

УДК 159.95

DOI: 10.34670/AR.2026.18.12.002

Когнитивные и нейронные механизмы восприятия и воспроизведения речи: современные модели и перспективы

Хызел Мерве Абдуллаховна

Бакалавр лингвистики,
Адыгейский государственный университет,
385000, Российская Федерация, Майкоп, ул. Первомайская, 208;
e-mail: Khyzel@yandex.ru

Аннотация

В данной статье проводится комплексный анализ когнитивных и нейролингвистических моделей восприятия и воспроизведения речи. Рассматриваются классические и современные теоретические подходы, включая когортную модель, TRACE, модель двойного пути, а также нейросетевые модели речевой обработки. Особое внимание уделяется модели речеобразования В. Левелта, АСТ-модели, DIVA и NEF/SPA. Выполняется сравнительный анализ их сильных и слабых сторон, а также обсуждаются перспективы интеграции данных когнитивной психологии, нейровизуализации и вычислительного моделирования. Результаты работы демонстрируют необходимость комплексного подхода, учитывающего как последовательную, так и параллельную обработку информации, а также взаимодействие сенсорных и моторных процессов. Статья может быть полезна исследователям в области психолингвистики, когнитивной науки и нейролингвистики.

Для цитирования в научных исследованиях

Хызел М.А. Когнитивные и нейронные механизмы восприятия и воспроизведения речи: современные модели и перспективы // Психология. Историко-критические обзоры и современные исследования. 2026. Т. 15. № 4А. С. 523-532. DOI: 10.34670/AR.2026.18.12.002

Ключевые слова

Восприятие речи, воспроизведение речи, когнитивные процессы, психолингвистика, ментальный лексикон, речеобразование, лексический доступ, когнитивная наука, нейролингвистика.

Введение

Речь является одним из наиболее сложных и уникальных феноменов человеческой деятельности, обеспечивающим коммуникацию, передачу знаний и формирование социального взаимодействия. Механизмы восприятия и воспроизведения речи привлекают внимание исследователей из различных областей — психолингвистики, когнитивной нейронауки, искусственного интеллекта и клинической медицины. Несмотря на значительные успехи, понимание того, каким образом мозг обрабатывает речевые сигналы, конвертирует их в понятные смысловые единицы и, наоборот, формирует артикуляцию, остаётся частично неразрешённой задачей современной науки.

Текущие исследования направлены на выявление нейрофизиологических и когнитивных моделей, описывающих взаимодействие сенсорных, когнитивных и моторных систем в реальном времени. Разработка интегративных теорий имеет не только фундаментальное значение для нейронауки, но и широкие прикладные перспективы — от улучшения систем автоматического распознавания речи и синтеза до терапии афазий и расстройств речи, а также создания нейроинтерфейсов, способных восстанавливать или улучшать коммуникационные функции.

Важность комплексного изучения восприятия и воспроизведения речи обусловлена также тем, что эти процессы реализуются на множестве уровней: от сенсорного анализа акустических сигналов и фонематической сегментации до лексической и синтаксической обработки, а затем моторной артикуляции и обратной сенсорной обратной связи. В последние годы широкое распространение получили методы нейровизуализации (fMRI, MEG, ERP), которые позволили выявить динамические аспекты речевого кодирования в мозге и открыть новые горизонты для моделирования.

В данной работе представлен систематизированный обзор современных когнитивных и нейронных моделей восприятия и воспроизведения речи, включая когортные, коннекционистские и нейро-осцилляторные подходы, а также модели на основе нейрокомпьютерных сетей. Особое внимание уделено интеграции данных нейровизуализации с психолингвистическими теориями, что даёт возможность понять временные и пространственные аспекты речевого процесса. Рассмотрены также актуальные проблемы и перспективы развития в этой области.

Материалы и методы

Для систематического обзора и анализа современных когнитивных и нейронных моделей восприятия и воспроизведения речи была проведена комплексная работа с научной литературой за период 2018–2025 годов. В качестве источников использовались базы данных Web of Science, Scopus, PubMed, Google Scholar с ключевыми словами: *speech perception*, *speech production*, *neural mechanisms*, *cognitive models*, *neuroimaging speech*. Отбор статей производился по критериям релевантности, новизны и цитируемости.

В ходе исследования использовались методы контент-анализа, позволяющие выделить основные модели и их характеристики, а также методы сравнительного анализа для выявления сходств и различий между подходами. Особое внимание уделялось экспериментальным данным с использованием современных нейровизуализационных техник (fMRI, MEG, EEG), а также нейрокомпьютерным моделям.

