

===== ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ТЕХНИЧЕСКИХ =====  
===== И СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМАХ =====

## Избыточность управления процессами речеобразования

В.Н.Сорокин

*Институт проблем передачи информации, Российская академия наук, Москва, Россия*

Поступила в редколлегию 05.05.2009

**Аннотация.** Система управления артикуляцией характеризуется избыточностью, что позволяет достичь одних и тех же или близких значений параметров на каждом уровне за счет различных параметров на нижележащих уровнях. Сокращение различных двигательных единиц создает одно и то же мышечное сокращение, активность разных мышц приводит к одному и тому же смещению артикулятора или форме языка, одно и то же расстояние между подвижными и неподвижными частями речевого тракта в чувствительных местах обеспечивается смещением разных артикуляторов, а один и тот же вектор резонансных частот тракта соответствует бесконечному множеству функций площади поперечного сечения тракта. Избыточность наблюдается также и на уровне критериев оптимальности в системе управления артикуляцией. Конкретные примеры использования этой избыточности исследуются путем анализа патологии речеобразования и в специальных экспериментах по возмущению артикуляции и восприятия речи.

Управление артикуляцией обладает гибкостью, которая обеспечивает сохранение акустических параметров звуков данного языка в пределах, необходимых для понимания речи.

Размеры и форма речевого тракта у разных людей различны, так же, как различны и свойства мышц, управляющих артикуляторами. Если бы для понимания речи требовалось, чтобы акустические образы информативных элементов речи, например, фонем, были строго одинаковыми в любом контексте, то разнообразие физиологических и геометрических параметров речевых трактов разных людей не позволило бы удовлетворить этому условию, и речевое общение стало бы невозможным.

Гибкость управления артикуляцией обеспечивается возможностью достижения одного и того же акустического или артикуляторного образа путем использования различных управлений. Известно, например, что при поражениях мышечного аппарата нижней челюсти, в результате которых она практически неподвижна, наблюдается повышенная активность губ при артикуляции губных звуков. Наблюдаются и другие, иногда поразительные явления компенсации патологии артикуляторного аппарата.

Существуют разнообразные источники избыточности в процессах речеобразования. Рассмотрим сначала влияние площади поперечного сечения речевого тракта на акустические характеристики речевого сигнала. Пусть для функции площади поперечного сечения речевого тракта  $S(x)$  и соответствующего уравнения Вебстера

$$\frac{1}{S} \frac{\partial}{\partial x} \left( S \frac{\partial P}{\partial x} \right) - \frac{1}{c_0^2} \frac{\partial^2 P}{\partial t^2} = 0 \quad (1)$$

имеется множество собственных чисел  $\{\lambda_i^{(0)}\}$  и собственных функций  $\{\psi_i^{(0)}(\lambda_i^{(0)}, x)\}$ . Обозначим через  $S_0(x)$  исходную функцию площади поперечного сечения, подвергнем  $S_0(x)$  небольшому изменению и оценим его влияние на собственные числа  $\lambda_i^{(0)}$ , которые, как известно, определяют частоту резонансных колебаний. Пусть новая площадь поперечного сечения

$$S(x) = S_0(x) + S_1(x), \quad (2)$$

где  $S_1(x)$  — возмущение, причем  $S_1(x) \ll S_0(x)$  при любом  $x$ . Тогда можно применить метод Шредингера, и определить новые собственные числа  $\lambda_i^{(m)}$  и собственные функции для  $\psi_i^{(m)}$  для  $m$ -го приближения, представив неизвестные собственные функции возмущенной системы в виде разложения по собственным функциям невозмущенной системы

$$\psi_i^{(m)}(x) = \psi_i^{(0)}(x) + \sum_{m \neq i} \alpha_{mi} \psi_m^{(0)}(x), \quad (3)$$

где  $\alpha_{mi}$  — малая величина. Возмущенные значения собственных чисел представим как

$$\lambda_i^2 = \lambda_i^{(0)2} (1 + \eta_i), \quad (4)$$

где  $\eta_i$  — также малая величина.

Разделим переменные в уравнении (1) и получим систему

$$\begin{aligned} (SP')' + \lambda^2 SP &= 0; \\ T'' + \lambda^2 c_0^2 T &= 0. \end{aligned} \quad (5)$$

Подставим (2) - (4) в (5), пренебрежем членами второго порядка малости, и примем во внимание, что, по определению,

$$(S_0 P')' + \lambda^{(0)2} S_0 P = 0; \quad (6)$$

и, как следствие,

$$\left( S_0 \psi_n^{(0)'} \right)' + \lambda_n^{(0)2} S_0 \psi_n^{(0)} = 0. \quad (7)$$

Умножим каждый член получившегося уравнения на  $\psi_i^{(0)}$  и возьмем определенный интеграл на интервале от 0 до  $l$ . В силу ортогональности собственных функций  $\psi_k^{(0)}$  с весом  $S_0$  справедливо соотношение

$$\int_0^l S_0 \psi_n^{(0)} \psi_i^{(0)} dx = 0 \text{ при } i \neq n, \quad (8)$$

откуда легко можно получить

$$\int_0^l (S_0 \psi_m^{(0)'})' \psi_i^{(0)} dx = 0 \text{ при } i \neq m. \quad (9)$$

Тогда для поправки первого порядка имеем

$$\eta_i = -\frac{1}{\lambda_i^{(0)2}} \int_0^l \left[ S_1(x) \psi_i^{(0)'}(x) \right]' \psi_i^{(0)}(x) dx - \int_0^l S_1(x) \psi_i^{(0)2}(x) dx. \quad (10)$$

Возьмем первый интеграл по частям и получим

$$\eta_i = \frac{1}{\lambda_i^{(0)2}} \int_0^l S_1(x) \left[ \psi_i^{(0)'}(x) \right]^2 dx - \int_0^l S_1(x) \psi_i^{(0)2}(x) dx - S_1(x) \psi_i^{(0)}(x) \psi_i^{(0)'}(x) \Big|_0^l \quad (11)$$

или

$$\eta_i = \frac{a_i}{\lambda_i^{(0)2}} - b_i - c_i, \quad (12)$$

где

$$a_i = \int_0^l S_1(x) \left[ \psi_i^{(0)'}(x) \right]^2 dx, \quad (13)$$

$$b_i = \int_0^l S_1(x) \psi_i^{(0)2}(x) dx, \quad (14)$$

$$c_i = S_1(l) \psi_i^{(0)}(l) \psi_i^{(0)'}(l) - S_1(0) \psi_i^{(0)}(0) \psi_i^{(0)'}(0). \quad (15)$$

Отсюда видно, что поправка на собственные частоты может оказаться неизменной при разных значениях параметров  $a_i$ ,  $b_i$  и  $c_i$ . Кроме того, параметры  $a_i$ ,  $b_i$  могут оказаться одними теми же для бесконечного множества функций  $S_1(x)$ , если интегралы в (13) и (14) сохраняют одно и то же значение, а  $S_1(l)$  и  $S_1(0)$  постоянны или  $c_i=0$  в силу определенных граничных условий.

Таким образом, достаточно строго устанавливается, что один и тот же вектор собственных чисел, а с ним и вектор резонансных частот, может соответствовать бесконечному множеству функций площади поперечного сечения речевого тракта. Отсюда следует, что отображение пространства собственных чисел на пространство функций площади поперечного сечения тракта неоднозначно.

В свою очередь, и одни и те же значения площади поперечного сечения тракта могут достигаться разными смещениями артикуляторных органов.

Расстояние между губами определяется суммарным смещением нижней челюсти, верхней и нижней губы. При этом малое смещение одного артикулятора может компенсироваться более значительным смещением других артикуляторов. Выше упоминалось о компенсации паралича мышц нижней челюсти. У некоторых людей верхняя губа практически не принимает участия в образовании смычных согласных ( $n$ ,  $b$ ), назального ( $m$ ) или фрикативных звуков ( $\phi$ ,  $\psi$ ), и губная смычка или щель образуется за счет смещения нижней губы и нижней челюсти, а иногда и смещениями только нижней губы.

Вертикальные смещения нижней челюсти и кончика языка дополняют друг друга при формировании смычки или щели в передней области твердого неба. Кроме того, на расстояние между кончиком языка и твердым небом влияет и активность мышцы *styloglossus*, которая при определенных условиях поворачивает все тело языка относительно его корня так, что кончик языка приподнимается. Активность разных отделов мышцы *genioglossus* также может привести к подъему кончика языка.

