

ВРЕМЯ ПОЯВЛЕНИЯ ЗВУЧАЩЕЙ РЕЧИ ПО ДАННЫМ АНТРОПОЛОГИИ

С.А. Бурлак

Институт востоковедения РАН, филологический факультет МГУ, Москва

В статье суммируются данные, проливающие свет на время возникновения звучащей речи. Рассматриваются анатомические черты, с которыми различные исследователи связывали переход гоминид к членораздельной звучащей речи (опущенная гортань, величина базикраниального угла, толщина канала подъязычного нерва, форма и расположение подъязычной кости, подбородочный выступ, ширина позвоночного канала), и степень их релевантности для обоснования таких гипотез: наиболее релевантными оказываются низкое положение гортани (в сочетании со сравнительно небольшой длиной ротовой полости, что обеспечивает языку большую подвижность в разных плоскостях и, тем самым, большие возможности для различения звуков при помощи артикуляции) и толщина позвоночного канала (свидетельствующая о расширении возможностей управления дыханием). Сопоставляются слуховые возможности как современных людей и человекообразных обезьян, так и ископаемых гоминид (по реконструкции на основе формы слуховых косточек), а также их возможности управлять дыханием (при помощи спинного мозга) и обрабатывать информацию (при помощи головного мозга). В обеспечении функционирования языка принимают участие не только зона Брока и зона Вернике, большое значение имеют лобные доли (в особенности, префронтальные формации коры). В то же время само по себе увеличение объема мозга в тех областях, где у современного человека находятся «речевые зоны», не является однозначным свидетельством в пользу наличия у соответствующего вида гоминид развитого звукового языка. Распознавание речи имеет комплексный характер, в нём участвуют мозговые структуры, распознающие не только отдельные смыслоразличительные признаки фонем, но и переходы от одной фонемы к другой, общий звуковой облик слова, языковой и внеязыковой контекст. Рассматриваются также когнитивные приспособления, необходимые для формирования звучащей речи: способность к самоподдерживающемуся звукоподражанию, желание слышать понимаемую речь и интерпретировать звуки, произносимые окружающими, как знаки. Формирование всего этого комплекса признаков происходит лишь в гоминидной линии и не может быть вызвано одиночной макромутацией. Всесторонний анализ данных позволяет сделать вывод, что звучащей речью обладал уже *Homo heidelbergensis*: у него отсутствуют горловые мешки (сильно уменьшающие влияние артикуляции на акустические характеристики звука), слух постепенно настраивается на лучшее распознавание тех частот, которые могут быть усилены или ослаблены при помощи артикуляции; у него же увеличивается толщина позвоночного канала. Но при этом нет оснований считать эту речь настоящим человеческим языком, поскольку многие аспекты языка связаны с лобными долями, которые у *Homo heidelbergensis* были меньше, чем у *Homo sapiens*.

Ключевые слова: антропология, происхождение языка; звучащая речь, *Homo heidelbergensis*

В числе ключевых свойств человеческого языка, отличающих его от коммуникативных систем других видов, непременно называют членораздельную звучащую речь. В середине XX века с лёгкой руки американского психолога Олвина Либермана эта идея обрела форму изящного афоризма – *Speech is special* (букв. «речь [видо]специфична»).

Человеческая речь – не просто издавание звуков, имеющих определенный смысл. Звуковая

сторона речи имеет сложную, иерархически устроенную организацию [см., например: Кодзасов, Кривнова, 2001]. Из отдельных звуков – фонем – складываются слоги, из слогов – фонетические слова (они могут не совпадать с грамматическими словами, например, в лесу – одно фонетическое слово, но два грамматических), из фонетических слов – фонетическим синтагмы (сочетания из нескольких фонетических слов без пауз между ними), а из них – фонетические предложения, или

периоды. Непросто устроены и сами звуки. При помощи движений языка, губ, нижней челюсти, нёбной занавески, надгортанника, человек, изменяя резонансные свойства речевого тракта, ослабляет одни частоты получающегося в итоге звука и усиливает другие. Каждый гласный характеризуется своим собственным «узором» усиленных частот (формант). Согласные также имеют свои частотные максимумы и минимумы, но распознаются в значительной степени по тому влиянию, которое они оказывают на форманты соседствующих с ними гласных. Например, после заднеязычного согласного (*g* или *k*) у последующего гласного сближаются начальные точки контуров второй и третьей формант.

Вообще говоря, прямое отождествление человеческого языка и членораздельной звучащей речи не вполне верно, поскольку жестовые языки глухонемых ни в коей мере не являются «менее человеческими», чем языки устные. Вопреки распространённому заблуждению, жесты этих языков передают не отдельные буквы (хотя пальцевая азбука – дактилология – тоже имеется, прежде всего для передачи имен собственных), а целые слова (или морфемы – значимые части слов). Каждый жест-слово состоит из незначимых элементов – хирем, а из слов, как и в устном языке, составляются словосочетания и предложения [см.: Stokoe, 1960; Зайцева, 1993]. В них имеется грамматика (например, средства выражения множественного числа, разнообразных видовых различий и мн. др.), различные стили речи, на них можно вести диалоги и составлять монологические рассказы на любые темы (например, можно пересказать приятелю сюрреалистический мультфильм).

И всё же, по-видимому, человеческий язык формировался в первую очередь именно как язык устный – об этом говорят имеющиеся у *Homo sapiens* многочисленные приспособления для производства и восприятия членораздельного звука. Главное из них – гортань, расположенная ниже, чем у современных человекообразных обезьян. Низкое положение гортани открывает возможность для чёткого произнесения звуков человеческого языка, но при этом создаёт риск подавиться. Как пишет С. Пинкер, «до недавнего изобретения приёма Геймлиха попадание еды в дыхательные пути было шестой лидирующей причиной смерти от несчастного случая в Соединенных Штатах, уносившей шесть тысяч жизней в год» [Пинкер, 2004]. Следует уточнить, что для членораздельной речи важно не положение гортани само по себе, а соотношение длины ротовой полости и длины глотки: у современного человека эти дли-

ны приблизительно одинаковы, так что части языка, расположенные во рту и в глотке, примерно уравниваются, и это даёт возможность чётко различать в произношении все звуки, включая «крайние» гласные – [i], [u] и [a] [Lieberman, 2002]. У других видов гоминид такой возможности, видимо, не было, поскольку они, как видно по сохранившимся черепам, обладали достаточно сильно выступавшими вперёд челюстями и, соответственно, длинной ротовой полостью. Чтобы уравновесить такую длину рта соответствующей длиной глотки, гортань должна была бы располагаться где-то в груди, что крайне маловероятно [ср.: Lieberman, 2002]. Впрочем, невозможность произнесения «крайних» гласных сама по себе не может свидетельствовать ни об отсутствии языка (язык мог в принципе быть и жестовым), ни даже об отсутствии членораздельной звучащей речи – просто, если таковая была, она должна была в большей степени полагаться на различия согласных звуков [Deacon, 1997].