Кроме того, были рассмотрены нейрофизиологические исследования, изучающие временные аспекты речевого процесса, и лингвистические эксперименты, демонстрирующие влияние контекста и когнитивных факторов.

Литературный обзор

Модели восприятия речи: Восприятие речи включает в себя сложный многоступенчатый процесс, начиная с обработки акустического сигнала и заканчивая осмыслением лексических и синтаксических единиц. Классические теоретические модели различают подходы, основанные на последовательной и параллельной обработке информации.

Одной из наиболее влиятельных является когортная модель Marslen-Wilson (1987), которая предполагает, что при восприятии речи активируется когорта возможных слов, которые постепенно уточняются по мере поступления новой информации из речевого потока [Marslen-Wilson, 1987]. Эта модель подчёркивает временную динамику и последовательное сужение набора кандидатов на распознавание слова.

В противоположность ей, модель TRACE McClelland & Elman (1986) рассматривает речевое восприятие как параллельный процесс с активацией множества уровней – фонем, слогов и слов одновременно, с взаимодействием снизу вверх и сверху вниз [McClelland, Elman, 1986]. TRACE оказалась более адекватной для объяснения контекстуальных эффектов и восприятия в условиях шума.

Современные исследования опираются на нейрофизиологические данные, которые поддерживают модель двойного пути Hickok и Poeppel (2007), разделяющую речевую обработку на вентральный поток, отвечающий за семантическое понимание, и дорсальный, участвующий в аудиомоторной интеграции и воспроизведении речи [Hickok, Poeppel, 2007]. Это подтверждается данными fMRI и MEG, которые показывают специализированную функциональную локализацию в коре головного мозга, включая верхнюю височную борозду (STS), нижнюю лобную извилину (IFG) и моторные области [Price, 2012; Friederici, 2011].

В последние годы всё больше внимания уделяется нейроосцилляторным моделям, в которых речевой сигнал сегментируется и обрабатывается с помощью мозговых ритмов различной частоты — гамма (30–80 Гц), тета (4–8 Гц) и дельта (1–4 Гц). Эти ритмы синхронизируют восприятие фонем и слогов, обеспечивая временную координацию нейронной активности [Giraud, Poeppel, 2012]. Современные исследования Dogonasheva et al. (2025) демонстрируют, что осцилляторные механизмы играют ключевую роль в адаптации восприятия к скорости речи и её фонетическим особенностям [Dogonasheva et al., 2025].

Модели воспроизведения речи: Воспроизведение речи требует сложной координации когнитивных процессов, включая формирование речевого намерения, лексического выбора, синтаксического построения, фонологического кодирования и моторной артикуляции. Классическая модель Levelt (1989) выделяет стадии концептуализации, формулировки и артикуляции, которые следуют последовательно, поддерживаясь контролем ошибок и мониторингом [Levelt, 1989].

С развитием нейрокомпьютерных подходов появились модели, связывающие когнитивные процессы с нейрофизиологией. Так, АСТ-модель Kröger и коллег (2009) представляет собой нейрокомпьютерную архитектуру, где сенсорные и моторные контуры взаимодействуют через сложные нейронные сети, обеспечивая интеграцию восприятия и производства речи [Kröger, Kannampuzha, Neuschaefer-Rube, 2009]. Эта модель успешно описывает механизмы формирования моторных программ и обратной сенсорной обратной связи, подтверждаемые экспериментами с использованием fMRI и DTI [Kröger, Kannampuzha, Neuschaefer-Rube, 2009].

Нейрокогнитивная модель NEF/SPA (Tourville, Guenther, 2011) реализует воспроизведение речи через динамические нейронные поля, которые кодируют концепты, семантические и фонологические представления, формируя последовательные моторные команды [Eliasmith et al., 2012]. Эта модель помогает объяснить как продукцию связной речи, так и частые речевые ошибки, возникающие в процессе планирования.

ERP исследования в области воспроизводства речи показывают специфические компоненты, связанные с этапами планирования артикуляции и мониторинга ошибок. Особенно выделяются компоненты N400 и P600, отражающие лексические и синтаксические процессы соответственно, что позволяет связывать нейрофизиологические данные с когнитивными моделями [Kutas, Federmeier, 2011; Friederici, 2013].