Расстояние между кончиком языка и твердым небом может изменяться и за счет совместных движений нижней челюсти и смещении корня языка в горизонтальном направлении. Это связано с особенностями формы твердого неба вблизи верхних зубов.

Расстояние между поверхностью языка и небом складывается из вертикальных смещений корня языка, поворота языка как твердого тела и изменения формы языка в его передней части.

Расстояние между задней поверхностью тракта и языком определяется горизонтальным смещением корня языка, изменением формы языка и поворотом нижней челюсти относительно челюстного сустава.

Подъем нижней челюсти осуществляется активизацией трех мышц: жевательной (*masseter*), височной (*temporalis*) и внешней крыловидной. При опускании нижней челюсти действуют четыре мышцы: внутренняя крыловидная, двубрюшная (*digastric belly*), подбородочно-подъязычная и челюстно-подъязычная. Эти же семь мышц, но уже в других комбинациях, участвуют в движениях нижней челюсти вперед или назад в горизонтальном направлении. Как видно, движение нижней челюсти в любом направлении требует активности нескольких мышц, создавая, таким образом, избыточность управления и взаимозаменяемость действий мышц, а один и тот же артикуляторный акт (например, движение губ) может быть осуществлен за счет различной пространственно-временной активности группы мышц [1].

Рассматривая язык, как изогнутую пластину, уравнение его упругих деформаций под воздействием внешних и внутренних мышц можно представить как [2]

$$\frac{1}{R_0^2} \frac{\partial^2}{\partial \varphi^2} \left( EJ_z \frac{\partial^2 u}{\partial \varphi^2} + u \right) + cu + r \frac{\partial u}{\partial t} + \rho \frac{\partial^2 u}{\partial t^2} = - \frac{\partial^2 M}{\partial \varphi^2}, \quad (16)$$

где  $\varphi$  – угол в полярной системе координат,  $u$  – форма средней линии языка,  $E$  – модуль упругости мышечных тканей,  $J_z$  – момент инерции поперечного сечения языка,  $c$  – упругость подстилающих

тканей,  $r$  – вязкое сопротивление подстилающих тканей,  $\rho$  – плотность тканей языка,  $M$  – изгибающий момент, создаваемый внешними и внутренними мышцами. Все эти параметры являются функциями координаты вдоль средней линии языка.

Форма верхней поверхности языка определяется распределением изгибающего момента  $M(\varphi)$ , а этот момент, в свою очередь, создается совместным действием усилий мышц *styloglossus*, *genioglossus*, *longitudinalis inferior* и *longitudinalis superior*.

Как видно, в управлении формой языка возможно различное распределение активности разных мышц.

Имеются сообщения о том, что одна и та же степень сокращения мышцы может быть достигнута за счет возбуждения разных двигательных единиц [3].

Как было установлено в [4], разные виды критерия оптимальности обеспечивают примерно одинаковую точность управления. Это заставляет рассмотреть критерии оптимальности не на уровне движения артикулятора, а на уровне команд управления, что имеет более определенный физиологический смысл.

Предположим, что при управлении каждым артикулятором действует критерий минимума энергии команды управления  $f_k^2 \rightarrow \min$  для каждого момента времени. Под командой управления теперь понимается равнодействующая мышечных сил, приложенных к артикулятору. Это эквивалентно

$$[m_k \ddot{x}_k + r_k \dot{x}_k + c_k (x_k - x_k^0)]^2 \rightarrow \min. \quad (17)$$

Здесь легко видеть частные критерии, рассмотренные ранее в предположении доминирующей роли ускорения, скорости или смещения артикулятора:

$$[m_k \ddot{x}_k]^2 \rightarrow \min, [r_k \dot{x}_k]^2 \rightarrow \min \text{ или } [c_k (x_k - x_k^0)]^2 \rightarrow \min,$$

с той лишь разницей, что здесь масса, жесткость или коэффициент вязких потерь представлены в квадрате.

Известно, что управление сокращением двигательных единиц может выполняться как относительно степени сокращения (т.е. силы, развиваемой двигательной единицей), так и относительно скорости сокращения [5]. Примем для определенности, что управление артикуляторами описывается классом кусочно-линейных функций [6]

$$f_k = a_k + b_k(t - t_{ki}). \quad (18)$$

Здесь  $t_{ki}$  - время начала очередной команды. Параметр  $a_k$  в (18) может рассматриваться как аналог команды управления по величине силы мышечного сокращения, а параметр  $b_k$  - как аналог команды управления по скорости изменения силы мышечного сокращения. Такая математическая модель управлений позволяет провести аналитическое исследование свойств критериев оптимальности.

Если система управления движением осведомлена об интервале времени, на котором выполняется управление, то фактор времени может быть включен в критерий оптимальности. Исследования временных характеристик процессов речеобразования дают основание полагать, что система управления артикуляцией планирует речевое высказывание на определенном интервале времени  $T$ . В частности, при изменении контекста, темпа или стиля речи наблюдаются разнообразные явления изменения длительности отдельных сегментов [2, 7, 8, 9]. При планировании произвольных движений точное время исполнения может быть неизвестно, и выгоднее может оказаться использование мгновенных или кратковременных критериев. Это и было обнаружено в экспериментах [4].

Для одной команды, генерируемой в момент времени  $t_k=0$ , и поддерживаемой на интервале времени  $T_k$ , при кусочно-линейном управлении получим

$$\int_0^{T_k} (a_k + b_k t)^2 dt = T_k [a_k^2 + a_k b_k T_k + \frac{b_k^2 T_k^2}{3}] = T_k [(a_k + b_k T_k)^2 - a_k b_k T_k - \frac{2}{3} (b_k T_k)^2]. \quad (19)$$

Параметры  $a_k$  и  $b_k$  могут быть и меньше нуля:  $a_k$  - потому, что в этой команде объединены усилия как агонистов, так и антагонистов, создающих как бы "отрицательные" усилия, а  $b_k < 0$  для ко-

манд с падающим уровнем. В то же время, для каждой мышцы управление только положительное. Поэтому в критерии необходимо взять модуль значения  $a_k$ . Тогда имеем

$$\int_0^{T_k} (|a_k| + b_k t)^2 dt = T_k (a_k^2 + |a_k| b_k T_k + \frac{b_k^2 T_k^2}{3}) \quad (20)$$

и энергетический критерий оптимальности есть

$$\Omega_e = \sum_k \sum_i^{n_k} [T_{ki} (a_{ki}^2 + |a_{ki}| b_{ki} T_{ki} + \frac{b_{ki}^2 T_{ki}^2}{3})] \rightarrow \min, \quad (21)$$

Но  $b_{ki} T_{ki} = \Delta a_{ki}$ , т.е. это приращение силы на интервале  $T_{ki}$ . Тогда, вместо (7), имеем

$$\Omega_e = \sum_k \sum_i^{n_k} [T_{ki} (a_{ki}^2 + |a_{ki}| \Delta a_{ki} + \frac{\Delta a_{ki}^2}{3})] \rightarrow \min, \quad (22)$$

С целью более ясного определения взаимоотношений между длительностью команды и ее энергией примем, что  $b_{ki} = 0$ , т.е. что управляющая сила представляет собой ступенчатую функцию времени. Тогда критерий оптимальности приобретает вид

$$\Omega_e^{step} = \sum_k \sum_i^{n_k} T_{ki} a_{ki}^2 \rightarrow \min, \quad (23)$$

Отсюда видно, что для одного и того же значения критерия (23) возможен обмен длительности команды на ее энергию. Это может означать, что какой-то энергетический критерий оптимальности в системе управления артикуляцией имеет отношение к явлению подобного обмена, обнаруженному в экспериментах с измерением ЭМГ потенциалов лицевых мышц, описанному в следующем разделе.

Из (21) также следует, что возможен обмен между уровнем скачка команды и скоростью ее изменения.

Имеются свидетельства того, что временные параметры в речевом сигнале определяются множеством факторов. Похоже, что сначала планируется длительность высказывания, по крайней мере, синтагмы, т.е. участка речи от паузы до паузы. Во всяком случае, вариации длительности фразы оказываются в несколько раз меньше вариации длительности входящих в нее сегментов речи. Длительность фразы определяется темпом речи, стилем, важностью сообщения, а также индивидуальными характеристиками диктора – типом нервной системы, массами артикуляторов и мышечным аппаратом. Если число сегментов во фразе не превышает 12, то средняя длительность сегментов обратно пропорциональна их числу [10]. При большем числе сегментов их длительность нелинейно зависит от их числа. Известно также, что обычно начальные сегменты короче конечных [11, 12]. Ударные гласные обладают большей длительностью, чем безударные, а в некоторых языках имеется три типа длительности гласных, играющих смысловозначительную роль.