Отметим, что у человеческих младенцев гортань тоже, как и у шимпанзе, расположена высоко (это позволяет одновременно сосать и дышать). Примерно к трём годам гортань опускается – и это приблизительно совпадает со временем полного овладения звуковой стороной языка. Впрочем, справедливости ради следует сказать, что положение гортани не остается неизменным в течение жизни не только у человека: по данным группы японских ученых, некоторое опускание гортани наблюдается и у шимпанзе [Nishimura et al., 2003].

О том, для чего нужно низкое положение гортани, существует по меньшей мере две гипотезы. Согласно одной точке зрения, оно необходимо только для членораздельной звучащей речи [см., например, Lieberman, 2002; Вишняцкий, 2004], поскольку дает возможность языку двигаться внутри речевого тракта как в горизонтальной, так и в вертикальной плоскости. Это позволяет создавать различные конфигурации ротовой полости и глотки независимо и тем самым сильно расширяет набор возможных фонем, различающихся по тому, на каких частотах звук усилен, а на каких, наоборот, приглушён. Согласно другой точке зрения, главная роль понижения гортани – это обеспечение возможности издавать более низкие звуки и тем самым создавать у слушателей впечатление, что говорящий имеет большие размеры, чем на самом деле [Hauser et al., 2002. P. 1574]. По-видимому, эта точка зрения принципиально неверна. Дело не только в том, что преувеличение собственного размера – слишком ничтожное приобретение для такой огромной «цены», как риск подавиться. Главное, как кажется, то, что приматы

(а значит, по-видимому, и ранние гоминиды) – групповые животные с достаточно высоким уровнем интеллекта. Они долгие годы живут вместе, часто встречаются и хорошо знают друг друга «в лицо» – как показывают наблюдения, роль межличностных контактов в обезьяньем сообществе весьма велика [см., например, Бутовская, Файнберг, 1993]. В такой ситуации пытаться создавать ложное впечатление о размере (который виден невооружённым глазом и всей группе давно известен) просто бесполезно (показательно, что М. Хаузер и его соавторы в подтверждение своей точки зрения ссылаются на лягушек и птиц, которые производят коммуникативные действия на таких расстояниях и в такой среде, что размер того, кто издаёт звук, слушающему не виден). Предположение, что понижение тембра голоса было необходимо в межгрупповых конфликтах (для того, чтобы издавляя устроить членов соседней группировки) тоже едва ли может претендовать на обоснованность: во-первых, такая задача должна была обусловить понижение гортани у взрослых особей мужского пола, но не у женщин и трёхлетних детей, для которых устрашение соседей неактуально, а во-вторых, слух человека настроен на преимущественное восприятие частот, слишком высоких для тех расстояний, на которых осуществляется межгрупповое общение (см. подробнее ниже). Таким образом, остается лишь одна возможность: низкое положение гортани как видовой признак – это одно из приспособлений для членораздельной звучащей речи.

Сложность заключается в том, что гортань не содержит костей, а мягкие ткани не сохраняются, так что все имеющиеся в науке сведения о положении гортани у того или иного вида гоминид – это реконструкции, выполненные на основании предположения, что положение гортани связано с величиной базикраниального угла (угла основания черепа): чем сильнее изогнуто основание черепа, тем ниже расположена гортань [Lieberman, Crelin, 1971; Lieberman et al., 1972; Laitman, Crelin, 1976; Crelin, 1987]. У австралопитеков и *Homo habilis* изгиб основания черепа примерно такой же, как у современных человекообразных обезьян, у архантропов он выражен значительно сильнее – поэтому некоторые исследователи предполагают, что «первыми «говорящими» гоминидами были, возможно, *Homo ergaster*» [Зубов, 2004]. Впрочем, исследования указывают на отсутствие (у современных людей) надёжной корреляции между величиной изгиба базикраниума и положением гортани [Lieberman, McCarthy, 1999].

Но низкое положение гортани – далеко не единственное анатомическое приспособление,

необходимое для языка. Большое значение для звучащей речи имеет точное управление мышцами языка. Движения всех этих мышц, кроме одной, обеспечиваются подъязычным нервом, канал которого расположен в толще затылочной кости между ярёмным отростком и мыщелком. Чем толще этот канал, тем, как считается, больше нейронов задействовано в управлении языком, поэтому на основании этого признака делаются предположения о наличии у того или иного вида гоминид звучащей речи: в частности, известно, что у представителей рода *Homo* этот канал как абсолютно, так и относительно больше, чем у человекообразных обезьян и австралопитековых [Kay et al., 1998]. Однако, по данным других исследователей, с учётом различий в размере языка толщина канала подъязычного нерва варьирует практически в одних и тех же пределах даже у человека и шимпанзе [Fitch, 2000. P. 262; Lieberman, 2002. P. 177]; кроме того, отсутствует корреляция между толщиной канала подъязычного нерва и числом аксонов, из которых этот нерв состоит [DeGusta et al., 1999].

В обеспечении членораздельной звучащей речи участвует подъязычная кость. У человека она расположена ниже, чем у других приматов, благодаря чему сильно расширяется спектр возможных движений глотки, гортани и языка друг относительно друга. Если бы подъязычная кость располагалась у нас иначе, мы были бы способны произносить не больше различающихся звуков, чем, например, шимпанзе. Как была расположена эта кость у других представителей класы человека – неизвестно, поскольку эта небольшая кость, не прикреплённая к остальному скелету, обычно не сохраняется. К настоящему времени исследователи располагают лишь очень немногими образцами подъязычной кости. Наиболее известны найденная в Израиле (пещера Кебара) подъязычная кость неандертальца [Arensburg et al., 1989] и подъязычная кость гейдельбергского человека из Испании (район Атапуэрка, пещера Сима де лос Уэсос); кроме того, в пещере Сидрон в Астурии, близ Пилона, была найдена частично сохранившаяся подъязычная кость гоминида, принадлежавшего к неандертальцам (или гейдельбержцам) [Cabo et al., 2003]. Все эти кости, хотя и несколько отличаются по своему строению, чрезвычайно сходны с теми, которые можно наблюдать у современных людей [Зубов, 2004] (в частности, в них отсутствуют отверстия для горловых мешков, свойственных современным шимпанзе [Fitch, 2000]), и это даёт возможность высказать гипотезу, что голосовые аппараты неандертальца, гейдельбержца и современного человека

анатомически чрезвычайно близки. Напротив, найденная недавно подъязычная кость афарского австралопитека оказалась такой же, как у шимпанзе [Alemseged et al., 2006].

