КИ сталкиваются с большими сложностями в обычных акустических средах, чем люди с нормальной способностью слышать (НС). Шум от источников звука, которые не интересуют слушателя, может игнорироваться слушателями НС, но может отвлекать пользователя КИ. Авторы обсуждают методы подавления шума КИ и представляют способ управления шумом в новой системе имплантатов. Моноуральный алгоритм подавления шума на основе оценки соотношения сигнал / шум «eVoice», который встроен в процессоры «Nurotron® Enduro™», проходил оценку в двух экспериментах по восприятию речи.

Результаты показывают, что четкость восприятия речи в стационарном речевом шуме может быть значительно улучшена с помощью eVoice. Аналогичные результаты наблюдались и в других устройствах КИ с одноканальными методами шумоподавления. В частности, среднее снижение порога восприятия речи в настоящем исследовании составило 2,2 дБ. Сообщество «Nurotron» уже насчитывает более 10 000 пользователей, при этом eVoice является стартовым элементом для управления шумом в новой системе. Данный алгоритм гарантирует дальнейшее подавление нестационарного шума, разделение источников звука в пространстве, двустороннее восприятие

звуков, конфигурацию микрофона и учет условий окружающей среды. Имеющиеся данные, в том числе настоящее исследование, предполагают, что в системах КИ должны применяться методы подавления шума. Искусственный слух слушателей с КИ требует более совершенных методов обработки сигналов, чтобы уменьшить нагрузку на мозг и повысить четкость восприятия речи в шумной обстановке.

**Ключевые слова:** кохлеарный имплантат, шумоподавление, проблема выделения сообщения из потока речевой информации, моноуральное восприятие, речь в шуме, четкость восприятия, Nurotron, eVoice.

<sup>1</sup>Части этой работы были представлены в разделе «Внедрение и оценка одноканального метода шумоподавления в кохлеарных имплантатах» («Implementation and evaluation of a single-channel noise reduction method in cochlear implants») на конференции по имплантируемому слуховому протезам в 2017 г., Лейк-Тахо, Калифорния, США, июль 2017 г.; «Нейронный интерфейс: Границы и приложения» («Neural Interface: Frontiers and Applications») в журнале «Достижения экспериментальной медицины и биологии» («Advances in Experimental Medicine and Biology»), том 1101; и «Тестирование четкости восприятия речи «eVoice», нового алгоритма шумоподавления в системах «Nurotron Enduro» («Speech intelligibility test of “eVoice”, a new noise-reduction algorithm in Nurotron Enduro systems») на Азиатско-Тихоокеанском симпозиуме 2019 года по кохлеарным имплантатам и смежным наукам, Токио, Япония, ноябрь 2019 года.

## ВВЕДЕНИЕ

Кохлеарный имплантат (КИ) — один из наиболее успешных из когда-либо разработанных протезов, направленный на восстановление способности слышать посредством передачи акустической информации в мозг людей с тяжелыми и очень тяжелыми нарушениями слуха за счет электрической стимуляции слуховых нервных волокон (Шеннон (Shannon), 2014). Искусственно стимулированная при помощи электротехнических средств способность слышать, обеспечиваемая существующими в настоящее время КИ, подходит для речевого общения, но все еще не дает удовлетворительного результата по сравнению с нормальной способностью слышать (НС), особенно с точки зрения распознавания речи в шуме.

Проблема шумов — частая жалоба пользователей КИ (см., например, Рен и соавт. (Ren et al.), 2018). Из-за различий, связанных со временем имплантации, анамнезом потери способности слышать, реабилитацией и обучением, хирургическими условиями, устройствами, обработкой сигналов и т. д., пользователи КИ всегда сообщали о значимых различиях в слуховых способностях. Эти причины расхождения между КИ и НС и разницы восприятия КИ людьми можно разделить на «восходящий» и «нисходящий» типы (Моберли и Рид (Moberly and Reed), 2019; Тамати и соавт. (Tamati et al.), 2019).

С практической точки зрения знания о нисходящей памяти и познании полезны для реабилитации и принятия хирургических решений (Крал и соавт. (Kral et al.), 2019), а взаимосвязь между результативностью речи и функциями обработки «восходящего» сигнала, особенно на интерфейсе электрода, определяет инженерные подходы, используемые в современных КИ-системах (Вилсон и соавт. (Wilson et al.), 1991; Лоизу (Loizou), 1999, 2006; Рубинштейн (Rubinstein), 2004; Цзен (Zeng), 2004; Цзен и соавт. (Zeng et al.), 2008; Вутерс и соавт. (Wouters et al.), 2015; Нogueйра и соавт. (Nogueira et al.), 2018). Хотя восходящий подход предлагалось включить в системы КИ для формирования адаптивного нейронного протеза с замкнутым контуром (МакЛафлин и соавт. (Mc Laughlin et al.), 2012), авторы задействуют только нисходящие методы, которые могут быть полезными для пользователей КИ при решении проблемы маскировки шума, согласно дальнейшему обсуждению.

Как направить более полезную информацию в восходящем потоке? Формы колебаний давления звукового сигнала разбираются здоровыми улитками уха на тонкие временно-спектральные «слуховые образы». КИ пытаются получить и доставить те же образы, но, к сожалению, грубым образом. Теории группировки, анализа сцены, демаскировки и внимания продемонстрировали важность точного кодирования акустических сигналов, включая высоту тона или различаемые гармоники, общее вступление и пространственные сигналы. Для большинства систем КИ только временные огибающие от ограниченного числа каналов могут быть переданы в нерв, а текущие взаимодействия между каналами являются основным ограничением многоканальной структуры КИ.

Было изучено несколько направлений исследований для улучшения эффективности КИ при распознавании речи в шуме путем обновления технологии современных многоканальных устройств: (1) стимуляция слуховых нервов новыми физическими способами, такими как оптическая стимуляция (Йешке и Мозер (Jeschke and Moser), 2015) и проникающая стимуляция

нерва (Миддлбрукс и Снайдер (Middlebrooks and Snyder), 2007); (2) разработка массивов внутрикочлеарных электродов разной длины, формы электродов и механических характеристик (Дханасингх и Джолли (Dhanasingh and Jolly), 2017; Ребшер и соавт. (Rebscher et al.), 2018; Сю и соавт. (Xu et al.), 2018); (3) управление и акцентирование распространения тока посредством одновременной активации нескольких электродов (Беренштайн и соавт. (Berenstein et al.), 2008); Бонхам и Литвак (Bonham and Litvak), 2008); (4) уточнение стратегий во времени путем введения гармоник (Ли и соавт.) (Li et al.), 2012), частотой пересечения нулевой отметки (Цирхофер (Zierhofer), 2003) или пиков (Ван Хозель (Van Hoesel), 2007) и медленно меняющихся тонких временных структур (Ни и соавт. (Nie et al.), 2005; Менг и соавт. (Meng et al.), 2016); и (5) улучшение речи или подавление шума до или в рамках основных стратегий обработки сигналов. Первое и второе направления развиваются с позиций нейрофизиологии; третье основано на психофизических тестах; в четвертом используется комбинация обработки сигналов и психофизики, а пятое в основном концентрируется на обработке сигналов. Все эти аспекты заслуживают дальнейшего изучения.