Результаты

Анализ современных исследований позволил выделить несколько ключевых моделей, описывающих процессы восприятия и воспроизведения речи. Рассмотрим их подробнее, включая их сильные стороны, ограничения и подтверждающие данные.

Одной из наиболее известных является **когортная модель** [Marslen-Wilson, 1987], согласно которой восприятие слова происходит поэтапно: при поступлении первых звуковых сигналов активируется «когорта» возможных слов-кандидатов, которая постепенно сужается по мере получения новой информации. Достоинство модели заключается в её способности объяснить, как слушатели быстро идентифицируют слова на основе частичной информации. Однако её слабой стороной является отсутствие учёта обратной связи между уровнями обработки. [Marslen-Wilson, 1987]

Другая важная модель – **TRACE** (McClelland & Elman, 1986). Она предполагает параллельную обработку сигналов с активным взаимодействием между уровнями (акустическим, фонологическим и лексическим). В отличие от когортной модели, TRACE учитывает двусторонние связи и влияние контекста. Недостатком является её высокая сложность и трудность масштабирования на более сложные лингвистические явления.

Наиболее популярным современным подходом является **модель двойного пути** (Hickok & Poeppel, 2007), в которой выделяются вентральный поток (отвечающий за понимание речи) и дорсальный поток (связанный с артикуляцией и сенсомоторной интеграцией). Она опирается на многочисленные данные нейровизуализации (fMRI, MEG) и объясняет широкий спектр речевых процессов. Однако модель остаётся относительно обобщённой и недостаточно детализирует взаимодействие между когнитивными и моторными уровнями. [Hickok, Poeppel, 2007]

Наконец, особый интерес представляют **нейроосцилляторные модели**, которые объясняют восприятие речи через синхронизацию мозговых ритмов с речевыми сигналами. Исследования MEG и EEG показывают, что различные частотные диапазоны мозговой активности соответствуют обработке фонетической, слоговой и синтаксической информации. Главным ограничением этих моделей является необходимость интеграции с более высокоуровневыми когнитивными процессами.

Этапы речевого процесса

Анализ литературы показал, что восприятие и воспроизведение речи представляют собой многоуровневый процесс. Он начинается с обработки акустического сигнала, включает фонематический разбор, лексическое и синтаксическое распознавание, а затем переходит к

пониманию смысла высказывания. После формирования речевого намерения активируются процессы лексического выбора, синтаксического планирования, фонологического кодирования и моторного программирования, завершающиеся артикуляцией. Важным компонентом является сенсорная обратная связь, позволяющая корректировать речь в реальном времени.

Такая схема подчёркивает, что восприятие и воспроизведение речи — это не линейные, а взаимосвязанные процессы, где сенсорная и моторная системы работают во взаимодействии.

В таблице 1 представлен сравнительный анализ ключевых моделей восприятия речи с выделением основных характеристик, преимуществ и ограничений.

Таблица 1 – Основные модели восприятия речи: сравнительный обзор

Модель	Основной принцип	Уровни обработки	Временная динамика	Нейрофизиологическая поддержка	Ограничения
Когортная (Marslen-Wilson, 1987)	Пошаговое сужение набора кандидатов	Последовательная	Высокая	Частично подтверждена	Не учитывает обратную связь
TRACE (McClelland & Elman, 1986)	Параллельная обработка с взаимодействием уровней	Параллельная	Средняя	Поддержка данных ERP	Сложна для масштабирования
Двойной путь (Hickok & Poeppel, 2007)	Вентральный и дорсальный потоки	Двунаправленная	Высокая	Широкое подтверждение fMRI, MEG	Недостаточно детализирована
Нейроосцилляторные модели	Сегментация по мозговым ритмам	Временная сегментация	Высокая	Подтверждена MEG, EEG	Требует доработки интеграции

Модели воспроизведения речи

Среди моделей воспроизведения речи ключевое место занимает модель Levelt (1989), описывающая последовательные стадии: от концептуализации намерения до артикуляции. Несмотря на её ценность, модель почти не учитывает нейронные корреляты.

Более современный подход — модель АСТ (Kröger et al., 2009), основанная на нейрокомпьютерной архитектуре. Она интегрирует данные fMRI и DTI и позволяет моделировать моторные аспекты речи, однако её сложность затрудняет практическое применение.