Поэтому разумно предположить, что при планировании речевого высказывания сначала задается длительность сегментов, а затем вычисляются параметры команд. Вместе с тем, скорость речи зависит и от общего мышечного тонуса организма. У больного или усталого человека можно наблюдать замедление темпа речи, включающее не только более медленные движения, но и более длительные сегменты речи. Это может быть связано с необходимостью уменьшения амплитуды ЭМГ потенциалов. Высказывается предположение, что удлинение сегментов во фразе от ее начала к концу может быть связано с тем, что для каждой фразы не только заранее планируется длительность, но и выделяется определенный запас энергии, и по мере расходования этого запаса происходит обмен затрат на сокращение мышц на длительность сегментов [10].

Наблюдения за динамикой артикуляции дают основание полагать, что оптимизация выполняется на достаточно большом интервале времени  $T_{opt}$ , таком, что

$$\sum_i^{n_k} T_{ki} \approx T_{opt}. \quad (24)$$

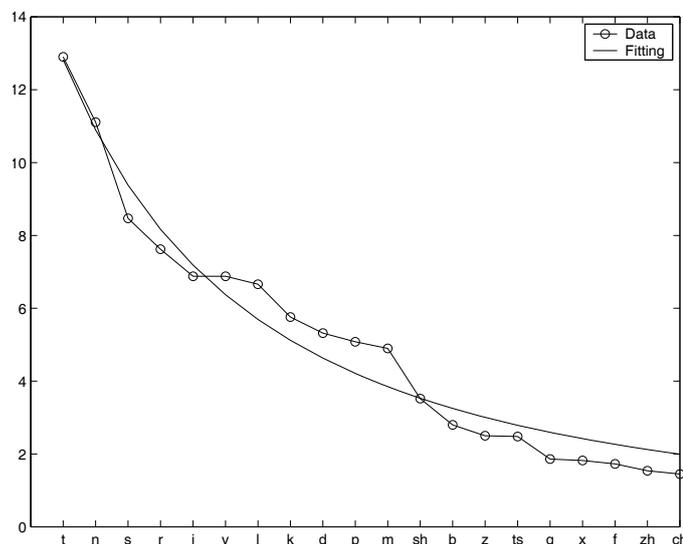
При этом команды управления на разные артикуляторы не обязательно синхронны, т.е. моменты их начала и конца и даже само число команд могут быть различны для разных артикуляторов, и эти параметры изменяются при смене темпе или стиля речи [2] или возмущения артикуля-

ции. Так, явление сдвига момента раскрытия голосовой щели при возмущении движения нижней губы во время артикуляции взрывных наблюдали [13].

Предварительные эксперименты с моделированием управления одним артикулятором в задаче аппроксимации реально измеренной траектории движения артикулятора на некотором интервале времени  $T_{opt} > T_i$  показали, что энергетический критерий для ступенчатых команд, т.е. при  $b=0$ , позволяет найти не только амплитуды команд  $a_i$ , но и моменты их начала и длительность  $T_i$ . Однако следует отметить, что произведение переменных, по которым ищется минимум критерия (20), делает задачу минимизации плохо обусловленной в том смысле, что одно и то же значение критерия может быть достигнуто при разных значениях переменных. При минимизации управлений для полного набора артикуляторов ситуация становится еще более неопределенной в силу явления взаимной компенсации положений артикуляторов.

Таким образом, обнаруживается еще один источник избыточности в системе управления артикуляцией. Разрешение неопределенности при планировании управлений может быть выполнено либо произвольным образом, либо с учетом каких-то дополнительных ограничений. Например, длительности сегментов речи могут быть ограничены, и тогда амплитуда управлений может увеличиться. Или, наоборот, в силу особенностей распределения двигательных единиц в мышце, амплитуда или скорость сокращения могут быть ограничены. Тогда критерий минимума энергии приведет к увеличению длительности речевых сегментов. Одним из проявлений избыточности на уровне критериев оптимальности является тот факт, что при необходимости говорить быстрее некоторые люди действительно ускоряют движение артикуляторов, тогда как другие просто сокращают длительность речевых сегментов.

Существует мнение, что, помимо экономии энергии, в системе управления движениями человека минимизируется и сложность управления, хотя само понятие сложности не имеет четкого определения. Наблюдения за совместными изменениями углов в различных суставах привели [14] к формулированию понятия синергии управления. Предполагается, что для достижения какой-либо цели, например, положения кончика пальца в пространстве, одна и та же последовательность команд управляет сразу несколькими группами мышц. Синергетическое управление действительно наблюдается во многих физиологических экспериментах. Понятие синергии трактует сложность управления как некоторую меру, зависящую от числа пространственно независимых команд управления.



**Рис. 1.** Ранжированные по частоте встречаемости фонемы русского языка и аппроксимация законом Мандельброта.

К другому определению сложности управления приводит интерпретация закона Мандельброта для распределения частоты появления фонем в некотором языке [2]. Закон Мандельброта выводится в результате решения задачи максимизации информации на символ кода при условии ограничения "сложности" генерирования кодовой последовательности, в нашем случае - звуков речи. В частности, сложность генерации звуков речи может быть связана и с числом независимо управляемых артикуляторов. На Рис. 1 показана ранжированная последовательность фонем русского языка и ее аппроксимация законом Мандельброта.

По оси абсцисс представлены символы фонетической транскрипции, а по оси ординат – вероятность появления каждого символа в речи в процентах.

Первые четыре места по частоте встречаемости занимают звуки, образованные подъемом кончика языка, управление которым требует наименьших усилий. На первом и третьем местах находятся звуки /m/ и /c/, которые артикулируются без участия голосового источника, на втором месте – назальный /n/, для которого не требуется подъема небной занавески, а на четвертом месте – очень короткий звук /p/. Глухие звуки /n, m, κ/ встречаются чаще звуков с тем же местом артикуляции /b, d, z/, для образования которых используется голосовой источник. Это же относится и глухим и звонким фрикативным. В [2] было показано, что частота встречаемости фонем не связана с их по- мехоустойчивостью.

В [15] было обнаружено, что редко встречающиеся слова не ассимилируются, тогда как частые слова воспринимаются либо как полностью звонкие, либо как глухие. Это явление может быть интерпретировано, как указание на то, что артикуляторные усилия минимизируются путем ликвидации разницы между звонкими и глухими в частых словах.

Известно также, что чем короче слово, тем чаще оно встречается в речи, причем наблюдается тенденция к упрощению произношения слова по сравнению с его литературным стилем. Эти явления указывают на доминирующую роль затрат на генерирование речи.

В [16] критерий оптимальности вычислялся на интервале  $T$  существования звуко сочетания

$$\sum_{n=2}^N (\|F_n - F_{n-1}\|^2 + w\|X_n - X_{n-1}\|^2) \rightarrow \min, \quad (25)$$

где  $F_n$  – вектор силы, приложенной к артикуляторам в каждый момент времени  $n$ ,  $N=T/\Delta t$ ,  $\Delta t$  – шаг по времени. В этот критерий входит не только скорость изменения положения артикулятора, но и скорость изменения управляющей силы. Последнее, подразумевает предпочтение как можно более редкому изменению управлений.

Число действующих артикуляторов не только определяет энергетические затраты на их управление, но также связано и с числом операций по планированию и исполнению программы управления артикуляцией. Поэтому можно предположить, что в критерий сложности входят и затраты на программирование изменения команд во времени, т.е. приращения параметров  $a_k$  и  $b_k$ , и число таких изменений за единицу времени.

Из закона Мандельброта для частоты встречаемости фонем следует, что чем меньше артикуляторов участвует в образовании звука, тем чаще этот звук появляется в речи.

Численная оценки сложности в зависимости от приращения параметров команд управления определяется свойствами центральной и периферической нервной системы. Она, конечно же, неизвестна. Поэтому, чисто формально примем его в виде квадратичного функционала

$$(\Delta a_k^2 + \mu \Delta b_k^2) \rightarrow \min. \quad (26)$$

Для системы управления, состоящей из  $k$  артикуляторов, критерий сложности постулируется в виде

$$\Omega_c = \sum_k \sum_i^{n_k} (\Delta a_{ki}^2 + \mu \Delta b_{ki}^2), \quad (27)$$

$n_k$  – число команд для  $k$ -го артикулятора на заданном интервале времени. Этот критерий предполагает стремление системы управления изменять параметры команд не только как можно меньше по сравнению с предыдущим состоянием, но и как можно реже. Множитель  $\mu$  формально описывает вес сложности управления по скорости относительно сложности управления по амплитуде команды. Различные значения этого веса приводят к различным тактикам управления артикуляцией.