За последние два десятилетия пятый подход к улучшению восприятия речи или подавлению шума до или в рамках основных стратегий обработки сигналов стал актуальным в научных и промышленных исследованиях. Подавление шума и улучшение восприятия речи — это две составляющие одного направления исследований, целью которого является улучшение четкости восприятия или качества речи в шуме, в большинстве случаев с помощью улучшения отношения сигнал/шум (ОСШ) системы обработки сигналов. Некоторые методы шумоподавления в телекоммуникациях и слуховых аппаратах использовались для обработки шумных речевых сигналов, а затем обработанные сигналы представлялись через громкоговорители пользователям КИ (например, классическое одноканальное спектральное вычитание) (Янг и Фу (Yang and Fu), 2005) для проверки реализуемости метода. В настоящее время существуют более сложные одноканальные алгоритмы шумоподавления (АШП) (Чен и соавт. (Chen et al.), 2015), направленный микрофон или формирователи луча слуховых аппаратов на основе мультимикрофона (Чунг и соавт. (Chung et al.), 2004; Бухнер и соавт. (Buechner et al.), 2014), а в последнее время появились алгоритмы на основе глубоких нейронных сетей (Лаи и соавт. (Lai et al.), 2018; Геринг и соавт. (Goehring et al.), 2019), которые были апробированы со слушателями с КИ. Другое направление исследований заключается в стремлении специально оптимизировать параметры алгоритма с учетом различий между слушателями с КИ и НС. Параметры обычно связаны с оценкой шума или функцией усиления для снижения шума (Ху и соавт. (Hu et al.), 2007; Кастури и Лоизу (Kasturi and Loizou), 2007; Маугер и соавт. (Mauger et al.), 2012a, b; Ван и Хансен (Wang and Hansen), 2018). Все эти исследования продемонстрировали значительные улучшения, которые можно объяснить более высоким ОСШ, полученным с помощью методов, применявшихся ранее или задействованных в рамках основных стратегий КИ.

В новейших версиях процессоров КИ от действующих коммерческих компаний, таких как «Кохлеар» («Cochlear®») (Херсбах и соавт. (Hersbach et al.), 2012), «Эдвэнсд Бионикс» («Advanced Bionics®») (Бухнер и соавт. (Buechner et al.), 2010) и «МЕД-ЭЛ» («MED-EL®») (Хаген и соавт. (Hagen et al.), 2019), были реализованы и оценены один или несколько алгоритмов моноурального шумоподавления на основе ОСШ и направленных формирователей луча микрофона или мультимикрофонов на базе пространственных сигналов. Последние значительно улучшают четкость восприятия речи для пользователей КИ при шуме. Однако подход основан на предположении, что целевые источники речи и шума пространственно разделены. Таким образом, АШП с одним микрофоном в системах КИ по-прежнему заслуживают внимания для улучшения восприятия речи при шуме, особенно в сценариях, когда целевая речь и источники шума пространственно не разделены.

В литературе сообщалось о некоторых АШП с одним микрофоном, которые уже реализованы в коммерческих продуктах КИ. «ClearVoice» представляет собой моноуральный АШП, реализованный со стратегией обработки речи «HiRes 120» (Бухнер и соавт. (Buechner et al.), 2010; Холден и соавт. (Holden et al.), 2013). Сначала он оценивает шум, предполагая, что амплитуда энергии речи часто изменяется, а энергия фонового шума менее модулирована. Затем

коэффициент усиления уменьшается для каналов, идентифицированных как передающие (в основном) шумовую энергию.

Оценка шума работает при временном окне 1,3 секунды, которое является временем активации данного алгоритма. Эксперименты показали, что ClearVoice может улучшить четкость восприятия речи в условиях стационарного шума (Бухнер и соавт. (Buechner et al.), 2010; Кам и соавт. (Kam et al.), 2012). Другой моноуральный АШП реализован с помощью стратегии ACE (кодирование расширенных комбинаций) в устройствах «Nucleus». В ней используется алгоритм минимальной статистики с оптимальным методом выравнивания для оценки шума (Мартин (Martin), 2001) и априорной оценкой ОСШ (МакОлей и Малпасс (McAulay and Malpass), 1980) в сочетании с модифицированным методом усиления Винера (Лоизу (Loizou), 2007). Сообщалось о значительном улучшении способности слышать при стационарном шуме (Досон и соавт. (Dawson et al.), 2011).

Авторы представляют «eVoice» —

недавно разработанный одноканальный АШП на основе проектного ОСШ, который был реализован в исследовательском процессоре второго поколения «Enduro™» компании «Nurotron». Nurotron, молодая компания, находящаяся в г.Ирвин, Калифорния, США, и Ханчжоу, Чжэцзян, Китай, в настоящее время установила имплантаты более 10000 пациентов. Система «Nurotron» характеризуется 24 каналами электродов, а эффективность восприятия речи ее пользователями в состоянии тихой и послеоперационной разработки сопоставимы с предыдущими данными других производителей (Цзен и соавт. (Zeng et al.), 2015; Гао и соавт. (Gao et al., 2016). Оценка шума в eVoice осуществляется посредством покадровой обработки с использованием относительно короткого временного окна. eVoice основан на классических алгоритмах обработки сигналов и не является первым устройством КИ, где используется такой подход. Целью этого исследования в том числе является отчет о результатах эксперимента по четкости восприятия с использованием eVoice и переосмысление управления шумом в новой системе КИ, в данном случае — системе «Nurotron».

### **EVOICE СИСТЕМЫ «NUROTRON»: ОДНОКАНАЛЬНЫЙ АШП**

Базовой стратегией Nurotron по умолчанию является стратегия расширенного выбора пиков (РВП), которая аналогична стратегии «*n*-из-*m*» («*n*-of-*m*») (Цзен и соавт. (Zeng et al., 2015). Стратегия РВП основана на оконном преобразовании Фурье (STFT) и обычно выбирает восемь максимальных значений (автоматический процесс, определенный в стратегии кодирования) для стимуляции в каждом кадре (Пинг и соавт. (Ping et al.), 2017). Блок-схема стратегии РВП и eVoice показана на **Рисунке 1**. В РВП акустический входной сигнал сначала предварительно усиливается, за ним следует полосовая фильтрация (номер полосы *m* обычно равен номеру активного электрода, т. е. *m* = 24 в устройствах «Nurotron») и вычисление огибающей. Затем при выборе пика указываются *n* полос с наибольшей амплитудой для дальнейшего нелинейного сжатия и электростимуляции (обычно *n* = 8 в устройствах «Nurotron»). eVoice — это метод шумоподавления на основе огибающей, который реализуется между вычислением огибающей и выбором пика. Он состоит из двух этапов: оценка шума и расчет усиления (Ван и соавт. (Wang et al.), 2017).