Не менее важен для использования звучащей речи тонкий контроль дыхания [MacLarnon, Hewitt, 1999]. Дело в том, что при речи, в отличие от нечленораздельного крика, воздух надо подавать на голосовые связки не сразу, а небольшими порциями – слогами. Это позволяет строить длинные высказывания на одном выдохе, перемежая его короткими вдохами в моменты значимых для смысла и/или синтаксиса пауз. В рамках одного такого высказывания можно произнести большое количество слогов – тем самым, возникает эволюционная задача снабдить эти слоги необходимым количеством различий, что позволит придать высказыванию большую информационную наполненность. Если бы воздух подавался на голосовые связки сразу весь, возможности изменения звучания в процессе одного выдоха-высказывания были бы крайне ограничены (читатель может убедиться в этом сам, попытавшись снабдить членораздельными изменениями звучания, скажем, вопль ужаса). Как следствие, в таком языке оказалось бы очень мало слов: слишком малые возможности варьирования звука не позволяли бы проводить большое число различий. Кроме того, при произнесении фонем сближающиеся органы артикуляции ослабляют акустическую мощность – по-разному в разных случаях, так что при одинаковой силе воздушного потока, подаваемого на голосовые связки, одни звуки оказались бы настолько громче других, что заглушили бы последние (при восприятии наблюдается «маскирующий» эффект: тихий звук, непосредственно предшествующий громкому или следующий за ним, не распознаётся). Таким образом, речевое дыхание должно не только квантовать выдох на слоги, но и регулировать силу выдоха в рамках одного слога так, чтобы соседствующие звуки не заглушали друг друга. Как было показано ещё Н.И. Жинкиным [Жинкин, 1958] при помощи рентгенокимографической съёмки, это обеспечивается движениями диафрагмы: «в процессе речевого произнесения диафрагма на выдохе совершает резкие и отчётливо заметные вдыхательные и выдыхательные движения. Она модулирует с определённой амплитудой на каждом речевом звуке, то поднимаясь вверх, то опускаясь вниз, при этом экспирация не прекращается» [Жинкин, 1958]. Например, при произнесении слова *скалы* «на слоге *ска* диафрагма вначале делает два движения вверх (*ск*), потом опускается на *а*. После этого идёт краткое нижнее слогостояние диафрагмы и новый слог *лы*,

который начинается малым подъёмом диафрагмы на *л* и вторым, большим подъёмом на *ы*» [там же], при этом «в момент падения диафрагмы на слогоразделе вдоха не происходит» [там же]. Диафрагма иннервируется диафрагмальными нервами, отходящими от шейного отдела спинного мозга на уровне третьего, четвёртого и пятого шейных позвонков. В речевом дыхании задействованы также межрёберные мышцы, которые иннервируются из грудных спинно-мозговых нервов. Таким образом, для эффективного управления дыханием при речи необходим достаточно широкий позвоночный канал. По имеющимся данным, у неандертальца и гейдельбергского человека этот канал был примерно таким же широким, как у неоантропа, тогда как у архантропа и даже у *Homo antecessor* – значительно уже [MacLarnon, Hewitt, 1999; Walker, Shipman, 1996; Дробышевский, 2004].

Иногда можно встретить утверждение о том, что значительную роль для членораздельной речи играет подбородочный выступ. Но это не вполне верно. Подбородочный выступ – это просто результат неравномерной редукции челюстей, происходившей в процессе эволюции человека. Другое дело, что при развитии речи мышцы языка совершали всё больше разнообразных тонко дифференцированных движений, и именно необходимость в прикреплении этих мышц, возможно, убрала нижнюю челюсть от редукции. Более того, на ней возникли подбородочные ости и выступ. В становлении же членораздельной речи сыграл роль не подбородочный выступ как таковой, а изменение способа прикрепления подбородочно-язычной мышцы с мясистой на сухожильный. Впрочем, как отмечал В.В. Бунак, для развития членораздельной речи уменьшение размеров нижней челюсти сыграло положительную роль, поскольку «при быстрой смене артикуляции массивная нижняя челюсть и мускулатура создавали бы большую инерцию в работе речевого аппарата, основанной, как известно, именно на быстрой смене артикуляции» [Бунак, 1980]. Из всех гоминид подбородочный выступ имеется лишь у неоантропа, но у некоторых неандертальцев прослеживается определённая тенденция к его формированию [Зубов, 2004]. Присутствует в зачаточном состоянии подбородочный выступ, например, у неандертальца Кебара 2, обладателя сохранившейся подъязычной кости. Впрочем, для выводов о системе коммуникации, возможно, более показательно строение внутренней поверхности нижней челюсти [Дробышевский С.В., устное сообщение]: в её середине (в области симфиза) у человека имеются подбородочные ости (места прикрепле-

ния подбородочно-язычной мышцы); у обезьян же на этом месте ямка (поскольку эта мышца прикрепляется к кости не сухожилием, а мясистой частью). Челюсти ископаемых гоминид демонстрируют спектр переходных форм.

Анатомические изменения, связанные с развитием членораздельной звучащей речи, коснулись не только речевого аппарата. У человека иначе, чем, например, у шимпанзе, устроен слуховой анализатор. У нас имеется область лучшей слышимости в диапазоне от 2 до 4 кГц – именно на этих частотах сосредоточены значимые характеристики фонем. Шимпанзе же лучше всего слышат звуки частотой около 1 кГц – для них это очень важно, поскольку примерно такую частоту имеют их «долгие крики» (один из типов коммуникативных сигналов). Настройка слуха на повышенную чувствительность к высоким частотам происходила ещё у предков человека современного типа – у *Homo heidelbergensis*. И. Мартинес и его коллеги исследовали слуховые косточки *Homo heidelbergensis*, найденных в Испании (местонахождение Сима-де-лос-Уэсос) и реконструировали, каким слухом обладали представители этого вида [Martinez et al., 2004]. Оказалось, что область лучшей слышимости в диапазоне от 2 до 4 кГц к этому времени уже начала формироваться, но ещё не установилась полностью, поскольку в степени её развития разные индивиды обнаруживают достаточно значительную вариативность.