Объединяя энергетическую компоненту  $\Omega_e$  и компоненту сложности управления  $\Omega_c$ , получим критерий оптимальности для системы управления артикуляцией в виде

$$\Omega = \Omega_e + \nu \Omega_c = \sum_k \sum_i^{n_k} \left\{ [T_{ki} (a_{ki}^2 + |a_{ki}| b_{ki} T_{ki} + \frac{b_{ki}^2 T_{ki}^2}{3})] + \nu (\Delta a_{ki}^2 + \mu \Delta b_{ki}^2) \right\}. \quad (28)$$

Здесь множитель  $v$  определяет относительную роль факторов экономии энергии и сложности управления.

Критерий оптимальности, подобный (28), может использоваться при планировании речевого высказывания и контроле реализации этого плана во внутренней модели. При этом не возникает проблема плохой обусловленности из-за связанности длительности команд и их амплитуды, поскольку система управления осведомлена о планируемой временной диаграмме управлений и ожидаемых артикуляторных и акустических параметрах.

Как видно, критерий (28) допускает разнообразные соотношения между длительностью команд, их количеством на заданном интервале времени, амплитудой скачка и скоростью изменения команды, а также приращением этих параметров для соседних команд. Это дает возможность удовлетворить критерию оптимальности при различных требованиях на темп речи, сегментную длительность, а также силу мышц, соотношение быстрых и медленных моторных единиц в мышцах и т.д. Тем самым обеспечивается гибкость системы управления артикуляцией, ее способность к адаптации к изменению условий речеобразования. Все это отображается в разных тактиках управления артикуляцией, наблюдаемых у разных людей и разных стилях речи.

Таким образом, на всех уровнях управления речеобразованием присутствует избыточность, которая обеспечивает компенсацию возмущения процессов артикуляции или поражения отдельных элементов системы управления. Эта избыточность также создает условия для реализации различных тактик управления речеобразованием.

Конечной целью процесса речеобразования является создание такого речевого сигнала, параметры которого обеспечивают передачу информации слушателю. Поэтому всевозможные возмущения, как процесса артикуляции, так и процесса восприятия, вызывают реакцию, направленную на сохранение информативности речевого сигнала. Исследование процессов компенсации или долговременной адаптации к таким возмущениям выявляет конкретные механизмы использования избыточности в управлении речеобразованием.

#### *Искажение формы твердого неба*

Известно, что трубка в зубах или зубной протез, искажая привычный динамический стереотип артикуляции говорящего, все же позволяет генерировать речь с достаточной разборчивостью. Эти наблюдения послужили основанием для постановки экспериментов с использованием искусственного твердого неба, в которых изучалась кратковременная компенсация и долговременная адаптация к искажению формы речевого тракта. В [17, 18] измерялись формантные частоты гласных диктора без использования искусственного неба и в процессе его ношения. Во многих случаях наблюдалась постепенная адаптация, и формантные частоты приближались к исходным значениям. Вместе с тем отмечались случаи, когда компенсация искажений формы тракта не достигалась даже после длительного периода ношения этого протеза.

В [19] исследовалась изменения артикуляции с помощью электропалатографа, заметно искажающего форму твердого неба. Выяснилось, что при прознесении фрикативного /с/, на условия турбулентных шумов которого этот протез оказывает наибольшее влияние, происходит сдвиг места артикуляции назад и изменение ширины поперечного прогиба языка. Было также замечено, что положение языка, определяющее место артикуляции, характеризуется повышенной нестабильностью относительно нормальных условий артикуляции.

#### *Препятствия движению артикуляторов – bite-block, lips-block*

Аналогично экспериментам с искусственным небом, где возмущение формы речевого тракта было постоянным, были разработаны специальные экспериментальные методики с препятствием движению нижней челюсти или губ. Между коренными зубами вставлялся твердый предмет (байт-блок, *bite-block*), который не позволял сблизиться челюстям больше, чем на 1 см [20-23]. В этих экспериментах было установлено, что формантные частоты гласных звуков, произведенных с байт-блоком и без него, могут быть весьма близки вследствие компенсаторного изменения формы речевого тракта.

Более того, были описаны случаи, когда первые три резонансные частоты гласного, произносимого с байт-блоком, на первом же импульсе голосового источника были идентичны формантным частотам гласного без возмущения артикуляции. Следовательно, система управления артикуляцией выполнила компенсацию возмущения еще до того, как испытуемый услышал искажение собственного гласного. В этом случае источником информации для компенсации возмущения

могла служить только информация от механорецепторов речевого тракта, каким-то образом установивших невозможность движения нижней челюсти при произнесении заданного гласного, в результате чего программа управления была соответственно реорганизована.

Вместе с успешной компенсацией возмущения в виде байт-блока, наблюдались и случаи, когда испытуемый не смог восстановить требуемые акустические характеристики звука [24]. Это свидетельствует о том, что, несмотря на известную избыточность, система управления артикуляции конкретного человека не всегда может найти необходимые преобразования для компенсации возмущений.

Аналогичные эксперименты по оценке способности системы управления к компенсации возмущений артикуляции были выполнены путем блокирования смыкания губ (*lips-block*) с помощью трубочки при произнесении гласного /y/ [25]. При первом испытании ни один диктор не смог компенсировать возмущение, но 7 из 11 предприняли изменение формы тракта так, чтобы частично компенсировать изменение F1, F2. В последующих 19 испытаниях только один диктор полностью компенсировал возмущение путем продвижения сужения в сторону голосовой щели, 6 дикторов частично компенсировали возмущение, а остальные не смогли.

В развитие методики с постоянным во времени возмущением артикуляции, проводились эксперименты с динамическим вмешательством в движения артикуляторов. Исследовалась реакция артикуляторов на неожиданное механическое возмущение движений губ и нижней челюсти [18, 26 - 30], внезапную электрическую стимуляцию мышц, управляющих артикуляторами [31], или изменение формы твердого неба [32]. Результаты экспериментов такого типа показывают, что компенсация возмущений может происходить достаточно быстро, с задержкой 10 – 40 мс, что не оставляет времени для пробных артикуляторных движений, и сопоставимо с общей задержкой прохождения сигнала от периферии до центральной нервной системы и обратно. Это означает, что сигналы компенсации при таких видах возмущения вычисляются практически мгновенно.

При создании сопротивления закрытию губ при артикуляции губных взрывных увеличение мышечного напряжения задерживается всего лишь на несколько миллисекунд. В аналогичных экспериментах получены задержки порядка 15—20 мс [26]. В экспериментах по внезапному растяжению углов в [33] губ было установлено, что рефлекторная активность губных мышц появляется через 14—17 мс, а амплитуда рефлекса зависит от логарифма скорости растяжения (или отпущения), что, кстати, отличается от рефлекса на растяжение, наблюдаемого у животных, где амплитуда отклика зависит от амплитуды смещения.

В этих экспериментах также была обнаружена реорганизация управлений в ответ на динамическое возмущение артикуляции. Так, в [13] исследовалось влияние неожиданного возмущения движений губ при артикуляции первого согласного в звукосочетании /i'pip/ на сведение и разведение голосовых складок. Было найдено, что при задержке сближения губ разведение складок задерживается, а длительность сведения складок увеличивается. Это наблюдение может быть интерпретировано, как стремление системы управления артикуляцией сохранить акустические характеристики глухого взрыва.

#### *Подавление обратной связи - анестезия механорецепторов*

Анестезия поверхности языка или нервов, управляющих речедвигательным аппаратом, в течение некоторого времени широко применялась для выяснения роли механической обратной связи. Метод анестезии основан на том, что нервные волокна меньшего диаметра, передающие, например, сигналы  $\gamma$ -управления к веретенам, поддаются анестезии быстрее и угнетаются дольше, чем волокна большого диаметра. Анализ результатов этих исследований выявил некоторые методологические недостатки, свойственные этому методу. Во-первых, при анестезии неизвестно, какие виды обратной связи и в какой степени угнетены. Во-вторых, остается неизвестной и степень подавления сигналов из центральной нервной системы. Так, имеются свидетельства ослабления челюстно-подъязычной мышцы при анестезии подбородочного нерва [34] и влияния анестезии на центральные механизмы при наркозах в клинических условиях. Анестезия влияет на пассивные и активные движения по-разному. Установлено, что при анестезии  $\gamma$ -путей в режиме пассивного растяжения веретена ведут себя так же, как и до анестезии, но при произвольных сокращениях  $\alpha$  и  $\gamma$ -отклики уменьшаются [5].