#### **Оценка шума**

Алгоритм оценки шума основан на усовершенствованном алгоритме минимально контролируемого рекурсивного усреднения (MCRA-2) (Рангачари и Лоизу (Rangachari and Loizou), 2006). Мощность шума в каждом канале оценивается покадрово, а не временным окном, куда входит несколько кадров для сокращения времени отклика. Если предположить, что шум суммируется, тогда во временной области входной сигнал  $y(n)$  можно обозначить как

$$y(n) = x(n) + d(n) \quad (1)$$

где  $x(n)$  — чистый речевой сигнал, а  $d(n)$  — суммарный шумовой сигнал. Авторы используют  $Y(\lambda, k)$ , т. е. STFT для  $y(n)$ , чтобы представить суммарную величину канала  $k$  в кадре  $\lambda$  в частотной области. Спектр мощности шумового сигнала можно выравнивать и обновлять покадрово, используя представленную далее рекурсию:

$$P(\lambda, k) = \eta P(\lambda - 1, k) + (1 - \eta) |Y(\lambda, k)|^2 \quad (2)$$

где  $\eta$  — коэффициент выравнивания. В таком случае локальный минимум спектра мощности в каждом канале можно отследить следующим образом:

$$P_{min}(\lambda, k) = \begin{cases} P(\lambda, k) & , P_{min}(\lambda - 1, k) \geq P(\lambda, k) \\ \gamma P_{min}(\lambda - 1, k) + \frac{1-\gamma}{1-\beta} (P(\lambda, k) - \beta P(\lambda - 1, k)), & P_{min}(\lambda - 1, k) < P(\lambda, k) \end{cases} \quad (3)$$

где  $P_{min}(\lambda, k)$  — локальный минимум спектра мощности зашумленной речи, а  $\gamma$  и  $\beta$  — постоянные параметры. Отношение спектра мощности зашумленной речи к его локальному минимуму можно рассчитать следующим образом:

$$S_r(\lambda, k) = \frac{P(\lambda, k)}{P_{min}(\lambda, k)} \quad (4)$$

Это соотношение сравнивается с пороговым значением  $T(\lambda, k)$  для определения вероятности наличия речи  $I(\lambda, k)$  с использованием приведенного далее критерия:

$$I(\lambda, k) = \begin{cases} 1, & S_r(\lambda, k) \geq T(\lambda, k) \\ 0, & S_r(\lambda, k) < T(\lambda, k) \end{cases} \quad (5)$$

где  $T(\lambda, k)$  — пороговое значение, которое динамически обновляется в соответствии с оцененным ОСШ предыдущего кадра. Стоит отметить, что это пороговое значение установлено в литературе на постоянном уровне, на основе пилотных данных авторы обнаружили, что динамические пороговые значения во время оценки показали большую эффективность, чем постоянные, поэтому было решено использовать динамические пороговые значения. Вероятность наличия речи  $I(\lambda, k)$  можно выровнять следующим образом:

$$K(\lambda, k) = \alpha K(\lambda, k) + (1 - \alpha) I(\lambda, k) \quad (6)$$

где  $K(\lambda, k)$  — выровненная вероятность наличия речи, а  $\alpha$  — константа выравнивания. Коэффициент выравнивания, который будет использоваться для оценки шума, может быть обновлен с использованием вычисленной выше вероятности наличия речи:

$$\alpha_s(\lambda, k) = \alpha_d + (1 - \alpha_d) K(\lambda, k) \quad (7)$$

где  $\alpha_s$  — коэффициент выравнивания, который будет использоваться для оценки шума, а  $\alpha_d$  — константа. Наконец, мощность шума каждого канала оценивается следующим образом:

$$D(\lambda, k) = \alpha_s(\lambda, k) D(\lambda - 1, k) + (1 - \alpha_s(\lambda, k)) |Y(\lambda, k)|^2 \quad (8)$$

### Функция усиления для шумоподавления

Используя полученную в рамках оценки мощность шума, ОСШ можно оценить согласно следующей формуле:

$$SNR(\lambda, k) = \delta SNR(\lambda - 1, k) + (1 - \delta) \left( \frac{P(\lambda, k)}{D(\lambda, k)} - 1 \right) \quad (9)$$

Затем используется функция усиления, например:

$$G(\lambda, k) = \frac{SNR(\lambda, k)}{SNR(\lambda, k) + 1} \quad (10)$$

Для максимального подавления шума можно дополнительно настроить усиление:

$$G_0(\lambda, k) = \begin{cases} g & , G(\lambda, k) < T_g \\ G(\lambda, k) & , G(\lambda, k) \geq T_g \end{cases} \quad (11)$$

где  $g$  — значение меньшей константы, а  $T_g$  — динамическое пороговое значение, определяемое ОСШ.  $T_g$  также является одним из ключевых факторов, определяющих чувствительность алгоритма.

Наконец, мощность сигнала после шумоподавления рассчитывается следующим образом:

$$S(\lambda, k) = G_0(\lambda, k) P(\lambda, k) \tag{12}$$

**Пример**

Пример работы eVoice в речевом шуме (РШ) при ОСШ +5 дБ показан на **Рисунке 2**. eVoice был внедрен с помощью стратегии кодирования РВП с выбором каналов 8 из 24 при частоте дискретизации 16 000 Гц. На **Рисунке 2** разными цветами показано сравнение мощности в восьмом канале, включая сигналы для чистой речи, зашумленной речи, обработанной речи и полученного в рамках оценки шума.



**РИСУНОК 1** | Блок-схема стратегии РВП (черн.) и eVoice (красн.)

**ЭКСПЕРИМЕНТ 1: СУБЪЕКТИВНОЕ ПРЕДПОЧТЕНИЕ И РАСПОЗНАВАНИЕ РЕЧИ В ШУМЕ**

Этот эксперимент был разработан для оценки четкости восприятия речи с помощью eVoice (далее «ШП1») по сравнению с другим АШП (далее «ШП2»), в котором для снижения шума использовалось наложение двоичной маски, а также стратегия РВП без АШП (обозначается как «РВП»). ШП2 использует тот же метод оценки шума, что и ШП1; метод описан в разделе «Оценка шума». После оценки шума ШП2 вычисляет ОСШ, которое используется для установки усиления. Поэтому, если ОСШ выше порогового значения, коэффициент усиления устанавливается на 1 (преобладает речевой сигнал), или задается небольшая константа, если ОСШ ниже такого значения (преобладает шум). ШП2 был выбран для сравнения, потому что он так же эффективен с вычислительной точки зрения, как eVoice, а метод идеального наложения двоичной маски был изучен в других системах КИ (Маугер и соавт. (Mauger et al.), 2012b). Четкость распознавания речи измерялась с помощью теста на распознавание речи в шуме и субъективной оценочной анкеты.

**Методы**

**Участники**

В этом эксперименте участвовали 11 опытных пользователей КИ (шесть женщин и пять мужчин) в возрасте от 20 до 59 лет (средний возраст = 41,2 года). Все они являлись постлингвально оглохшими взрослыми людьми, которым на одном ухе установили имплантат CS-10A и использовали звуковой процессор «Venus™» (т. е. процессор первого поколения),

запрограммированный в рамках стратегии РВП. В этом эксперименте вместо Venus<sup>TM</sup> был установлен звуковой процессор «Enduro<sup>TM</sup>». На пульте дистанционного управления есть опция, позволяющая выбрать активацию АШП (один вариант с шумоподавлением ШП1-eVoice, а другой — с наложением двоичной маски для шумоподавления ШП2-Binary Masking). Демографические данные отдельных участников представлены в **Таблице 1**. Родным языком участников был мандаринский диалект китайского языка, участникам оплачивали затраченное время и командировочные расходы. Перед началом эксперимента было получено письменное информированное согласие, все процедуры были одобрены этическим советом соответствующей местной организации.

### Процедуры и материалы

В этом эксперименте характеристики ШП1 и ШП2 оценивались сначала в рамках субъективной оценки, а затем в тесте на распознавание речи в шуме.

Субъективная оценка проводилась на протяжении 2 недель. В начале недели 1 участникам установили процессоры «Enduro<sup>TM</sup>», который был встроен в ШП1, и попросили протестировать его в домашних условиях в течение 1 недели. В данный период участники могли свободно включать и выключать ШП1 и использовать его в различных сценариях повседневного прослушивания.