Анализатор речевых звуков работает у человека чрезвычайно быстро (быстрее, чем распознаются неречевые звуки) – до 20–30, а при искусственном ускорении речи – до 40–50 фонем в секунду [Пинкер, 2004]. При этом люди могут проводить достаточно тонкие фонетические различия, например, мы способны не перепутать такие похожие звуки, как *b* и *p* (физически отличающиеся друг от друга тем, что колебания голосовых связок начинаются либо одновременно с тем, как разомкнутся губы, либо после этого).

Важным свойством человеческой коммуникации является то, что она подконтрольна воле, а не эмоциям (т.е. управляется структурами коры больших полушарий): чтобы заговорить, нам не обязательно приходить в сильное возбуждение (это скорее помешает), надо лишь захотеть нечто сказать.

Главную роль в обеспечении функционирования языка играют две области левого (в норме у правой) полушария: зона Брока и зона Вернике. Зона Вернике, соседствующая с зоной визуального распознавания, хранит образы отдельных языковых элементов, зона Брока, прилегающая к премоторной коре, – программы обращения с

ними. Но не менее важны и другие отделы мозга, в особенности лобные доли – они обеспечивают возможность подавлять лишние эмоции, а также сосредоточивать внимание на главном, отвлекаясь при этом от несущественных деталей [см., например, Дерягина, 2003]. В отсутствие такой возможности люди никогда не смогли бы распознать, например, какие элементы фонетической реализации звука несут смысловоразличительную нагрузку, а какие – нет. При поражениях лобных долей человек не теряет дара речи, но утрачивает возможность строить поведение по словесной инструкции [см., например, Шульговский, 2003]. Если поражены префронтальные отделы коры, больной может повторять слова и целые фразы, но не в состоянии самостоятельно высказать какую-либо мысль или задать вопрос. Следует отметить, что в коре больших полушарий связаны между собой самые разные стороны восприятия одного и того же объекта: его внешний вид, запах и вкус (если он их имеет), звуки, называющие этот объект, звуки, производимые этим объектом (если он производит звуки), ощущение этого объекта в руке (если его можно взять в руку), представление о манипуляциях с ним и т.д. Словом, все то, что позволяет нам, видя (слыша, обоняя, осязая) этот объект, понимать, чего от него можно ожидать, что с ним можно (или даже нужно) делать, а чего нельзя [ср.: Calvin, Bickerton, 2000]. В хранение наших знаний о различных объектах оказываются вовлечены те отделы мозга, которые регулируют связанное с этими объектами поведение: например, в распознавании инструментов участвует премоторная кора, которая управляет рабочими движениями, а «при категоризации и назывании изображений животных, напротив, активируются прежде всего затылочно-височные области, ответственные за сложные формы зрительной обработки и восприятие движения» [Величковский, 2006].

Это оказывается важным при распознавании звучащей речи. В мозге, как показывают данные мозгового картирования, существуют специальные участки, предназначенные для обработки речевых звуков (отличные от тех, которые используются для распознавания звуков неречевых) [Hickok, Poeppel, 2000; Poeppel, 2001; Trout, 2001; Vouloumanos et al., 2001]. Эти участки позволяют обнаруживать различные простые характеристики акустических событий: наличие звучания на определенной частоте, увеличение энергии звука, уменьшение энергии звука, скорость изменения энергии звука, повышение частоты, понижение частоты и некоторые другие [Кодзасов, Кривнова, 2001]; различные комбинации показаний детекто-

ров складываются в смысловоразличительные признаки фонем (набор которых уникален для каждой фонемы). Но человек, воспринимая и понимая речь, не распознаёт фонему за фонемой, складывая их потом в слова (подобно компьютеру, побуквенно распознающему отсканированный текст), распознавание речи устроено значительно сложнее. Во-первых, от звука к звуку в потоке речи имеются вполне акустически заметные переходы (поэтому, если звуки в слоге поменять местами, человек услышит не слог, произнесённый наоборот, а бессмысленную абракадабру – за счёт того, что привычные ему правила перехода от звука к звуку не будут соблюдены). Формантные переходы между соседствующими звуками нередко позволяют людям «услышать» нужный звук даже в том случае, когда он не был реально произнесён [Sinnott, Saporita, 2000], – и человек вполне может не осознать, что вместо, скажем, *Он – человек ответственный* услышал ...*чек отвесный*. Во-вторых, звуки речи встречаются – если не брать экспериментов – в словах, а «информация, достаточная для опознания слова по звуковому облику, включает в себя его общую длину, просодический контур, несколько гласных и согласных звуков, следующих друг за другом в определенном порядке» [Бурлак, Старостин, 2005]. Кроме того, слова употребляются в высказываниях, а высказывания – в тех или иных жизненных ситуациях, тем самым количество «контекста» (как языкового, так и внеязыкового) увеличивается. К распознаванию звуков речи может подключаться также зрительный анализатор, чему свидетельством известен «эффект Мак-Гурка»: если дать человеку прослушать слог *ba* и при этом показать ему губы, произносящие *ga*, он, автоматически сделав соответствующую поправку, воспримет услышанное как слог *da* [McGurk, MacDonald, 1976] (разомкнутые губы не могли произносить *b*, а шум на тех частотах, которые характерны для *b*, можно с некоторым напряжением принять за *d*, но никак не за *g*). Все это позволяет людям понимать друг друга даже при наличии ошибок.

Всем этим анатомо-физиологическим приспособлениям сопутствуют приспособления когнитивные: дети приходят в мир с желанием обнаруживать слова, т.е. интерпретировать звуки, произносимые окружающими, как знаки. Желание слышать понимаемую речь столь велико, что подчас заставляет человека обнаруживать слова в шумах природы (например, песня птицы чечевицы описывается обычно как «Витю видел?»). Даже новорожденные младенцы предпочитают звуки речи (включая такие, которые они не могли различать в утробе) неречевым звукам, похожим на

речевые по своему тембру и ритмике [Vouloumanos, Werker, 2004a; 2004b]. Большое значение для развития речи имеет способность к звукоподражанию: притом, что люди, в массе своей, достаточно плохие имитаторы (далеко не каждый способен адекватно изобразить пение синицы, мяуканье кошки или правильно воспроизвести даже несложную мелодию), подражание речевым звукам даётся им гораздо лучше: уже к трём-четырёх-пяти годам дети научаются правильно произносить все согласные и гласные своего родного языка, воспроизводить тоны (в тех языках, где они есть), интонационную структуру разных типов предложений и т.д. Важно, что звукоподражание это является самоподдерживающимся: детям, овладевающим языком, не нужно специальное поощрение за выученные элементы коммуникативной системы.