В [35] было найдено, что при анестезии подбородочной ветви тройничного нерва движения нижней челюсти были нарушены, а речь воспринималась как нормальная. При этом подъем нижней челюсти уменьшился в среднем на 3 мм, но это было компенсировано увеличенной амплитудой

дой движения губ. Это является экспериментальным подтверждением использования возможности обмена смещений нижней челюсти и губ.

Результаты экспериментов по анестезии дали основание утверждать, что, хотя достигаемые эффекты выражены довольно слабо, а иногда и противоречивы, анестезия уменьшает скорость и увеличивает амплитуду движений. При общей малой повреждаемости речи согласные все же повреждаются сильнее, чем гласные.

Наибольшая повреждаемость переднеязычных звуков и особенно /s/ отмечается при двусторонней анестезии подбородочного [36] и тройничного [37] нервов. При этом в [36] все ошибки являются ошибками места артикуляции с наиболее частым переходом /s → sh/, хотя пересечение границ незначительно. В [37] параллельно с анестезией проводились измерения ЭМГ. Было обнаружено, что активность челюстно-подъязычной мышцы и передней головки двубрюшной мышцы резко падала, а активность некоторых других мышц, например подбородочно-язычной, повышалась. Тот факт, что в этих экспериментах больше всего повреждались фрикативные звуки, свидетельствует о наиболее высоких требованиях к точности системы управления артикуляцией этих звуков.

Обстоятельные исследования влияния анестезии на процессы артикуляции были проведены в [38], где в различных комбинациях применялись поверхностная анестезия языка и блокировка язычного нерва наряду с бинауральной маскировкой белым - шумом высокого уровня (100 дБ). Одновременно регистрировались спектрограммы и палатограммы. Наибольшее влияние проявилось при совместном воздействии всех факторов. Оно состояло в увеличении длительности звуков, большей площади контакта языка с нёбом, сужении центрального желобка языка и сдвиге места артикуляции вперед. На основании результатов этих экспериментов был сделан вывод о том, что для временной регуляции сегментов речи и площади контакта языка с нёбом важна не столько тактильная рецепция, сколько акустическая обратная связь. Однако малое влияние нарушений отдельных видов обратной связи скорее указывает на высокую надежность и гибкость комплекса различных видов обратных связей, обеспечиваемых разными типами рецепторов. Наблюдения за клиническими случаями нарушений речевой деятельности, в которых происходит обширное нарушение механизмов обратной связи на высших уровнях, демонстрируют невозможность нормальной реализации процесса речеобразования без обратных связей.

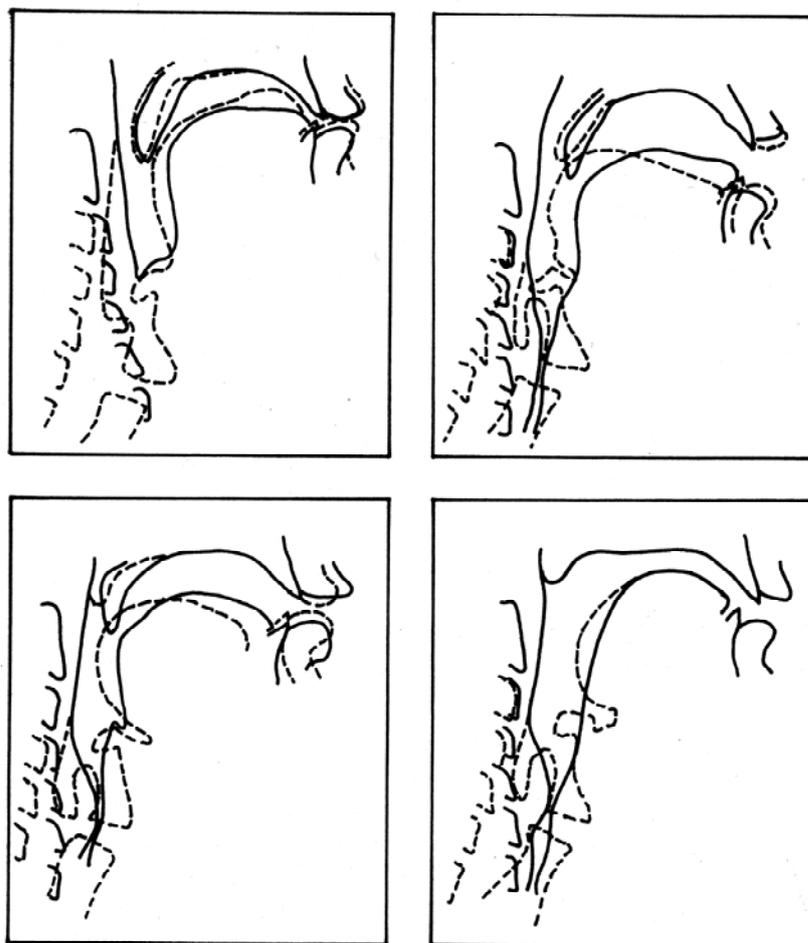
По мнению [9], малая эффективность анестезии язычного и тройничного нервов объясняется тем, что механорецепторы языка не включены в рефлекторную цепь в процессах речеобразования. Однако, роль тактильных рецепторов в контроле места и степени касания языком поверхности речевого тракта очевидна.

#### *Ларингэктомия*

Наблюдения за параметрами речи у людей, претерпевших операцию удаления гортани (ларингэктомию), дают поразительные примеры способности системы управления артикуляцией к адаптации к значительным изменениям в форме речевого тракта и отсутствию нормального источника голосового возбуждения. Обычно оцениваются такие параметры, как плавность речи, динамический диапазон, разборчивость [39, 40, 41]. В [42] оценивалась разборчивость речи до ларингэктомии, через четыре дня и год после операции. Было обнаружено значительное улучшение разборчивости речи у большинства пациентов. Аналогичные оценки приводятся и в [43], где сообщается об улучшении разборчивости речи до 61% у 47 пациентов через 6 месяцев после операции.

В [44] исследовалась речь больных до удаления гортани, через две недели после операции, и через год после операции. Были выполнены рентгенографические и акустические измерения. Пациенты произносили гласные /a, u, y/, спектры которых усреднялись на интервале около 200 мс. Измерения формантных частот и амплитуд показали заметные отклонения от среднестатистических параметров, характерных для этих гласных.

На Рис. 2 показаны контуры рентгенограммы в средне-сагиттальной плоскости четырех пациентов до операции и через два дня после операции. Как видно, форма речевого тракта при произнесении одного и того же гласного претерпела значительные изменения.



**Рис. 2.** Форма речевого тракта до операции (---) и после операции (—). Вверху: слева - нейтральная артикуляторная позиция, справа - артикуляция гласного /a/. Внизу: слева - артикуляция гласного /y/, справа - артикуляция гласного /u/.

В этих исследованиях была выявлена сравнительно быстрая компенсация акустических характеристик гласных звуков, когда уже через две недели их формантные частоты приблизились к области распределения для нормальных дикторов. Еще более полное восстановление фонетических параметров произошло после длительной адаптации при повторных измерениях через год после операции. Это произошло путем изменения формы речевого тракта при артикуляции гласных после удаления гортани.

Наиболее поразительное явление было обнаружено в функционировании суррогатного источника голосового возбуждения, образованного сфинктером, который был создан хирургическим путем из мышц – сжимателей глотки. Через некоторое время после операции в речи пациентов наблюдалось правильное управление звонкостью/глухостью согласных в соответствии с их фонетическим качеством, и, более того, формировалась правильная интонация во фразах различного типа. Например, в вопросительных фразах частота основного тона должна повышаться к концу фразы, что и наблюдалось в речи пациентов. Встречаются случаи, когда после операции пациенты даже могут петь, демонстрируя полное владение управлением частотой основного тона.

Это значит, что в управлении суррогатным голосовым источником произошли необычные изменения, поскольку после удаления гортани с мышцами, участвовавшими в голосообразовании, эти же функции были переданы мышцам – сжимателям глотки, которые никогда их не выполняли. Таким образом, даже такое катастрофическое вмешательство в процесс речеобразования, как удаление гортани, удается компенсировать.

Имеется, по крайней мере, один пример успешной компенсации удаленного языка за счет использования соответствующим образом изготовленного протеза. Разборчивость речи этого пациента оказалась вполне приемлемой для общения.