В конце недели 1 были получены субъективные оценки с использованием анкеты, приведенной в **Таблице 2**. Аналогичные процедуры были проведены для ШП2 на неделе 2. Анкета состоит из восьми вопросов о различных повседневных сценариях процесса слушания. Для получения субъективных оценок участников для ШП1 или ШП2 в каждом сценарии процесса слушания после каждого недельного домашнего использования использовалась 5-балльная шкала: 2 — полностью согласен; 1 — согласен; 0 — нейтральная позиция; -1 — не согласен; -2 — категорически не согласен.

В тесте на распознавание речи в шуме мы использовали два типа шума (речевой шум (РШ) и гул голосов) с тремя ОСШ (5, 10 и 15 дБ) для сравнения трех алгоритмов (АШП, ШП1 и ШП2). Всего получился 21 тестовый блок (два типа шума × три ОСШ × три алгоритма + базовые значения трех алгоритмов в условиях тишины). Сначала в случайном порядке были проведены три базовых блока (три алгоритма в условиях тишины), а затем также в случайном порядке реализовывали оставшиеся 18 блоков. Авторы использовали предложения из двух опубликованных баз данных речи на мандаринском наречии китайского языка: тест на распознавание предложений свода документов многопрофильной больницы НОАК (Си и соавт. (Xi et al.), 2012) и тест на распознавание предложений свода документов Института внутренних исследований (Фу и соавт. (Fu et al.), 2011). Свод документов многопрофильной больницы НОАК состоит из 12 списков, каждый из которых включает в себя 11 предложений, при этом каждое предложение сформировано из шести–восьми ключевых слов. Свод документов исследовательского института состоит из 10 списков, каждый из которых содержит 10 фонетически сбалансированных предложений по семь слов. Все предложения были произнесены женщинами. Были использованы одиннадцать из 12 списков свода документов 301 и все списки из свода документов института.

Из-за ограниченного количества материалов разные списки из свода документов многопрофильной больницы и свода документов исследовательского института были случайным образом распределены по блокам для каждого участника по одному списку на каждый блок. Особое внимание было уделено тому, чтобы блоки каждого алгоритма использовали списки из того же свода документов. В каждом блоке предложения были представлены в случайном порядке, рассчитывался процент правильности слов. Стимулы подавались в звуконепроницаемой комнате через динамик, расположенный на расстоянии 1 м перед участником, на удобном уровне громкости (примерно 65 дБА). Тесты проводились с использованием программного обеспечения «QuickSTAR4TR», разработанного Qianjie Fu (Фонд Эмили Фу (Emily Fu Foundation), 2019).

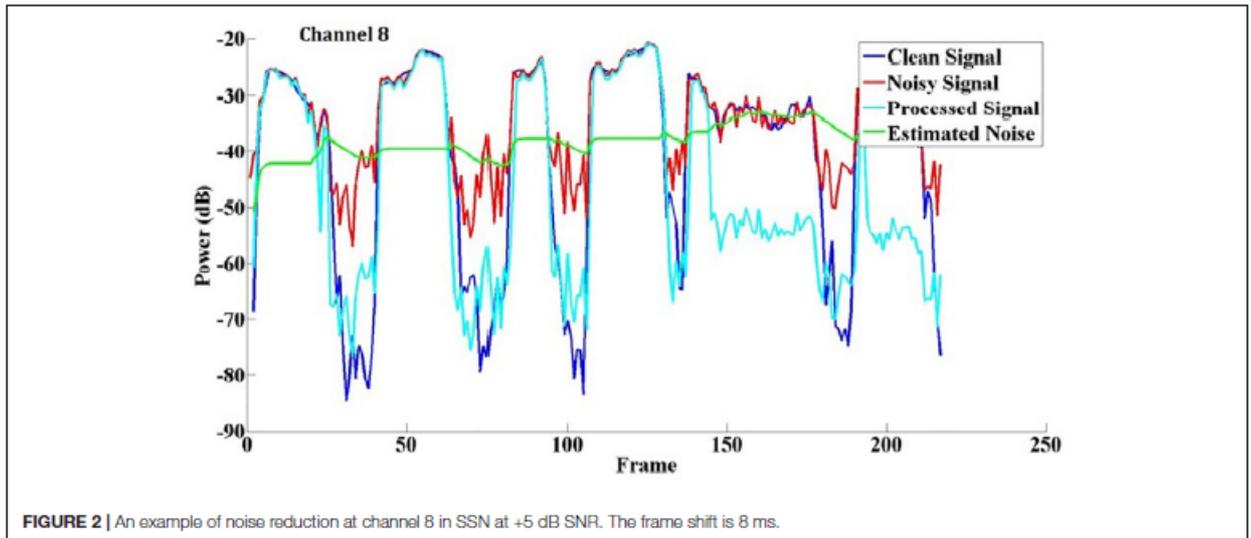


FIGURE 2 | An example of noise reduction at channel 8 in SSN at +5 dB SNR. The frame shift is 8 ms.

Clean Signal	Чистый сигнал
Noisy Signal	Шумный сигнал
Processed Signal	Обработанный сигнал
Estimated Noise	Оцененный шум
Channel 8	Канал 8
Power (dB)	Мощность (дБ)
Frame	Кадр

РИСУНОК 2 | Пример шумоподавления на канале 8 в РШ при ОСШ +5 дБ. Сдвиг кадра составляет 8 мс.

ТАБЛИЦА 1 | Демографические данные участников эксперимента 1.

Участник	Возрастная группа при тестировании	Этиология	Длительность глубокой глухоты на стороне уха с имплантатом (лет)	Ухо с имплантатом	Опыт КИ (лет)
№ 1	45–49	Внезапная глухота	10	Л	2
№ 2	25–29	Неизвестные причины	2	П	2
№ 3	20–24	Вызвано лекарственными средствами	17	П	4
№ 4	45–49	Неизвестные причины	4	П	2
№ 5	45–49	Вызвано лекарственными средствами	35	П	2
№ 6	45–49	Вызвано лекарственными средствами	28	П	4
№ 7	35–39	Вызвано лекарственными средствами	15	Л	6
№ 8	40–44	Внезапная глухота	0,25	Л	1,5
№ 9	55–59	Внезапная глухота	4	Л	3
№ 10	40–45	Внезапная глухота	8	П	6
№ 11	40–45	Вызвано лекарственными средствами	17	П	6

Сокращения: КИ — кохлеарный имплантат; Ж — женщина; Л — левое; М — мужчина; П — правое.

### Статистический анализ

Для анализа распознавания речи в тишине использовался односторонний дисперсионный анализ с многократными измерениями (ANOVA). Повторные измерения в рамках трехстороннего дисперсионного анализа были выполнены для оценки распознавания речи в условиях шума. Корректировки Бонферрони использовались для множественных сравнений.

## Результаты

### Субъективный оценочный тест

На **Рисунке 3** показаны результаты субъективных оценок для ШП1 и ШП2.