Формирование всего этого комплекса признаков начинается лишь в гоминидной линии. У современных человекообразных обезьян высоко расположена гортань (что позволяет есть и дышать / издавать звуки практически одновременно), имеются горловые мешки, иначе, чем у человека, крепится к нижней челюсти подбородочно-язычная мышца, отсутствует произвольная регуляция дыхания (зато имеется возможность издавать звуки как на выдохе, так и на вдохе), звукопроизводство подконтрольно эмоциям, а не воле (и управляется подкорковыми структурами мозга), слуховой анализатор настроен на преимущественное распознавание звуков только на низких частотах, желание обнаруживать слова, видеть за звуками имена всего сущего не отмечается даже в языковых проектах, способности и склонности к звукоподражанию нет. При этом у обезьян имеются гомологи зоны Брока и зоны Вернике [Зорина, Смирнова, 2006], соответствующие человеческим не только по своему расположению, но и по клеточному составу, а также по входящим и исходящим нейронным связям; кроме того, эти области – как у человека, так и у человекообразных обезьян – соединены между собой пучком волокон (это было показано как отечественными, так и зарубежными исследователями [Кочеткова, 1973; Galaburda, Pandya, 1982; Deacon, 1997]). Гомолог зоны Брока «отвечает» за автоматические комплексные поведенческие программы, осуществляемые мышцами лица, рта, языка и гортани, а также за координированные программы действий правой руки [Greenfield, 1991; Rizzolatti, Arbib, 1998]. Гомолог зоны Вернике (и соседние участки мозга) используются для распознавания звуковых сигналов, а также для того, чтобы различать сородичей по голосу. Кроме того, «различные подобласти этих гомологов получают данные от всех

частей мозга, задействованных при слушании, ощущении прикосновения во рту, языке и гортани и областях, где сливаются потоки информации от всех органов чувств» [Пинкер, 2004]. Таким образом, обнаружение на эндокранах ископаемых гоминоид развитых выпуклостей в области зоны Брока и зоны Вернике не может свидетельствовать о наличии у соответствующего вида членораздельной звучащей речи [ср.: Tobias, 1996].

При этом у *Homo sapiens* весь этот комплекс признаков закреплён в геноме, и, поскольку появление такого сложного комплекса признаков в результате единственной макромутации совершенно невозможно, это означает, что перед видом-предком задача приспособления к членораздельной звучащей речи уже стояла (и основу вида-потомка положили те, кто смог её успешно решить).

В качестве «гена речи» нередко рассматривается ген *FOXP2*, расположенный на седьмой хромосоме. Люди, имеющие дефектный вариант этого гена, страдают специфическим расстройством речи (англ. SLI), затрагивающим как фонетическую, так и грамматическую составляющие языка [см., например, Бурлак, 2011. С. 255, 256]. Кроме того, у них несколько нарушен волевой моторный контроль за мышцами рта (в области как отдельных движений, так и их последовательностей) – им сложно, например, по команде высунуть язык или несколько раз подряд сомкнуть и разомкнуть передние зубы [Alcock et al., 2000]. Как показали исследования, ген *FOXP2* является геном-регулятором высокого уровня (т.е. регулирует активность других генов-регуляторов) [Konopka et al., 2009]; он экспрессируется в различных отделах мозга, в частности, влияет на характер нейронных связей между корой больших полушарий и базальными ядрами, повышая синаптическую пластичность в них [Reimers-Kipping et al. 2011], что является чрезвычайно существенным для возможности выучивать последовательности действий [Jin, Costa, 2010].

Ген *FOXP2*, по всей видимости, был мишенью отбора в гоминоидной линии: со времени разделения предков людей и шимпанзе в этом гене произошли две замены, и обе они являются несинонимичными [Enard et al., 2002]. Как показали недавние исследования, таким же, как у нас, вариантом этого гена обладали и неандертальцы [Krause et al., 2007]. Тем самым, велика вероятность, что «человеческие» мутации в этом гене произошли уже у общего предка неандертальца и *Homo sapiens*, т.е. у *Homo heidelbergensis*, расширив его возможности как в области моторного контроля за органами артикуляции, так и в области автоматизации последовательностей действий,

что особенно важно для тех, кто может произносить длинные, многосложные реплики и перед кем, соответственно, стоит задача наделять эти реплики необходимым количеством различий между соседствующими элементами.

Важное значение для определения времени формирования членораздельной звучащей речи имеет недавнее исследование Б. де Бура [de Boer, 2011], направленное на выявление функции горловых мешков. Эти образования имеются у современных шимпанзе (и имелись, судя по строению подъязычной кости, у афарских австралопитеков), но отсутствуют у людей (а также у неандертальцев и *Homo heidelbergensis*). Построив модель речевого резонатора с горловыми мешками и без них, он показал, что при наличии горловых мешков во-первых, резонансы речевого тракта смещаются ближе друг к другу, а во-вторых, появляются дополнительные резонансы и антирезонансы – причём появляются независимо от производимой артикуляции. Из этого сразу видна отрицательная роль горловых мешков для членораздельности речи. Во-первых, если все области усиления звука приближены друг к другу, это значит, что звуки получаются более похожими один на другой, тогда как для членораздельной речи необходимо, наоборот, чтобы звуки достаточно сильно различались. Увеличение различающихся на слух звуков позволяет иметь коммуникативную систему с большим числом знаков (и тем самым, с большими выразительными возможностями). Во-вторых, наличие резонансов и антирезонансов, независимых от производимой артикуляции, сильно сужает возможности произвольного варьирования производимого звука. Такая задача актуальна для обезьян, которые, имея высоко расположенную гортань, могут есть и вокализовать одновременно: при наличии горловых мешков еда, находящаяся во рту, не мешает издавать необходимые звуки. Но членораздельная речь обуславливает противоположную задачу: при помощи органов артикуляции, доступных волевому управлению, обеспечить как можно большее количество различий в звучании. Ещё одна функция горловых мешков – понижение высоты звука. Эта задача также актуальна для обезьян, которые используют звуковую коммуникацию для общения с сородичами, находящимися сравнительно далеко и скрытыми густой листвой тропического леса (в общении на близком расстоянии более существенную роль играет мимика, жесты, позы и разнообразные прикосновения) и – в связи с этим – имеют слуховой анализатор, настроенный на преимущественное распознавание низкочастотных звуков. Но для гоминоид, обитающих в полутропических и открытых ландшаф-

тах, эта задача постепенно теряет актуальность. Слуховой анализатор *Homo heidelbergensis*, демонстрирующий формирование дополнительной области лучшей слышимости на высоких частотах, свидетельствует, что на первый план у гоминид постепенно выходило звуковое общение на близком расстоянии.