Следует отметить, однако, что такая компенсация наблюдается далеко не всегда. По нашему опыту и сообщениям других исследователей, близкая к полной компенсация встречается примерно у трети пациентов, тогда как еще треть оказывается не в состоянии восстановить голосовой источник. Это свидетельствует об объективной сложности перестройки системы управления артикуляцией.

#### *Эффект Ломбарда*

Одним из примеров реакции на изменение условий восприятия является эффект Ломбарда, проявляющийся при разговоре в присутствии шумов в окружающей среде. Оценивая присутствие шумов своей слуховой системой, говорящий непроизвольно повышает громкость речи, чтобы создать необходимое превышение уровня речевого сигнала над уровнем шумов. Это достигается повышением легочного давления и напряжением мышц гортани, в результате чего скорость раскрытия и схлопывания голосовых складок возрастает и увеличивается уровень сигнала голосового возбуждения. Побочным эффектом является повышение частоты основного тона и расширение спектра речевого сигнала. Увеличение натяжения мышц гортани может сопровождаться подъемом гортани, что приводит к уменьшению длины речевого тракта и, соответственно, к сдвигу его резонансных частот вверх.

#### *Задержка восприятия собственной речи (эффект Ли)*

Некоторые количественные оценки характеристик системы управления артикуляцией с нарушением акустической обратной связи были получены в экспериментах по прослушиванию собственной речи с задержкой во времени. Впервые такие эксперименты были описаны в [Lee, [45]. Впоследствии было обнаружено, что при увеличении времени задержки до 80 мс человек начинает «заикаться», повторяя отдельные фонемы. Максимальный эффект возникает при задержке 200 мс, причем этот интервал не зависит от длины слова, т. е. контекста. Такой интервал соответствует слогу, т. е. двум или трем артикуляторным состояниям и переходным процессам между ними. При задержке до 1 сек человек повторяет целые слова, а при дальнейшем увеличении задержки заикание обычно прекращается. Любопытно, что такое возмущение не поддается компенсации.

Это явление вызывает аналогии со свойствами технических систем автоматического регулирования, в которых при увеличении задержки в системе обратной связи сначала возникает неустойчивость в виде автоколебаний, а при дальнейшем увеличении задержки система вновь возвращается в область устойчивости, но уже с другими характеристиками по точности управления.

Искусственно вызванное заикание сопровождается явлениями, отличающимися от явлений, наблюдающихся во время естественного заикания. В частности, регистрация электрической активности мышц гортани в экспериментах с задержкой восприятия показывает возникновение колебаний мышечной активности, не похожих на колебания во время естественного заикания [46]. Вместе с тем схожесть внешнего эффекта позволяет проводить далеко идущие аналогии и относительно механизмов возникновения естественного заикания. Исследование свойств автономного контура управления артикулятором, проведенное в [2], показывает удивительное совпадение количественных характеристик автоколебаний и времени задержки с параметрами, полученными в экспериментах по задержке восприятия собственной речи.

#### *Возмущение частоты основного тона и формантных частот в обратной акустической связи*

Были также проведены эксперименты с возмущением акустических параметров речевого сигнала – основного тона [47 - 50] или формантных частот [51 - 56]. Влияние фильтрации воспринимаемого сигнала на назализацию гласных исследовалось в [57].

В этих экспериментах наблюдалась текущая адаптация формы речевого тракта с целью компенсации акустических возмущений и сохранения акустического образа, характерного для произносимого звука или звукосочетания

В [50] Основной тон возмущался на гласных с фиксированным питчем и на фразе с вопросительной интонацией, до повышения частоты ОТ к концу фразы (с понижением ОТ возмущения). Длительность возмущения – 200 мс. Компенсация для фраз была больше, чем для гласных, и компенсация для понижения питча была больше, чем для его повышения. Задержка отклика для фраз была меньше (122 мс), чем для гласных (184 мс). Предположительно, происходит предсказание влияния возмущения на вопросительный тип интонации.

В экспериментах, описанных в [54] при восприятии слогов СГС сдвигалась частота первой форманты F1. Наблюдалась компенсация вплоть до возмущений  $\pm 200$  Гц, порог для начала ком-

пенсации обычно >60 Гц. После прекращения возмущения, некоторое время оставалась гиперкомпенсация.

*Изменение темпа артикуляции как возмущение управлений*

Изменение темпа артикуляции вызывает реорганизацию команд управления. Одно время считалось, что изменение темпа артикуляции приводит к пропорциональному изменению длительности команд, тогда как относительные сдвиги фаз управления различными артикуляторными органами сохраняются инвариантными. Однако вскоре было обнаружено изменение фаз активизации мышц, управляющих небной занавеской, языком и губами при изменении темпа артикуляции и места.

В [58] исследовалось влияние скорости артикуляции и места слога в звукосочетании на сдвиг начала нейромоторных команд относительно начала смычки, оцениваемой по акустическому сигналу. Электрическая активность верхней и нижней кольцевых мышц губ (*OOS* – *orbicularis oris superior*) и (*OOI* – *orbicularis oris superior*), подбородочной мышцы (*ME* – *mentalis*) и мышцы, опускающей углы губ (*DA* (*depressor anguly*)), регистрировалась проволочными и поверхностными электродами.

Речевой материал состоял из отдельных звуков, бессмысленных звукосочетаний, слов и фраз. Звукосочетания оканчивались слогом *БА* или *ПА*, начинаясь соответственно с *Б* или *П*, а в середине находилась конструкция гласный — согласный — гласный или гласный — согласный — согласный — гласный с различными сочетаниями звуков (например, *БУБИБА*, *ПАПМАЛА*, *БАБДОБА* и т. Д.). Весь текст повторялся несколько раз в нормальном и примерно вдвое ускоренном темпах.

Результаты усреднения времени задержки ЭМГ потенциалов для поверхностных электродов в среднем по 100 реализациям звука *Б* в бессмысленных звукосочетаниях приведены в Табл. 1, 2, где показаны время начала команд относительно начала звонкой смычки (опережение) и длительность этих команд для начального, среднего и конечного положений звука. Из этих Таблиц видно, что на верхнюю и нижнюю ветви кольцевой мышцы губ *OOS* и *OOI* команды подаются почти одновременно и оканчиваются либо в момент смычки (*OOS*), либо незадолго до нее (*OOI*). Опережение ЭМГ потенциалов и их длительность перед первой смычкой в 2 – 3 раза больше, чем перед последующими смычками как при нормальном, так и при ускоренном темпе артикуляции.

**Таблица 1.** Длительность  $\Delta T$  и опережение  $T_m$  ЭМГ потенциалов относительно конца смычки для /б/, мс. Поверхностные электроды, нормальная скорость артикуляции.

Мышца	Первая смычка		Вторая смычка		Третья смычка	
	$T_{m1}$	$\Delta T_1$	$T_{m2}$	$\Delta T_2$	$T_{m3}$	$\Delta T_3$
<i>OOS</i>	-210	210	-75	95	-85	115
<i>OOI</i>	-200	175	-75	95	-80	105
<i>ME</i>	-170	125	-50	60	-55	60
<i>DA</i>	-125	90	-60	50	-40	50

**Таблица 2.** Длительность  $\Delta T$  и опережение  $T_m$  ЭМГ потенциалов относительно конца смычки для /б/, мс. Поверхностные электроды, быстрая артикуляция.

Мышца	Первая смычка		Вторая смычка		Третья смычка	
	$T_{m1}$	$\Delta T_1$	$T_{m2}$	$\Delta T_2$	$T_{m3}$	$\Delta T_3$
<i>OOS</i>	-175	150	-60	60	-65	100
<i>OOI</i>	-175	140	-55	60	-65	90
<i>ME</i>	-105	60	-50	35	-55	45
<i>DA</i>	-115	70	-60	35	-50	40

Ускорение темпа артикуляции для среднего и конечного положений звука почти не влияет на начало команд у мышц *ME* и *DA*, но зато укорачивает их длительность, хотя и не пропорционально темпу ускорения речи. Для начального положения звука в быстром темпе команды в 2—3 раза опережают команды для средних и конечных положений, но тем не менее все же сдвигаются по направлению к смычке по сравнению с командами в нормальном темпе. Таким образом, временные соотношения между началами команд различных мышц и их длительностями изменяются в зависимости от темпа артикуляции и положения звука в звукосочетании. Анализ сигналов, регистрируемых проволочными электродами, подтверждает этот вывод. В данном случае импульсы на

смыкание появляются значительно раньше у начальных звуков, но длительность их короче, чем у сигналов, регистрируемых поверхностными электродами. Время же появления и длительность импульса перед взрывом смычки почти не зависят от положения звука в слове.