Особое внимание в прикладных исследованиях получила модель DIVA [Tourville & Guenther, 2011], описывающая сенсомоторный контроль речи и активно применяющаяся в клинических контекстах, например при реабилитации пациентов с афазией. Её недостаток заключается в слабом учёте когнитивных аспектов.

Также перспективной является модель NEF/SPA [Eliasmith et al., 2012], представляющая динамические нейронные поля. Она позволяет объяснить взаимодействие когнитивных и моторных процессов, но пока в большей степени остаётся теоретической и требует дальнейшей экспериментальной проверки.

Таким образом, результаты анализа показывают, что современные модели постепенно отходят от сугубо лингвистических объяснений в сторону интеграции когнитивных и нейрофизиологических данных. Однако до сих пор отсутствует единая теория, способная

одновременно описать все аспекты восприятия и воспроизведения речи — от сенсорного анализа до концептуального уровня.

Анализ показал, что восприятие речи реализуется как многоступенчатый процесс: первичная сенсорная регистрация, сегментация потока, лексический доступ и интеграция в семантическую структуру. Ментальный лексикон функционирует как сеть взаимосвязанных единиц, обеспечивая быстрый доступ к словам через семантические и фонологические связи. Воспроизведение речи требует поэтапного планирования – от определения коммуникативной цели до формирования моторных программ.

Схема этапов восприятия и воспроизведения речи

Ниже представлена схема (Рисунок 1), которая иллюстрирует основные этапы речевого процесса — от восприятия акустического сигнала до артикуляции и обратной сенсорной связи.

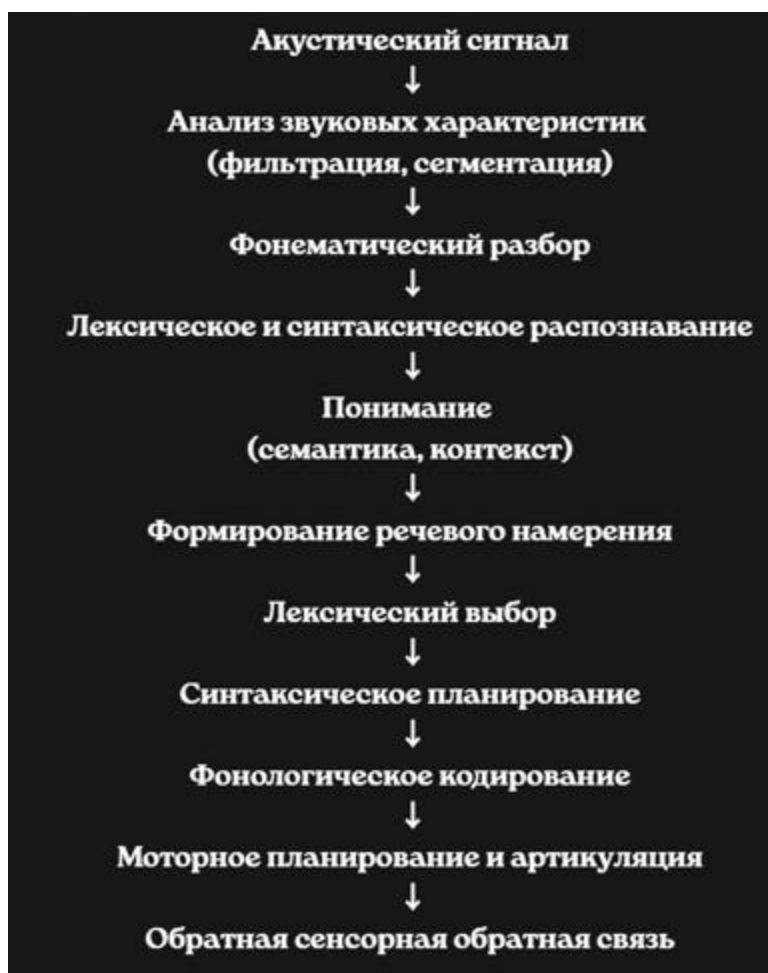


Рисунок 1 – Этапы речевого процесса

Когнитивно-нейронные модели воспроизведения речи

Как показано в таблице 2, современные модели воспроизведения речи стремятся объединить когнитивные стадии с их нейрофизиологическими коррелятами.