Всё это позволяет в общих чертах представить себе картину формирования звучащей речи в ходе эволюции. Первоначально основными носителями намеренно передаваемой сигнальной информации у гоминид, как и у современных приматов, вероятно, были жесты – они подчинены волевому контролю и могут использоваться для создания сигналов *ad hoc* (сигналов, создаваемых по ходу дела, у которых ни форма, ни смысл не являются врождёнными). Звуки же могли использоваться только в качестве эмоционального дополнения. Но когда объем манипулятивной активности гоминид возрос, в частности, за счет всё более частого изготовления и применения орудий, сочетание обычной и коммуникативной деятельности стало затруднено [Бутовская, 2004]: руки не могли одновременно делать орудия и знаки, мозгу приходилось выбирать, какой сигнал посылать на руки, какую информацию обрабатывать – от практических движений или от сигнальных (подобные затруднения легко смоделировать, попытавшись говорить и одновременно с этим жевать жвачку). Это, по-видимому, привело к эффекту замещения – сигнал из мозговых структур, управляющих коммуникацией, стал подаваться не только на руки, но и на органы звукопроизводства (и зона Брока, соседствующая с премоторной корой, получила в своё распоряжение контроль над вокализацией). Такое замещение могло быть облегчено тем, что у приматов управление ротовым аппаратом и управление руками связаны между собой, поскольку эти органы вместе задействованы в питании, груминге и т.д. [см., например, Pollick, de Waal, 2007].

Как было показано ещё советским археологом-экспериментатором С.А. Семёновым, самое примитивное «рубящее орудие возникало после двух-трёх ударов оббивки» [Семёнов, 1968], а для получения отщепы с острым краем бывает достаточно и одного (сходные результаты были получены и позднее, Н. Тотом и К. Шик, в том числе, в экспериментах, где орудия олдувайского типа изготавливал бонобо Канзи [Schick, Toth, 1993]); ашельская же технология предполагает гораздо большие трудозатраты: орудие обрабатывается со всех сторон, при этом сначала производится грубая оббивка каменным молотком, а затем края ретушируются с использованием мягких отбойни-

ков из кости или рога (подробнее о технологиях изготовления палеолитических орудий см. [Inizan et al., 1999; Schick, Toth, 1993]). Тем самым, с развитием технологии производство орудий требовало всё большего времени. Если при этом необходимо было коммуницировать, то в выигрыше оказывались те, кто мог ещё до начала собственно передачи информации (т.е. использования значащего, подконтрольного воле жеста), по звуку общего возбуждения или привлечения внимания догадаться (хотя бы в какой-то степени) о том, что будет сообщено. В принципе такое не невозможно: так, человек, слыша обращение к себе по имени, может по интонации предугадать часть смысла будущего сообщения – намерен ли говорящий просить его о чём-то, угрожать ему, стыдить, подзывать к себе, сообщить о каком-то поразившем его событии и т.д.; при становлении речи у ребёнка овладение интонацией происходит до овладения словами [Лепская, 1997, Цейтлин, 2000].

Соответственно, отбором будет поощряться всё более вариабельный исходный сигнал и всё более точное «угадывание» другими особями по этому сигналу смысла планируемого сообщения. В этом случае информационная нагрузка переместится на звуковой канал, использование же прочих каналов редуцируется. При этом увеличится важность для коммуникации высоких частот, поскольку общение происходит на близком расстоянии.

Вероятно, «архантропы вследствие массивности челюстей могли произносить только небольшое число различных выкриков» [Бунак, 1980]. Поскольку у них не было возможности произносить длинные высказывания (вследствие недостаточных анатомо-физиологических средств для управления дыханием), они могли общаться при помощи голофраз (однословных высказываний), возможно, как современные дети, не столько описывая таким образом те или иные ситуации, сколько выражая свои эмоции по их поводу [Лепская, 1997].

У *Homo heidelbergensis* ситуация сильно изменилась: увеличенный позвоночный канал свидетельствует о том, что они были в состоянии произносить реплики из нескольких слогов, сочетая при этом по-разному артикулируемые звуки. Это имеет смысл только при наличии возможности вкладывать в звучание максимум различий (максимизируя тем самым передаваемую информацию). И действительно, отсутствие в подъязычной кости отверстий для горловых мешков, а также настройка слуха на высокие частоты показывают, что приспособления для лучшего различения звуков при помощи артикуляции коснулись как

звукопроизводства, так и звуковосприятия. Вероятно, это имело генетическую подоснову, обеспеченную мутациями в гене *FOXP2*. Таким образом, можно утверждать, что в основе коммуникации гейдельбергского человека лежала звучащая речь, в которой различия между звуками обеспечивались при помощи артикуляции. Косвенным свидетельством этого может служить и увеличение у *Homo heidelbergensis* базикраниального угла (по сравнению с тем, что было у архантропов). Но всё же, по-видимому, эта речь ещё не была настоящим человеческим языком. Для языка очень важна способность делать выводы из нескольких посылок одновременно, сосредоточивать внимание на главном, отвлекаясь от несущественного (в том числе это касается и чисто звуковых различий), держать в оперативной памяти достаточно много единиц, чтобы мочь обобщить синтаксические правила, определённые на длинных предложениях. Всё это обеспечивают лобные доли коры больших полушарий головного мозга, которые у гейдельбергского человека были гораздо меньше, чем у *Homo sapiens* [см. Дробышевский, 2007].

Итак, анатомически современный человек, появившийся не менее, чем 195 ± 5 тыс. лет назад [McDougall et al., 2008], уже, по-видимому, пользовался настоящим человеческим языком. Но база для этого, овладение членораздельностью звучащей речи, была заложена ещё предшествующим видом, *Homo heidelbergensis*, на несколько сотен тысячелетий раньше.

Благодарность

Работа выполнена при финансовой поддержке гранта РФФИ № 11-06-12036-офи-м-2011.