Для ШП1 (т. е. eVoice) было получено много положительных оценок и несколько отрицательных. Большинство участников положительно ответили на вопросы В2, В4, В5 и В6, что указывает на лучшие впечатления от прослушивания с включенным ШП1, чем с выключенным, в таких сценариях, как общение с несколькими людьми, на перекрестке и в автомобиле. На вопросы В3, В7 и В8 большинство участников ответили нейтрально, что соответствовало сценариям, например, в ресторане или супермаркете и рядом с кондиционером или вентилятором. Этот результат предполагает сопоставимую эффективность при включенном и выключенном ШП1 в этих настройках. Несколько участников положительно ответили на В3, В7 и В8 (впечатление лучше с включенным ШП1 в сценариях прослушивания, таких как беседа вдвоем в ресторане, рядом с кондиционером или вентилятором, или в оживленном супермаркете). Что касается прослушивания в условиях тишины, большинство участников сообщили, что настройки ШП1 не повлияли на беседу вдвоем в тишине, и положительно ответили на В1 (АШП не оказал влияния на беседу вдвоем в тихих помещениях).

Для ШП2 (наложения двоичной маски) обратная связь была более разнообразной. В целом оценки распределились практически равномерно между отрицательными и положительными для всех восьми вопросов, кроме В2 и В8, что означает, что некоторые участники решили, что ШП2 был полезен в большинстве сценариев прослушивания. Однако сопоставимое количество участников сочли его бесполезным или никак не влияющим. Большинство участников ответили нейтрально на В2 и В8, что указывает на то, что большинство посчитало ШП2 не оказывающим влияния на общение множества людей в тихой обстановке или при разговоре двух собеседников в супермаркете.

**ТАБЛИЦА 2** | Анкета, используемая для субъективной оценки.

В1	АШП не влияет на беседу вдвоем в тихих помещениях.
В2	АШП помогает во время разговоров нескольких людей (минимум трое говорящих) в тихих помещениях.
В3	АШП помогает при беседах вдвоем в ресторанах.
В4	АШП заметно подавляет шум или помогает общаться с другими, когда рядом проезжают транспортные средства.
В5	АШП помогает во время разговоров вдвоем или позволяет более четко услышать объявления на остановке в переполненном автобусе.
В6	АШП помогает во время разговоров вдвоем или обеспечивает более четкий звук радио в автомобиле.
В7	АШП помогает во время беседы вдвоем рядом с кондиционером или вентилятором.
В8	АШП помогает при беседах вдвоем в оживленном супермаркете.

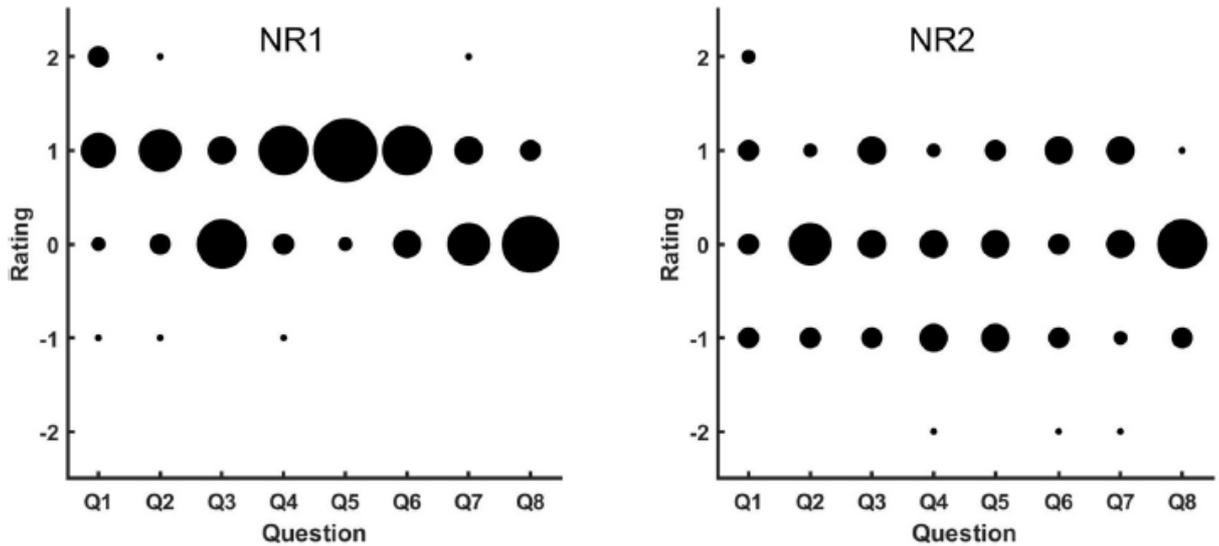
*АШП — алгоритм шумоподавления, который необходимо оценить (т. е. ШП1 на 1-й неделе или ШП2 на 2-й неделе).*

### Тест на четкость восприятия речи

Результаты распознавания речи в условиях тишины показаны на **Рисунке 4**. Односторонний дисперсионный анализ с повторными измерениями не выявил значимых различий между средними результатами (90%) трех алгоритмов ( $p = 0,452$ ).

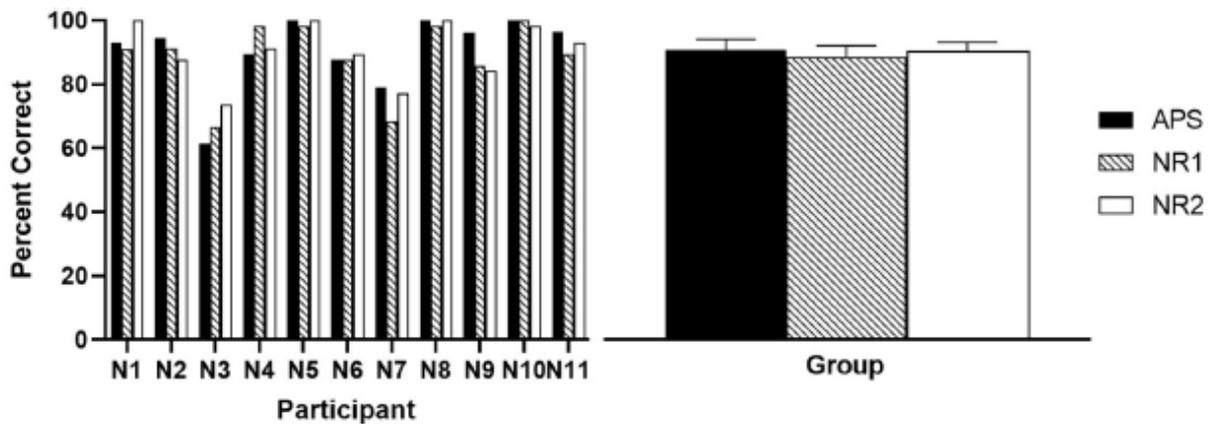
На **Рисунке 5** показаны результаты распознавания речи в РШ и гуле голосов. Статистическая значимость определялась с использованием ANOVA с процентном значением правильности восприятия речи в качестве зависимой переменной и типом шума (РШ или гул голосов), ОСШ (5, 10 или 15 дБ) и алгоритмом (РВП, ШП1 или ШП2) в качестве внутрисубъектных факторов. Тесты внутрисубъектных эффектов показали значительное воздействие типа шума ( $p = 0,022$ ), ОСШ ( $p < 0,001$ ) и алгоритма ( $p = 0,002$ ), а также значимую взаимосвязь между типом шума и ОСШ ( $p < 0,001$ ). Парные сравнения показали, что общая эффективность ШП1 была значительно выше, чем у АШП ( $p = 0,001$ ) и ШП2 ( $p = 0,016$ ), а значительной разницы между РВП и ШП2 ( $p = 0,612$ ) не было. Когда тип шума и ОСШ были зафиксированы для определения воздействия алгоритмов при специфических ОСШ в конкретном типе шума, ШП1 был значительно более эффективным, чем ШП2, при ОСШ 5 дБ в РШ ( $p = 0,010$ ), а также чем РВП ( $p = 0,027$ ) при ОСШ 5 дБ в гуле голосов. И для РШ, и для гула голосов три алгоритма не показали значительных различий при ОСШ 10 и 15 дБ. Однако более высокие средние баллы ШП1 можно было

наблюдать по сравнению с РВП и ШП2 при ОСШ 10 дБ в РШ (почти восемь процентных пунктов), а также при ОСШ 10 дБ (на восемь процентных пунктов выше, чем у РВП) и 15 дБ ОСШ (на 5 процентных пунктов выше, чем у РВП и ШП2) гуле голосов, хотя эти улучшения не были статистически значимыми.