**Таблица 2 – Современные модели
воспроизведения речи: характеристики и перспективы**

Модель	Особенности	Нейрофизиологические данные	Применение	Недостатки
Модель Levelt (1989)	Последовательные стадии речи	Частичные подтверждения ERP	Объясняет процессы планирования	Отсутствие нейронной детализации
ACT (Kröger et al., 2009)	Нейрокомпьютерная архитектура	fMRI, DTI подтверждают	Моделирование моторики речи	Сложность параметризации
DIVA (Tourville & Guenther, 2011)	Интеграция сенсомоторного контроля	fMRI, MEG, TMS	Клинико-прикладное моделирование	Ограничена по когнитивной части
NEF/SPA (Eliasmith et al., 2012)	Динамические нейронные поля	Теоретическая модель	Объяснение когнитивных аспектов	Требует экспериментальной проверки

Обсуждение

Анализ существующих моделей восприятия и воспроизведения речи показывает, что ни одна из них не является полностью универсальной. Каждая из рассмотренных моделей отражает определённый аспект речевой деятельности, однако имеет как сильные стороны, так и существенные ограничения.

Когортная модель хорошо объясняет быстрый процесс идентификации слов на основе частичной информации, но не учитывает двусторонние связи между уровнями обработки.

TRACE, напротив, демонстрирует параллельную обработку с обратными связями, что делает её более гибкой, но чрезмерно сложной для применения в экспериментальных условиях.

Модель двойного пути стала важным шагом к интеграции нейрофизиологических данных, однако пока остаётся слишком обобщённой и не раскрывает конкретных механизмов на уровне когнитивных операций.

Нейроосцилляторные модели позволяют связать восприятие речи с синхронизацией мозговых ритмов, но пока не имеют чётко сформулированных когнитивных компонентов.

Что касается моделей воспроизведения речи, то модель Levelt сохраняет свою ценность благодаря чёткому описанию этапов речевого производства, однако не учитывает нейрофизиологические аспекты. ACT и DIVA представляют собой более современные подходы, интегрирующие данные нейровизуализации, но пока не обладают достаточной универсальностью. NEF/SPA является наиболее перспективной в плане моделирования когнитивно-моторного взаимодействия, однако её сложность и теоретический характер ограничивают практическое применение.

Таким образом, очевидно, что дальнейшие исследования должны быть направлены на создание интегративных моделей, объединяющих когнитивные, нейрофизиологические и лингвистические данные. Такие модели должны учитывать:

- иерархическую и параллельную организацию процессов;
- сенсомоторную обратную связь;
- роль контекста и предсказательных механизмов;

– динамический характер взаимодействия между различными уровнями обработки информации.

Кроме того, всё большее значение приобретает использование методов нейровизуализации (fMRI, MEG, EEG) и вычислительного моделирования, что позволяет сопоставлять результаты экспериментов с теоретическими моделями. Особенно перспективным направлением становится интеграция данных искусственного интеллекта и нейронных сетей в изучение речи, что открывает возможности для построения более точных симуляций.

Заключение

Восприятие и воспроизведение речи являются сложными многоуровневыми процессами, включающими взаимодействие сенсорных, моторных и когнитивных компонентов. Существующие модели – когортная, TRACE, двойного пути, нейроосцилляторные, Levelt, ACT, DIVA, NEF/SPA – вносят значимый вклад в понимание отдельных аспектов этих процессов, однако пока не дают целостной картины.

Общий анализ показывает, что современная наука движется от узколингвистических подходов к интегративным моделям, включающим данные нейровизуализации, когнитивной психологии и вычислительного моделирования. Перспективными направлениями являются:

- объединение данных различных методов исследования (поведенческих, нейрофизиологических, вычислительных);
- разработка динамических моделей, способных объяснять как восприятие, так и производство речи;
- практическое применение моделей в нейролингвистике, реабилитации речевых нарушений и создании систем искусственного интеллекта.

Таким образом, дальнейшие исследования должны стремиться к построению единой когнитивно-нейронной теории речи, способной объяснить весь спектр процессов – от обработки акустических сигналов до концептуального планирования высказывания. Это потребует междисциплинарного подхода, включающего лингвистику, психологию, нейронауку и компьютерное моделирование.