Библиография

- Бунак В.В. Род *Номо*, его возникновение и последующая эволюция. М.: Наука, 1980.
- Бурлак С.А. Происхождение языка: Факты, исследования, гипотезы. М.: CORPUS, 2011.
- Бурлак С.А., Старостин С.А. Сравнительно-историческое языкознание. М.: Академия, 2005.
- Бутовская М.Л., Файнберг Л.А. У истоков человеческого общества. М.: Наука, 1993.
- Величковский Б.М. Когнитивная наука: Основы психологии познания: в 2 т. Т. 2. М.: Смысл: Издательский центр «Академия», 2006.
- Вишняцкий Л.Б. Человек в лабиринте эволюции. М.: Весь Мир, 2004.
- Дерягина М.А. Эволюционная антропология: биологические и культурные аспекты: учебное пособие. 2-е изд., испр. М.: Изд-во УРАО, 2003.
- Дробышевский С.В. Предшественники. Предки? Часть III: Архантропы. Часть IV: Гоминиды, переходные от архантропов к палеоантропам. М.: Едиториал УРСС, 2004.
- Дробышевский С.В. Эволюция мозга человека. Анализ эндокраниометрических признаков гоминид. М.: УРСС, 2007.
- Жинкин Н.И. Механизмы речи. М.: Издательство Академии педагогических наук, 1958.
- Зайцева Г.Л. Система жестового обучения глухих // Язык в океане языков / Сост. Донских О.А. Новосибирск: «Сибирский хронограф», 1993. С. 28–42.
- Зорина З.А., Смирнова А.А. О чем рассказали «говорящие» обезьяны: Способны ли животные оперировать символами? М.: Языки славянских культур, 2006.
- Зубов А.А. Палеоантропологическая родословная человека. М.: Россельхозакадемия, 2004.
- Кодзасов С.В., Кривнова О.Ф. Общая фонетика. М.: РГГУ, 2001.
- Кочеткова В.И. Палеоневрология. М.: Изд-во Моск. ун-та, 1973.
- Лепская Н.И. Язык ребенка (онтогенез речевой коммуникации). М.: Филол. фак. МГУ, 1997.
- Пинкер С. Язык как инстинкт. М.: УРСС, 2004.
- Семенов С.А. Развитие техники в каменном веке. Л.: Наука. Ленинградское отделение, 1968.
- Цейтлин С.Н. Язык и ребенок. Лингвистика детской речи. М.: Гуманитарный издательский центр ВЛАДОС, 2000.
- Шульговский В.В. Физиология высшей нервной деятельности с основами нейробиологии. М.: Академия, 2003.
- Alcock K.J., Passingham R.E., Watkins K.E., Vargha-Khadem F. Oral dyspraxia in inherited speech and language impairment and acquired dysphasia // *Brain and Language*, 2000. Vol. 75. P. 17–33.
- Alemseged Z., Spoor F., Kimbel W.H., Bobe R., Geraads D., Reed D., Wynn J.G. A juvenile early hominin skeleton from Dikika, Ethiopia // *Nature*, 2006. Vol. 443. P. 296–301.
- Arensburg B., Tillier A.M., Vandermeersch B., Duda H., Schepartz L.A., Rak Y. A Middle Palaeolithic human hyoid bone // *Nature*, 1989. Vol. 338. P. 758–760.
- Cabo L.L., Rodriguez L., Egocheaga J.E. Breve nota sobre el hioides neandertalense de Sidron (Piloa, Asturias) // *Antropología y biodiversidad: Actas XII Congreso de la SEAB*. Vol. 1 / Coord. por Malgosa A., Nogués R.M., Aluja M.P. Barcelona: Ediciones Bellaterra, 2003. Vol. 1. P. 484–493.
- Calvin W.H., Bickerton D. *Lingua ex machina: Reconciling Darwin and Chomsky with the human brain*. Cambridge: MIT Press, 2000.
- Crelin E. *The Human Vocal Tract: Anatomy, function, development and evolution*. New York: Vantage Press, 1987.
- Deacon T. *The symbolic species: The co-evolution of language and the brain*. N.Y.; L.: W.W. Norton & Company, Inc., 1997.
- de Boer B. Loss of air sacs improved hominin speech abilities // *J. Human Evol.*, 2011. P. 1–6.