Табл. 3 показывает смещение потенциалов ЭМГ мышц относительно *orbicularis oris superior*. Видно, что и при нормальном и при ускоренном темпе артикуляции *orbicularis oris superior* и *orbicularis oris inferior* активируются одновременно и разница между началами команд мала. В противоположность этому, при нормальном темпе артикуляции потенциал *depressor anguly* существенно запаздывает (на 55 мс) относительно потенциала *mentalis*, тогда как при быстром темпе артикуляции потенциал *depressor anguly* опережает его на 10 мс. Задержка потенциала *mentalis* относительно потенциала *orbicularis oris superior* на 30 мс больше, чем при нормальном темпе. Фаза потенциала *depressor anguly* значительно различается при нормальном и быстром темпе артикуляции.

**Таблица 3.** Задержка начала EMG потенциала лицевых мышц относительно потенциалов мышцы *OOI* (*orbicularis oris superior*), мс. Проволочные электроды.

Muscle	First Closure		Second closure		Third closure	
	Normal	Fast	Normal	Fast	Normal	Fast
<i>OOI</i>	10	0	0	5	5	0
<i>ME</i>	40	70	25	10	30	10
<i>DA</i>	95	60	15	0	45	15

Различная скорость сокращения и сила мышц требует изменения фаз команд при разном темпе артикуляции. Например, координация мышечных усилий необходима для того, чтобы достичь смычки губ в заданный момент времени. Если максимальное усилие для каждой мышцы ограничено, то переходные процессы при разных темпах артикуляции могут быть скоординированы только путем изменения фаз команд.

В [2] описываются эксперименты с отведением ЭМГ потенциалов от жевательной мышцы (*masseter*). Электрическая активность этой мышцы возникала в зависимости от того, произносится ли осмысленное или бессмысленное звуко сочетание в нормальном или быстром темпе. При нормальном темпе артикуляции никакой активности не наблюдалось независимо от осмысленности звуко сочетания. Не было активности и при артикуляции бессмысленного звуко сочетания при ускоренном темпе, тогда как при артикуляции осмысленных звуко сочетаний в быстром темпе такая активность появлялась.

#### *Взаимодействие визуального и акустического восприятия – эффект МакГурка*

Аудио-визуальные эффекты также могут свидетельствовать о связи акустического и артикуляторного анализа. Например, была обнаружена электрическая активность слуховой зоны коры головного мозга слушателя, наблюдающего за артикуляторными движениями диктора, тогда как неречевая мимика диктора такой активности не вызывала [59, 60]. Дальнейшие исследования активности слуховой и моторной зоны коры могут принести ценную информацию о свойствах восприятия речи.

Из ежедневной практики хорошо известно, что визуальная информация облегчает восприятие речи других людей, особенно в неблагоприятных акустических условиях или для иностранного языка [61]. Наблюдение за лицом диктора влияет на восприятие речи, и в случае противоречия между видимой артикуляцией и услышанным звуко сочетанием возникают разнообразные эффекты восприятия [62].

Информативность наблюдаемых проявлений артикуляции позволяет общаться глухо-немым. Имеются сведения о том, что наблюдение за речевой мимикой говорящего в условиях высокого уровня шумов в несколько раз повышает разборчивость воспринятой речи по сравнению с тем, когда лицо говорящего не видно. Даже в условиях нормальной слышимости понимание речи на иностранном языке существенно облегчается при визуальном контакте.

#### **Заключение**

Совокупность экспериментальных данных о естественных нарушениях или искусственных возмущениях процессов речеобразования приводит к важным выводам о механизмах управления артикуляцией и позволяет сформировать конструктивные подходы к количественной оценке па-

раметров системы управления. Наблюдаемые эффекты компенсации служат основой для поиска критериев оптимальности, которыми руководствуется система управления.

Избыточность в системе управления артикуляцией активно используется для компенсации как статических, так и динамических помех. Сигналы обратной связи от механорецепторов и слуховой системы участвуют в коррекции управлений, если выходные параметры на каком-либо уровне – артикуляторном или акустическом, не соответствуют целевым значениям. Задержка коррекции возмущений артикуляции меньше задержки реакции на возмущение акустических параметров, и это говорит о дополнительной обработке речевого сигнала в слуховой системе. Относительно малое значение времени задержки исключает пробные движения и указывает на существование специального механизма обработки сигналов обратной связи, который является основой теории внутренней модели [63, 64].

Установлено, что, несмотря на наличие избыточности, успешная компенсация помех речеобразованию достигается далеко не всегда, что свидетельствует об объективной сложности формирования необходимой реакции. Заметное разнообразие акустических характеристик элементов речевого сигнала, наблюдаемое в естественных условиях, как правило, не мешает правильному восприятию речевого высказывания. Это является следствием специальной организации речевого общения, позволяющей корректировать ошибки восприятия с использованием лексической, синтаксической, семантической и прагматической избыточности.

### Литература

1. Nelson W. L., Westbury J. L. (1982). Jaw movements during increasingly rapid articulation of single syllables: Implication on speech motor control. — *JASA*, v. 71, Suppl. 1, p. S33.
2. В.Н.Сорокин. *Теория речеобразования*, 1985, Радио и связь, М., 313 с.
3. Cooper D. S., Folkins J. W. (1982). The temporal sampling problem in electromyographic studies of speech musculature. — *JASA*, v. 71, Suppl. 1, p. S33.
4. Леонов А.С., Сорокин В.Н. (2003). Энергетические критерии оптимальности в речевых обратных задачах. Доклады Академии наук, т. 392, № 5.
5. Granit, R., *The Basis of Motor Control*, 1970, London and N.Y.: Academic Press.
6. Kelso, S.J., Stelmach, G.E. (1976). Central and peripheral mechanisms in motor control, in *Motor Control. Issues and Trends*, Stelmach, G.E., (Ed.), Academic Press, N.Y., pp. 3-40.
7. Leonov A.S., Sorokin V.N. (2004). Control in the Internal Model: Score Reorganization and Compensation, *Pattern Recognition and Image Analysis*, v. 14, N 3, 407-420.
8. Kuehn, D.P., Moll, K.L. (1976). A cineradiographic study of VC and CV articulatory velocities. *J. Phonetics* 4, 303-320.
9. Edwards, J., Beckman, M.E., Fletcher, J. (1991). The articulatory phonetics of final lengthening. *J. Acoust. Soc. Amer.* 89, 369-382.
10. Moon, S.-J., Lindblom, B. (1994). Interaction between duration, context, and speaking style. *J. Acoust. Soc. Amer.* 96, 40-55.
11. Сорокин В.Н. (1992). *Синтез речи*. Наука, М., , 392 с.
12. Klatt D.H. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence, *JASA*, N5, pp. 1208-1221.
13. Fowler C.A. (1981). A relation between coarticulation and compensatory shortening, *Phonetica*, v. 7, N10, pp. 35-50.
14. Munhall K.G., Lofqvist A. and Kelso J.A.S. (1994). Lip-larynx coordination in speech: Effects of mechanical perturbations to the lower lip. *J. Acoust. Soc. Am.*, vol. 95, N 6, pp. 3605-3616.
15. Бернштейн Н.А. (1947). *О построении движений*. Медгиз.
16. Ernestus M., Lahey M, Verhees F, Baayen R (2006). Lexical frequency and voice assimilation. *JASA*, v. 120, N2, 1040-1051.
17. Kaburagi T., Honda M. (1996). A model of articulator trajectory formation based on the motor tasks shapes. *J. Acoust. Soc. Amer.* 99, 3154-3170.
18. Hamlet S.L. and Stone M. (1976). Compensatory vowel characteristics resulting from the presence of different types of experimental dental prostheses. *J. of Phonetics*, 7, 196-218.
19. Hamlet, S.L., Stone, M. (1978). Compensatory alveolar consonant production induced by wearing a dental prosthesis. *J. Phonetics*, v. 6, pp. 227-248.
20. McFarland D.H., Baum S.R., Chabot C. (1996). Speech compensation to structural modification of the oral cavity. *JASA*, 100(2), 1093-1104.