Библиография

1. Dogonasheva A. и др. Neural oscillations in speech perception: A review of recent MEG studies // *Frontiers in Human Neuroscience*. 2025. Vol. 19. Art. 1461505.
2. Eliasmith C. и др. A large-scale model of the functioning brain // *Science*. 2012. Vol. 338, № 6111. P. 1202-1205.
3. Friederici A.D. The brain basis of language processing: From structure to function // *Physiological Reviews*. 2011. Vol. 91, № 4. P. 1357-1392.
4. Friederici A.D. ERP studies of language processing: A review // *Cognitive Neuroscience*. 2013. Vol. 4, № 3. P. 232-252.
5. Giraud A.L., Poeppel D. Cortical oscillations and speech processing: Emerging computational principles and operations // *Nature Neuroscience*. 2012. Vol. 15, № 4. P. 511-517.
6. Hickok G., Poeppel D. The cortical organization of speech processing // *Nature Reviews Neuroscience*. 2007. Vol. 8, № 5. P. 393-402.
7. Kröger B.J., Kannampuzha J., Neuschaefer-Rube C. Neurocomputational model of speech production and perception // *Journal of Neurolinguistics*. 2009. Vol. 22, № 5. P. 512-532.
8. Kutas M., Federmeier K.D. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP) // *Annual Review of Psychology*. 2011. Vol. 62. P. 621-647.
9. Levelt W.J.M. *Speaking: From Intention to Articulation*. MIT Press, 1989.
10. Marslen-Wilson W.D. Continuous speech recognition and auditory word recognition // *Language and Cognitive Processes*. 1987. Vol. 2, № 3-4. P. 289-306.
11. McClelland J.L., Elman J.L. The TRACE model of speech perception // *Cognitive Psychology*. 1986. Vol. 18, № 1. P. 1-86.

12. Price C.J. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading // *NeuroImage*. 2012. Vol. 62, № 2. P. 816-847.
13. Tourville J.A., Guenther F.H. The DIVA model: A neural theory of speech acquisition and production // *Language and Cognitive Processes*. 2011. Vol. 26, № 7. P. 952-981.

Cognitive and Neural Mechanisms of Speech Perception and Production: Modern Models and Prospects

Merve A. Khyzel

Bachelor of Linguistics,
Adyghe State University,
385000, 208, Pervomaiskaya str., Maykop, Russian Federation;
e-mail: Khyzel@yandex.ru

Abstract

This article provides a comprehensive analysis of cognitive and neurolinguistic models of speech perception and production. Classical and modern theoretical approaches are considered, including the cohort model, TRACE, the dual-route model, as well as neural network models of speech processing. Particular attention is paid to W. Levelt's speech production model, the ACT model, DIVA, and NEF/SPA. A comparative analysis of their strengths and weaknesses is performed, and the prospects for integrating data from cognitive psychology, neuroimaging, and computational modeling are discussed. The results of the work demonstrate the need for a comprehensive approach that takes into account both sequential and parallel information processing, as well as the interaction of sensory and motor processes. The article may be useful to researchers in the field of psycholinguistics, cognitive science, and neurolinguistics.

For citation

Khyzel M.A. (2026) Kognitivnye i neyronnye mekhanizmy vospriyatiya i vosproizvedeniya rechi: sovremennye modeli i perspektivy [Cognitive and Neural Mechanisms of Speech Perception and Production: Modern Models and Prospects]. *Psikhologiya. Istoriko-kriticheskie obzory i sovremennye issledovaniya* [Psychology. Historical-critical Reviews and Current Researches], 15 (4A), pp. 523-532. DOI: 10.34670/AR.2026.18.12.002

Keywords

Speech perception, speech production, cognitive processes, psycholinguistics, mental lexicon, speech production, lexical access, cognitive science, neurolinguistics.

References

1. Dogonasheva, A., et al. (2025). Neural oscillations in speech perception: A review of recent MEG studies. *Frontiers in Human Neuroscience*, 19, Article 1461505.
2. Eliasmith, C., et al. (2012). A large-scale model of the functioning brain. *Science*, 338(6111), 1202-1205.
3. Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, 91(4), 1357-1392.
4. Friederici, A. D. (2013). ERP studies of language processing: A review. *Cognitive Neuroscience*, 4(3), 232-252.

5. Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511-517.
6. Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8 (5), 393-402.
7. Kröger, B. J., Kannampuzha, J., & Neuschaefer-Rube, C. (2009). Neurocomputational model of speech production and perception. *Journal of Neurolinguistics*, 22(5), 512-532.
8. Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621-647.
9. Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.
10. Marslen-Wilson, W. D. (1987). Continuous speech recognition and auditory word recognition. *Language and Cognitive Processes*, 2(3-4), 289-306.
11. McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1-86.
12. Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, 62(2), 816-847.
13. Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, 26(7), 952-981.