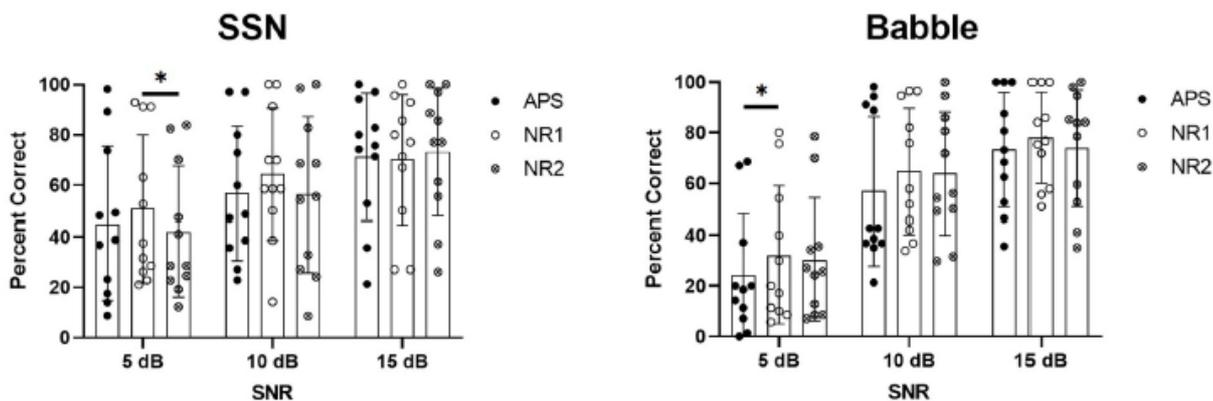
NR1	ШП1
NR2	ШП2
Rating	Оценка
Question	Вопрос

**РИСУНОК 3** | Результаты субъективных оценок ШП1 (слева) и ШП2 (справа). По оси абсцисс перечислены все восемь вопросов, используемых для субъективной оценки, по оси ординат — оценка, выставленная участниками. По оси ординат «-2» означает категорическое несогласие с утверждением, а «2» — абсолютное согласие. Чем больше число, тем более позитивным является субъективная оценка того, что ШП (шумоподавление) может помочь в различных сценариях шумной обстановки и не повлияет на прослушивание в условиях тишины. Размер кружков соответствует количеству участников, которые дали такие оценки, более крупные кружки указывают на большее количество участников.



APS	РВП
NR1	ШП1
NR2	ШП2
Group	Группа
Participant	Участник

**РИСУНОК 4** | Результаты оценки распознавания речи в тихой обстановке. Слева показано индивидуальное процентное значение правильности восприятия, а справа — средние значения группы с планками погрешностей, указывающими стандартную ошибку средних значений группы.



SSN	РШ
Babble	Гул голосов
Percent Correct	Процентное значение правильности восприятия речи
APS	РВП
NR1	ШП1
NR2	ШП2
SNR	ОСШ
dB	дБ

**РИСУНОК 5** | Результаты оценки распознавания речи в РШ (слева) и гуле голосов (справа). Результаты каждого участника представлены на графике, столбцы показывают средние значения, а планки погрешностей выдают стандартные отклонения.

### Резюме

В этом эксперименте авторы протестировали два АШП: eVoice (ШП1) и еще один алгоритм, в котором использовалось наложение двоичной маски (ШП2). В них применяется один и тот же процесс оценки шума, но различаются процессом шумоподавления. ШП1 использует функцию выравнивания усиления, а ШП2 — наложение двоичной маски. Субъективные оценки показывают, что ШП1 был воспринят положительно, а значение ШП2 в баллах было практически равномерно распределено от отрицательного к положительному восприятию с небольшим преобладанием нейтральных ответов. Результаты теста на распознавание речи указывают на общую более высокую эффективность ШП1 по сравнению с ШП2 и РВП. Однако значительное преимущество было выявлено только при ОСШ 5 дБ. Приведенные выше результаты демонстрируют, что ШП1 продемонстрировал большую эффективность, чем ШП2 по тестам на распознавание речи и по субъективным оценкам.

## ЭКСПЕРИМЕНТ 2: ТЕСТ НА ПОРОГОВЫЕ ЗНАЧЕНИЯ РАСПОЗНАВАНИЯ РЕЧИ

### Обоснование

Предполагаемое значительное преимущество eVoice не всегда подтверждается результатами первого эксперимента. Одна из причин может заключаться в процедуре с фиксированным ОСШ и большим разбросом эффективности в когорте. По результатам эксперимента 1 (слева на **Рисунке 5**) авторы заметили, что эффект верхнего предельного значения можно обнаружить у ряда пациентов при ОСШ 1 дБ, а эффект нижнего предельного значения можно наблюдать при ОСШ 5 дБ. Речевое восприятие в шумной обстановке резко изменялось среди участников, даже при одинаковом ОСШ и аналогичном шуме. Это указывает на предел процентного значения правильности восприятия речи при фиксированных ОСШ, т.к. данный тип испытания не способствует исключению возможных эффектов верхнего и нижнего предельных значений. Чтобы преодолеть это ограничение, авторы разработали эксперимент 2, в котором использовался адаптивный тест порогового значения восприятия речи (ПВР) для измерения потенциальных преимуществ eVoice. В первом эксперименте авторы четко показали, что ШП1 (eVoice) обеспечивает большую эффективность, чем ШП2 (идеальный бинарный алгоритм) в субъективном

тесте, хотя наблюдается небольшое улучшение в тесте распознавания речи в шумной обстановке. Чтобы дополнительно изучить потенциал eVoice и сэкономить время эксперимента, во эксперименте 2 оценивался только ШП1.

## Методы

### Участники

Для этого эксперимента были приглашены восемь опытных пользователей КИ (пять женщин и трое мужчин в возрасте от 23 до 62 лет, среднее значение — 43,6 лет). Для всех мандаринский диалект китайского языка являлся родным языком. Все они постлингвально оглохли во взрослом возрасте, им установили имплантат CS-10A на одно ухо, они использовали устройства Enduro™ в качестве клинических процессоров, запрограммированных стратегией кодирования РВП с возможностью дистанционного управления для включения или выключения eVoice.

Демографические данные отдельных участников представлены в **Таблице 3**. Участникам было оплачено время и командировочные расходы. Перед началом эксперимента от всех было получено письменное информированное согласие, все процедуры были одобрены этическим советом соответствующей местной организации.