20. Lindblom B.E.F., Lubker J. and Gay T. (1979). Formant frequencies of some fixed mandible vowels and a model of speech motor programming by predictive simulations. *J. Phonetics*, vol. 7, pp. 147-161.
21. Fowler C.A., Turvey M.T. (1980). Immediate compensation in bite-block speech. *Phonetica*, vol. 37, pp. 306-326.
22. Gay T.J., Lindblom B., Lubker J. (1981). Production of bite-block vowels : Acoustic equivalence by selective compensation. *J. Acous. Soc. Am.*, vol. 69, pp. 802-810.
23. Fledge I.E., Fletcher S.G., and Homiedan A. (1988). Compensating for a bite block in /s/ and /t/ production: Palatographic, acoustic, and perceptual data. *J. Acoust. Soc. Am.*, v. 83, N 1, pp. 212-228.
24. MacFarland D.H. and Baum S.R. (1995). Incomplete compensation to articulatory perturbation. *J. Acoust. Soc. Am.*, vol. 97, N 3, pp. 1865-1873.
25. Savariaux C., Perrier P. and Orliquet J.P. (1995). Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube: A study of the control space in the speech production. *J. Acoust. Soc. Am.*, vol. 98, No. 5, pp. 2428-2442.
26. Folkins J.W. and Abbs J.H. (1975). Lip and jaw motor control during speech: Responses to resistive loading of the jaw. *J. Hearing and Speech Res.*, vol. 18, pp. 207-220.
27. Abbs, J.H., Gracco, V.L. (1984). Control of complex motor gestures: Orofacial muscle responses to load perturbations of lip during speech, *J. Neurophysiology*, vol. 51, pp. 705-723.
28. Kelso J.A.S., Tuller B., Vatiokis-Bateson E., Fowler C. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *J.*
29. Shaiman S. (1989). Kinematic and electromyographic responses to perturbation of the jaw. *J. Acoust. Soc. Am.*, vol. 86, pp. 78-87.
30. Kollia H.B., Gracco V.L., Harris K.S. (1992). Functional organization of velar movements following jaw perturbation. *J. Acoust. Soc. Am.*, vol. 91, p. 2474
31. Folkins J.W. and Zimmerman G.N. (1982). Lip and jaw interactions during speech : responses to perturbation of lower-lip movement prior to bilabial closure. *J. Acoust. Soc. Am.*, vol. 71, N 4, pp. 1225 – 1233.
32. Honda, M., Kaburagi, T. (2000). Speech compensation to dynamical structural perturbation of the palate shape. In: *Proc. 5<sup>th</sup> Seminar on Speech Production*, Kloster Seeon, Bavaria, pp. 21-24.
33. Lurson C. R., Folkins J. W., McClean M. D., Muller E. M. (1978). Sensitivity of the human perioral reflex to parameters of mechanical stretch. — *Brain Res.*, N 146, p. 159—164.
34. Abbs J. H., Eilenberg G. R. (1976). Peripheral mechanisms of speech motor control. Contemporary issues in experimental phonetics. — New York: Academic Press, , p. 139-168.
35. Abbs J. H. (1973). The influence of the gamma motor system on jaw movements during speech: a theoretical framework and some preliminary observations. — *JSHR*, , № 16, p. 175-200.
36. Borden G. J., Harris K. S., Oliver W. (1973). Oral feedback. 1. Variability of the effect of nerve block nerve block anesthesia upon speech. — *J. of Phonetics*, , v. 1 № 4, p. 289-295.
37. Borden G. J., Harris K. S., Catena L. (1973). Oral feedback. 2. An electromyographic study of speech under nerve-block anesthesia.—*J. of Phonetics*, v. 1, № 4, p. 297-308.
38. Hardcastle W. J. (1975). Some aspects of speech production under controlled conditions of oral anesthesia and auditory masking and auditory masking. — *J. of Phonetics*, , № 3, p. 197—214.
39. M.Staffieri (1987). Prothèse et chirurgie dans □ehabilitation de la voix après laryngectomie totale. *Rev. Laryngol.*, vol. 95, N 2, pp. 145-147.
40. Zanoft D.J., Wold D., Montague J.C., Krueger K. and Drummond S. (1990), “Tracheoesophageal speech with and without tracheostoma valve”, *Laryngoscope*, Vol. 100, N 5, pp. 408-502.
41. Hilgers F.J.M. and Balm A.M.M. (1993). Long term results of vocal rehabilitation after total laryngectomy with the low-resistance, indwelling Provox<sup>TM</sup> voice prosthesis system. *Clin. Otolaryngol.*, vol. 18, pp. 721-749.
42. Blom E.D., Singer M.I. and Hamaker R.C. (1986). A prospective study of tracheoesophageal speech. *Archives Otolaryng. Head Neck Surg.*, Vol. 112, pp. 440-447.
43. Gates G.A., Ryan W., Cooper J.C. et al. (1982). Current status of laryngectomee rehabilitation: 1. Results of therapy. *Am. J. Otolaryngol.* Vol. 3, pp. 1-7.
44. Sorokin V., Olshansky V., Kozhanov L. (1998). Internal model in articulatory control: Evidence from speaking without larynx. *Speech Communication*, v. 25, N 3, pp. 249-268
45. Lee B.S. (1950). Effects of delayed speech feedback. *JASA*, N22, 824.
46. Freeman F.J., Ushijima T. (1976). Laryngeal muscles in the stuttering. *SRSR Haskins Labs*, N45-46, 211-236.

47. Elman J.L. (1981). Effects of frequency-shifted feedback on the pitch of vocal productions. *JASA*, 70, 45-50.
48. Burnett T.A., Freedland M.B., Larson C.R., Hain T.C. (1998). Voice  $F_0$  responses to manipulations in pitch feedback, *J. Acoust. Soc. Am.*, vol. 103, pp. 3153-3161.
49. Jones, J.L., Munhall, K.G. (2000). Perceptual calibration of  $F_0$  production: Evidence from feedback manipulation. *J. Acoust. Soc. Am.*, v. 108, pp. 1246-1251.
50. Chen S.H., H.Lin, Y.Xu, Larson C.R. (2007). Voice  $F_0$  responses to pitch-shifted voice feedback during English speech. *JASA*, 121(2), 1157-1163.
51. Max, L., Wallace, M.E., Vincent, I. (2003). Sensorimotor adaptation to auditory perturbations during speech: Acoustic and kinematic experiments. Proc. 15<sup>th</sup> ICPHS, Barcelona, v. 1, pp. 1053-1056.
52. Villacorta, V., Perkell, J.S., and Guenter, F.H. (2004). Sensorimotor adaptation to acoustic perturbations in vowel formants. *J. Acoust. Soc. Am.*, v. 115, p. 2618.
53. Purcell, D.W., Munhall, K.G. (2005). Compensation following real-time manipulation of formants in isolated vowels. *J. Acoust. Soc. Am.*, v. 119(4), pp. 2288-2297.
54. Purcell D.W., Munhall K.G. (2006). Adaptive control of vowel formant frequency: Evidence from real time formant manipulation, *JASA*, 120(2), 966-977.
55. Houde J.F., Jordan M.I. (1998). Sensorimotor adaptation in speech production. *Science*, 209, 1213-1216.
56. Houde J.F., Jordan M.I. (2002). Sensorimotor adaptation in speech I: compensation and adaptation. *J. Speech and Language and Hearing Research*, 45, 295-310.
57. Garber S.R., Moller K.T. (1979). The effects of feedback filtering on nasalization in normal and hypernasal speakers. *J. Speech and Hearing Research*, 22, 321-333.
58. Сорокин В.Н. (1980). Влияние скорости артикуляции на нейромоторные процессы лицевых мышц. *Физиология Сорокин человека*, N 1, с. 40-45.
59. Sams, M., Aulanko, R., Hamalainen, H., Lounasmaa, O., Lu S., and Simola, J. (1991). Visual information from lip movements modifies activity in the human auditory cortex. *Neuroscience Letters*, v. 127, pp. 141-145.
60. Calvert, G., Brammer, M., Bullmore, E., Campbell, R., Williams, S., McGuire, P., Woodruff, P., Iversen, S., and David, A. (1997). Activation of auditory cortex during silent lip-reading. *Science*, v. 276, pp. 593-596.
61. Sumbly W.Y., Pollack I. (1954). Visual contribution to speech intelligibility in noise, *J. Acoust. Soc. Am.* vol. 26, pp. 212-215.
62. McGurk, H., MacDonald, J. (1976). Hearing lips and seeing voices, *Nature*, v. 264, pp. 746-748.
63. Sorokin V.N. (2006). Speech inversion: Problems and Solutions. In: *Dynamics of Speech Production and Perception*, P.Divenyu, S.Greenberg, G.Meyer (eds), IOS Press, 263-282.
64. Сорокин В.Н. (2007). Моторная теория восприятия речи и теория внутренней модели. *Информационные процессы*, т. 7, №1, 1-12. [www.jip.ru](http://www.jip.ru).