- DeGusta D., Gilbert W.H., Turner S.P. Hypoglossal canal size and hominid speech // *Proc. National Academy of Science, USA*, 1999. Vol. 96. P. 1800–1804.
- Enard W., Przeworski M., Fisher S.E., Lai C.S.L., Wiebe V., Kitano T., Monaco A.P., Pääbo S. Molecular evolution of FOXP2, a gene involved in speech and language // *Nature*, 2002. Vol. 418. P. 869–872.
- Galaburda A.M., Pandya D.N. Role of architectonics and connections in the study of primate brain evolution // *Primate brain evolution: Methods and concepts* / Ed. by Armstrong E., Falk D. N.Y.: Plenum Press, 1982. P. 203–216.
- Greenfield P.M. Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior // *Behavior and brain sciences*, 1991. Vol. 14. P. 531–595.
- Hauser M.D., Chomsky N., Fitch W.T. The faculty of language: What is it, who has it, and how did it evolve? // *Science*, 2002. Vol. 298. P. 1569–1579.
- Hickok G., Poeppel D. Towards a functional neuroanatomy of speech perception // *Trends in Cognitive Sciences*, 2000. Vol. 4. N 4. P. 131–138.
- Inizan M.-L., Reduron-Ballinger M., Roche H., Tixier J. Technology and terminology of knapped stone. Nanterre : CREP, 1999.
- Jin X., Costa R.M. Start/stop signals emerge in nigrostriatal circuits during sequence learning // *Nature*. – 2010. Vol. 466. P. 457–462.
- Kay R.F., Cartmill M., Balow M. The hypoglossal canal and the origin of human vocal behavior // *Proc. National Academy of Science, USA*, 1998. Vol. 95. P. 5417–5419.
- Konopka G., Bomar J.M., Winden K., Coppola G., Jonsson Z.O., Gao F., Peng S., Preuss T.M., Wohlschlegel J.A., Geschwind D.H. Human-specific transcriptional regulation of CNS development genes by FOXP2 // *Nature*, 2009. Vol. 462. P. 213–217.
- Krause J., Lalueza-Fox C., Orlando L., Enard W., Green R.E., Burbano H.A., Hublin J.-J., Hänni C., Fortea J., de la Rasilla M., Bertranpetit J., Rosas A., Pääbo S. The derived FOXP2 variant of modern humans was shared with Neandertals // *Current Biology*, 2007. Vol. 17. P. 1908–1912.
- Laitman, J.T., Crelin, E.S. Postnatal development of the basicranium and vocal tract region in man // *Symposium on Development of the Basicranium* / Ed. by Bosma J.F. Bethesda: Department of Health, Education and Welfare, 1976. P. 206–219.
- Lieberman P. Human language and our reptilian brain: The subcortical bases of speech, syntax and thought. Cambridge (Mass.): Harvard Univ. Press, 2002.
- Lieberman P., Crelin E.S. On the speech of Neanderthal man // *Linguistic Inquiry*, 1971. Vol. 2 N 2. P. 203–222.
- Lieberman P., Crelin E.S., Klatt D.H. Phonetic ability and related anatomy of the newborn and adult human, Neanderthal man, and the chimpanzee // *American Anthropologist*, 1972. Vol. 74. P. 287–307.
- Lieberman D.E., McCarthy R.C. The ontogeny of cranial base angulation in humans and chimpanzees and its implications for reconstructing pharyngeal dimensions // *J. Human Evol.*, 1999. Vol. 36. N 5. P. 487–517.
- MacLarnon A.M., Hewitt G.P. The evolution of human speech: The role of enhanced breathing control // *Amer. J. Physical Anthropol.*, 1999. Vol. 109. N 3. P. 341–363.
- Martinez I., Rosa M., Arsuaga J.-L., Jarabo P., Quam R., Lorenzo C., Gracia A., Carretero J.-M., Bermudez de Castro J.-M., Carbonell E. Auditory capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain // *Proc. National Academy of Science, USA*, 2004. Vol. 101. N 27. P. 9976–9981.
- McDougall I., Brown F.H., Fleagle J.G. Saporrels and the age of hominins Omo I and II, Kibish, Ethiopia // *J. Human Evol.*, 2008. Vol. 55. N 3. P. 409–420.
- McGurk H., MacDonald J. Hearing lips and seeing voices // *Nature*, 1976. Vol. 264. P. 746–748.
- Nishimura T., Mikami A., Suzuki J., Matsuzawa T. Descent of the larynx in chimpanzee infants // *Proc. National Academy of Science, USA*, 2003. Vol. 100. N 12. P. 6930–6933.
- Pinker S. Language as an adaptation to the cognitive niche // *Language evolution* / Ed. by Christiansen M.H., Kirby S. Oxford: Oxford Univ. Press, 2003. P. 16–37.
- Poeppel D. Pure word deafness and the bilateral processing of the speech code // *Cognitive Science*, 2001. Vol. 21. N 5. P. 679–693.
- Pollick A.S., de Waal F.B.M. Ape gestures and language evolution // *Proc. National Academy of Sciences, USA*, 2007. Vol. 104. N 19. P. 8184–8189.
- Reimers-Kipping S., Hevers W., Pääbo S., Enard W. Humanized FOXP2 specifically affects cortico-basal ganglia circuits // *Neuroscience*, 2011. Vol. 175. P. 75–84.
- Rizzolatti G., Arbib M.A. Language within our grasp // *Trends in neurosciences*, 1998. Vol. 21. N 5. P. 188–194.
- Schick K.D., Toth N. Making silent stones speak: Human evolution and the dawn of technology. London: Weidenfeld and Nicolson, 1993.
- Sinnott J.M., Saporita T.A. Differences in American English, Spanish, and monkey perception of the say-stay trading relation // *Perception and Psychophysics*, 2000. Vol. 62. N 6. P. 1312–1319.
- Stokoe W.C. Sign language structure: An outline of the visual communication systems of the American deaf / *Studies in linguistics. Occasional papers*. Vol. 8. Buffalo, N.Y.: University of Buffalo, 1960.
- Tobias P.V. The brain of the first hominids // *Origins of the human brain* / Ed. by Changeux J.-P.; Chavillon J. Oxford: Oxford Univ. Press, 1996. P. 61–81.
- Trout J.D. The biological basis of speech: What to infer from talking to the animals // *Psychological Review*, 2001. Vol. 108. N 3. P. 523–549.
- Vouloumanos A., Kiehl K.A., Werker J.F., Liddle P.F. Detection of sounds in the auditory stream: Event-related fMRI evidence for differential activation to speech and nonspeech // *J. Cognitive Neuroscience*, 2001. Vol. 13. N 7. P. 994–1005.
- Vouloumanos A., Werker J.F. A neonatal bias for speech that is independent of experience. Paper presented at the Fourteenth Biennial international conference on infant studies. Chicago. 2004a.
- Vouloumanos A., Werker J.F. Tuned to the signal: The privileged status of speech for young infants // *Developmental science*, 2004b. Vol. 7. P. 270–276.
- Walker A., Shipman P. The wisdom of the bones: In search of human origins. N.Y.: Alfred A. Knopf, 1996.

Контактная информация:

Бурлак Светлана Анатольевна. E-mail: svetlana.burlak@bk.ru.

WHEN DID ARTICULATE SPEECH ARISE (ACCORDING TO ANTHROPOLOGICAL DATA)?

Burlak S.

*The Institute of Oriental Studies of the Russian Academy of Sciences, Moscow State University,
Moscow*

Author sums up the data which sheds light on the time of emergence of articulate speech and considers anatomical features that, in the opinion of some researchers, are relevant for the transition to articulate speech, such as lowered larynx, basicranial angle, hypoglossal canal, form and position of hyoid bone, mental protuberance, spinal canal width. The most relevant are lowered larynx (combined with shorter oral cavity, which allows a much greater range of sounds to be produced by movements of the vocal tract) and spinal canal width (indicating the increased capacity of breath control). Author compares auditory capacities of modern humans and apes with that of fossil hominids (reconstructed on the base of ear bones form) and also their breath control and information processing abilities. For the faculty of language, not only Broca and Vernicke areas are important, frontal lobes (especially, prefrontal cortex) also play a great role. At the same time, brain size growth in Broca and Vernicke areas does not provide, per se, the decisive argument for developed sound speech in a certain Hominid species. Speech processing has a complex nature, involving brain structures recognizing not only the distinctive features of single sounds but also formant transitions from one sound to another, the whole word-gestalt, linguistic and extra-linguistic context. Author also discusses cognitive adaptations necessary for sound speech to emerge: capacity for self-sustained vocal imitation, a strong desire to hear understandable speech and interpret sounds produced by others as symbols. This whole syndrome emerges only in Hominids and cannot be due to a single macromutation. Comprehensive study of the available data makes it possible to conclude that already Homo heidelbergensis was capable of articulate speech: this species had no air sacs (which reduced distinctiveness of articulations), maintained a relatively high sensitivity from 2 kHz up to 4 kHz (a region that contains acoustic features that can be modified by articulation) and had a wider spinal canal than Homo erectus. But there is no reason to believe that this species had a full-fledged human language, because many aspects of the latter are closely associated with frontal lobes, and in Homo heidelbergensis these lobes are smaller than in Homo sapiens.

Keywords: anthropology, language origin; articulate speech, Homo heidelbergensis