### Процедуры и материалы

Для дальнейшей оценки эффективности eVoice в тесте на шум был применен адаптивный ступенчатый ПВР. Этот метод измерения ПВР был адаптирован из предыдущих исследований авторов (Менг и соавт. (Meng et al.), 2016, 2019) с двумя небольшими изменениями: (1) время представления стимула было сокращено с трех максимум до двух, и (2) пороговое значение оценки правильности было изменено с 50% слов в предложении до 80% слов. Первое было сделано для сокращения времени эксперимента. Второе потребовалось для отслеживания более высокого порогового значения, что более показательнее для истинного понимания. Таким образом, авторы фактически отслеживали пороговое значение, при достижении которого испытуемые имеют 50% шанс воспринять 80% правильной информации.

Был использован свод документов теста для восприятия на слух мандаринского диалекта китайского языка (MHINT) (Вон и соавт. (Wong et al.), 2007), начитанный одним мужчиной. Всего 12 списков для формальных тестов и 2 списка для практики по 20 предложений в каждом списке и 10 слов в каждом предложении. В этом эксперименте 10 из 12 списков формальных тестов использовались в качестве целевой речи в формальных тестах, и оба списка для практики применялись на этапе обучения для ознакомления участников с процедурами тестирования.

Были протестированы ПВР для каждой ситуации с eVoice и без него. Для каждой ситуации использовались два типа фонового шума: РШ и гул голосов, которые были сгенерированы с использованием метода, описанного в разделе «Речевые стимулы и задачи» эксперимента 2 в работе Менг и соавт. (Meng et al.) (2019). ПВР для каждой комбинации «ситуация-фоновый шум» был протестирован дважды с использованием двух разных списков MHINT, результаты были усреднены между двумя списками для получения окончательного ПВР. Четкость распознавания речи для каждой ситуации в условиях тишины также измерялась с помощью одного списка MHINT. Таким образом, для тестирования было использовано всего 10 списков (два фоновых × две ситуации × два комбинированных списка + два списка для четкости восприятия речи в условиях тишины). Порядок списков и ситуаций был случайным для участников. Перед формальным тестом использовались два практических списка, чтобы ознакомить участников с процедурами тестирования на ПВР и четкость восприятия речи в тихом помещении. Во время теста каждое предложение звучало не более двух раз по запросу участников; участникам было предложено повторить слова, которые образуют предложение со смыслом, обратной связи не было.

ОСШ в каждом испытании адаптировали путем изменения уровня целевой речи с фиксированным фоновым шумом. Участники должны были повторить столько слов, сколько они смогут, и целевой уровень был снижен, если не менее 8 слов повторялись правильно; в противном случае целевой уровень повышался. Размер шага составлял 8 дБ перед вторым воспроизведением, затем 4 дБ перед четвертым воспроизведением и 2 дБ для остальных. Вычислялось среднее

арифметическое ОСШ последних восьми предложений и фиксировалось как итоговое значение ПВР.

Стоит отметить, что гул голосов, использованный в этом исследовании, состоял из голосов того же говорящего, что и целевая речь (Менг и соавт. (Meng et al.), 2019), что стало чрезвычайно сложной задачей для любого АШП. Дополнительная информация о процедурах и материалах представлена в работе Менг и соавт. (2016, 2019).

## Результаты

В этом эксперименте участвовали восемь пользователей КИ, указанных в **Таблице 3**, но у № 17 была выявлена слуховая нейропатия. Поэтому, данные участника № 17 были исключены из анализа.

Результаты распознавания речи в условиях тишины показаны на **Рисунке 6**. Среднее значение в баллах по группе составило 93,1 и 93,3% при выключенном eVoice и включенном eVoice соответственно.

Двусторонний парный критерий Стьюдента не показал существенных различий между двумя ситуациями ( $t(6) = -0,162$ ,  $p = 0,877$ ).

На **Рисунке 7** показаны результаты ПВР для РШ (слева) и гула голосов (справа). При РШ у каждого участника были более низкие значения ПВР при включенном eVoice, чем при выключенном eVoice. Средние значения ПВР составили 7,9 и 5,7 дБ при выключенном eVoice и включенном eVoice соответственно. Эта разница в 2,2 дБ стала статистически значимым улучшением ( $t(6) = 6,892$ ,  $p < 0,001$ ). При гуле голосов среднегрупповые значения ПВР составили 10,9 дБ при выключенном eVoice и 10,7 дБ при включенном eVoice, соответственно. Двусторонний парный критерий Стьюдента не показал существенных различий между двумя ситуациями ( $t(6) = 0,249$ ,  $p = 0,812$ ).

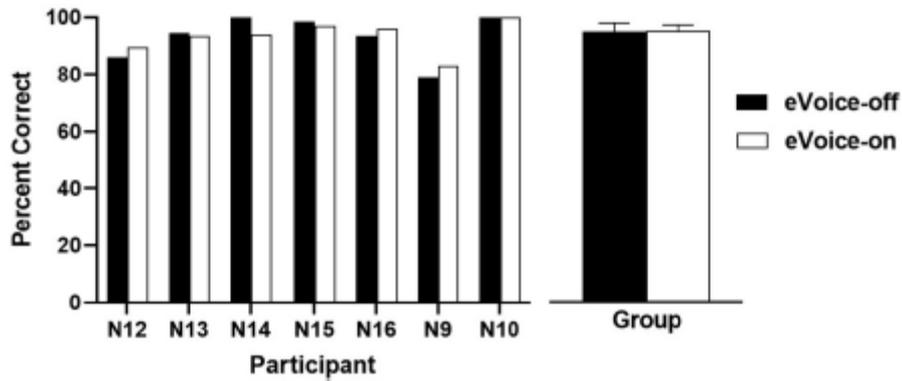
## Резюме

Цель этого эксперимента заключалась в количественной оценке преимущества eVoice для обеспечения четкости восприятия речи и исключения потенциальных эффектов верхнего и нижнего пороговых значений. Четкость восприятия речи измерялась с помощью адаптивного ПВР-теста с двумя фоновыми условиями: РШ и гул голосов. Значительного различия в скорости распознавания речи в тихой обстановке выявлено не было. Этот результат показывает, что eVoice не повлияет на восприятие речи в условиях тишины. Использование eVoice привело к снижению значений ПВР на 2,2 дБ при РШ, при этом значения ПВР значительного не изменились при гуле голосов.

ТАБЛИЦА 3 | Демографические данные участников эксперимента 2.

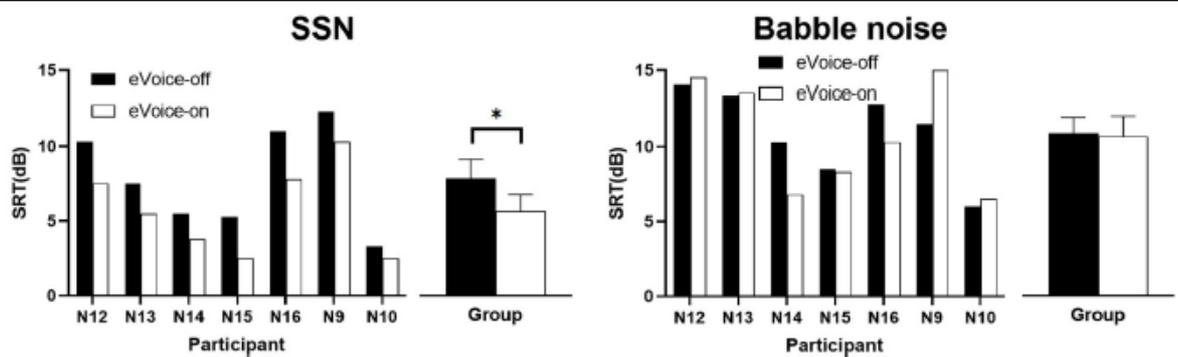
Участник	Возрастная группа при тестировании	Этиология	Длительность глубокой глухоты на стороне уха с имплантатом (лет)	Ухо с имплантатом	Опыт КИ (лет)	Опыт использования Enduro (лет)
№ 12	20–24	Вызвано лекарственными средствами	1	П	6	1
№ 13	50–54	Неизвестные причины	0,4	Л	6	0,5
№ 14	30–34	Внезапная глухота	1	П	1,5	1
№ 15	45–49	Неизвестные причины	5	Л	7	3
№ 16	50–54	LAVS	1	Л	6	1
№ 9	60–64	Внезапная глухота	4	Л	6	3
№ 10	45–49	Внезапная глухота	8	П	9	3
№ 17	35–39	Неизвестные причины	7	Л	8	0,5

Сокращения: КИ — кохлеарный имплант; Ж — женщина; Л — левое; LAVS — синдром расширенного водопровода преддверия внутреннего уха; М — мужчина; П — правое.



Percent Correct	Процентное значение правильности восприятия речи
Participant	Участник
Group	Группа
eVoice-off	eVoice выключен
eVoice-on	eVoice включен

**РИСУНОК 6** | Оценка распознавания речи в условиях тишины с включенным и выключенным eVoice. Слева показаны индивидуальные оценки, а справа — средние значения группы с планками погрешностей, указывающими стандартную ошибку средних значений группы.



Percent Correct	Процентное значение правильности восприятия речи
Participant	Участник
Group	Группа
eVoice-off	eVoice выключен
eVoice-on	eVoice включен
SSN	РШ
Babble noise	Гул голосов
SRT (dB)	ПВР (дБ)

**РИСУНОК 7** | Результаты ПВР для РШ (слева) и гула голосов (справа). Индивидуальные ПВР показаны слева, а среднегрупповые ПВР - справа. Планки погрешностей показывают стандартную ошибку средних значений группы. Существенное различие показано звездочкой ( $p < 0,05$ ).

### ОБСУЖДЕНИЕ

В этом исследовании авторы изучали eVoice, первый метод подавления шума в КИ «Nurotron®». eVoice — это одноканальный АШП, реализованный в рамках стратегии РВП в процессоре «Enduro». Для оценки эффективности этого алгоритма было проведено два эксперимента. Сначала эффективность eVoice сравнивалась с другим методом наложения двоичной маски в тесте на распознавание речи, а также подвергалась субъективной оценке в

эксперименте 1 ( $N = 11$ ). eVoice работает немного лучше, чем АШП с наложением двоичной маски. Затем был проведен более показательный адаптивный ПВР-тест для количественной оценки воздействия шумоподавления eVoice на четкость восприятия речи в эксперименте 2 ( $N = 7$ ). При сравнении эффективности с включенным и выключенным eVoice, было выявлено преимущество ПВР на 2,2 дБ при стационарном шуме и отсутствие различия в условиях тишины и при нестационарном шуме.

По сравнению с другими одноканальными АШП, реализованными в стратегиях КИ, eVoice имеет сопоставимую эффективность с представленными в литературных источниках. Например, одноканальный АШП, реализованный в стратегии ACE, показал преимущество ПВР до 2,14 дБ при стационарном шуме (Досон и соавт. (Dawson et al.), 2011). ClearVoice, реализованный в стратегии HiRes 120, использовал временное окно в 1,3 секунды для оценки шума и дало позволило повысить процентное значение правильности восприятия речи до 24 процентных пунктов (Бухнер и соавт. (Buechner et al.), 2010). Это может привести к уменьшению ПВР на 1,3–3,4 дБ, согласно литературным данным, при этом для стандартных речевых материалов уменьшение ПВР на 1 дБ приводит к увеличению процентного значения правильности восприятия речи на 7–19 процентных пунктов (Мур (Moore), 2007). Однако о значительных преимуществах при нестационарном шуме в литературе сообщается редко, что может указывать на ограничение традиционных одноканальных АШП. Необходимо разработать более инновационные методы для улучшения восприятия речи при нестационарном шуме для пользователей КИ.

Эта статья важна с точки зрения людей, которым были установлены имплантаты, и сурдологов. С точки зрения новой системы с быстро растущим количеством пользователей этот отчет по eVoice полезен для понимания системы и нового метода шумоподавления. Пользователей и сурдологов очень волнуют два вопроса о новом АШП в КИ: (1) действительно ли данный АШП работает при различных типах шумов; и (2) в какой степени пользователи могут воспользоваться его преимуществами. Полученные результаты показывают, что eVoice может улучшить четкость восприятия речи при стационарном шуме и не влияет на восприятие речи в тихом помещении и при нестационарном шуме. Это связано с тем, что eVoice является моноуральным алгоритмом оценки ОСШ, который предполагает, что шум относительно стационарен по сравнению с речью. Авторы обнаружили, что некоторые пользователи процессора Enduro, возможно, не заметили существования этого АШП, и их сурдологи могут посоветовать или напомнить им включить eVoice, чтобы улучшить качество восприятия речи в условиях шума.

Данная статья также может побудить людей переосмыслить управление шумом в системах КИ. Исследователям следует учитывать допущения о направленности и сложности нелинейных структур, которые могут быть смоделированы с помощью вычислений посредством обработки сигналов или машинного обучения (см., например, Бьянко и соавт. (Bianco et al.), 2019; Гон и соавт. (Gong et al.), 2019). Более ранние исследования и настоящая работа оказывают значительную поддержку в оптимизации и методов подавления шума для улучшения распознавания речи в условиях шума для пользователей КИ.

### **ЗАЯВЛЕНИЕ О ДОСТУПНОСТИ ДАННЫХ**

Серии данных, созданные для этого исследования, доступны по запросу, направленному отвечающему за переписку автору.

### **ЗАЯВЛЕНИЕ О СОБЛЮДЕНИИ ЭТИЧЕСКИХ ТРЕБОВАНИЙ**

Исследования с участием людей были проанализированы и одобрены Комитетом по медицинской этике Шэньчжэньского университета. Пациенты/участники предоставили письменное информированное согласие на участие в этом исследовании.

### **УЧАСТИЕ АВТОРОВ**

Все авторы участвовали в создании и разработке анализа, а также подготовке рукописи. ХЧ и НВ осуществляли сбор данных.

---

### **ФИНАНСИРОВАНИЕ**

Эта работа проводилась при поддержке Национального фонда естественных наук Китая (National Natural Science Foundation of China) (11704129, 11574090 и 61771320), Фонда естественных наук провинции Гуандун (Natural Science Foundation of Guangdong), Китай (2020A1515010386 и 2018B030311025) и Шэньчжэньского фонда науки и инноваций (Shenzhen Science and Innovation Funds) (JCYJ 20170302145906843).

### **ВЫРАЖЕНИЕ БЛАГОДАРНОСТИ**

Мы выражаем благодарность всем участникам с КИ за их терпение и сотрудничество в рамках этого исследования. Мы хотели бы поблагодарить Вэньхэ Ту (Wenhe Tu), Суй Хуан (Sui Huang), Пейао Ван (Peiyao Wang), Кэрол Пэн (Carol Peng), Шаньсянь Гао (Shanxian Gao) и Цзяци Чжан (Jiaqi Zhang) из компании «Nurotron» за помощь при сборе данных. Мы также выражаем благодарность двум рецензентам за полезные комментарии.