

Из истории Лаборатории экспериментальной фонетики ИСАА при МГУ

Румянцев М.К.,

Каплун М.И.



Лаборатория экспериментальной фонетики (ЛЭФ) Института стран Азии и Африки при МГУ с самого её начала — с осени 1962 г. — по своему фактическому предназначению была и есть Лаборатория межкафедральная: в ней работают и учатся студенты, аспиранты и преподаватели всех филологических кафедр института. Лаборатория создавалась как единица учебно-научная: студенты работали и работают над своими курсовыми и дипломными работами, аспиранты и преподаватели — над кандидатскими, докторскими диссертациями и любыми другими экспериментальными исследованиями — статьями, монографиями, докладами.

* * *

Возникла Лаборатория на базе Фонетического кабинета Института восточных языков (ИВЯ), который использовался, как правило, лишь в учебных целях. Однако именно своими научными изысканиями на материале восточных и африканских языков Лаборатория достаточно быстро стала заметна в лингвистическом мире.

Всё в те годы как-то удивительно совпало: общая атмосфера предчувствия свободы 60-х годов будила какую-то необъяснимую тягу ко всему новому, необычному. Экспериментальная фонетика — к тому времени наука ещё очень молодая — как раз была той областью лингвистики, которая предоставляла возможность исследователю решать традиционные проблемы нетрадиционными методами. Поэтому у нас происходил невероятный **взрыв интереса к экспериментальным работам**. Вспомним, сколько докторских и кандидатских диссертаций было подготовлено и защищено в Лаборатории экспериментальной фонетики и психологии речи МГПИИЯ им. М. Тореца, в нашей Лаборатории ИВЯ МГУ, на филологическом факультете МГУ, в Ленинградском университете. Лаборатории, подобные нашей, появлялись в крупных городах бывшего Советского Союза, буквально как грибы после дождя.

Конечно, заслуживает внимания и самого пристального анализа научная проблематика того времени. Технические новшества (спектроанализаторы, интонографы и другие специальные приборы) давали возможность исследовать акустическую природу звуков, роль просодических средств языка. Особенно много было тем, связанных с интонацией. Почему? Ответ находим в словах проф. В.А. Артёмова: «Экспериментально-фо-

нетические и экспериментально-психологические исследования интонации с применением современных электроакустических приборов и электронных анализаторов в корне изменили наши представления об интонации» [Артёмов, 1966]. В те годы было защищено много диссертаций по различным аспектам интонации языков разной типологии. Вот характерные образцы: О.Г. Козьмин «Интонация побудительных предложений в современном немецком языке», З.Л. Григорова «Интонация вопросительных предложений в современном китайском языке», В.Т. Михеева «Интонация предложения с обособленным определением в современном французском языке».

* * *

Характерной приметой того времени была естественная **тяга к научной кооперации**. Ярким примером такой кооперации было плодотворное сотрудничество двух Лабораторий — МГПИИЯ им. М. Тореза (научный руководитель проф. В.А. Артёмов) и ИВЯ МГУ (научный руководитель проф. М.К. Румянцев). Мы не только обсуждали научные проблемы на знаменитых на весь Союз артёмовских коллоквиумах, но и выпустили несколько научных сборников, в тематике которых, как в зеркале, был отражён вектор наших исследовательских поисков. В одном из сборников, например, были опубликованы материалы симпозиума по спектральному анализу звуков речи, который состоялся в 1966 году. Конечно, Лаборатория МГПИИЯ им. М. Тореза уже накопила к тому времени громадный научный потенциал, о чём свидетельствует широкий спектр публикаций сотрудников Лаборатории по самым актуальным проблемам звукового состава и просодии различных языков. ИВЯшные возможности были значительно скромнее, поэтому круг наших исследований был несколько уже и ограничивался чаще всего изучением элементов просодии: тона, ударения и интонации.

Тематика публикаций, представленных в этих сборниках, приводит к мысли, что уже тогда, в 60-е годы, инженеры, специализирующиеся на исследовании звукового состава речи (анализ, синтез), понимали необходимость тесного научного сотрудничества с лингвистами. Вот почему наличие в лингвистическом сборнике таких статей, как «Многофункциональный анализатор спектра речевых сигналов типа «Спектр-1» Б.М. Лобанова и «Анализ спектров речи при помощи ЦВМ» В.Н. Соболева было веянием времени.

К нам часто заходили посидеть, поговорить, поработать А.А. Реформатский, Е.Н. Драгунова с Т.П. Задоевко. Эти посещения всегда выливались в заинтересованное обсуждение проблем фонетики и фонологии, методов и возможностей их экспериментального исследования, особенно на материале восточных и африканских языков. Нередко в этих обсуждениях принимал участие и проф. П.С. Кузнецов, очень интересовавшийся языками Азии и Африки и ставший впоследствии постоянным рецензентом и оппонентом работ, выполненных на материале восточных и африканских языков.

Пристально интересовался проблемами Фонетического кабинета ИВЯ, его энтузиастами и В.А. Артёмов — глава московской школы фонетиков-экспериментаторов, руководитель Лаборатории экспериментальной фонетики и психологии речи 1-го Московского государственного пединститута иностранных языков (ныне Московский государственный лингвистический университет).

* * *

Ко времени организации ЛЭФ в ИВЯ один из авторов данной статьи М.К. Румянцев уже более десяти лет работал в семинаре по экспериментальной фонетике под руководством В.А. Артёмова.

Обсуждение на семинаре **фонетических и фонологических проблем экзотических языков** — китайского и др. — всегда привлекало особое внимание и вызывало немалое удивление индоевропейцев. Эта заинтересованность и удивление хорошо иллюстрировали общую ситуацию в языкознании того времени, характеризовавшуюся острым взаимным интересом друг к другу так называемого общего языкознания и языкознаний частных.

Общее языкознание, сформировавшееся на материале индоевропейских языков, уже к началу 30-х гг. вполне осознало свою фактологическую недостаточность и крайне нуждалось в достоверных данных об устройстве языков другой — нефлективной — типологии. Именно этот фактор в первую очередь и объясняет появление Лаборатории в Институте восточных языков.

Работа в «артёмовском семинаре», обсуждение фонетико-фонологической проблематики в беседах с А.А. Реформатским, П.С. Кузнецовым, А.А. Драгуновым — научным руководителем М.К. Румянцева по теме кандидатской диссертации — определило и основную проблематику исследований ЛЭФ в ИВЯ при МГУ. Она также диктовалась общей ситуацией, сложившейся в языкознании к 50-м — 60-м гг. Фонетико-фонологический раздел в языкознании к этому времени выглядел сложившейся дисциплиной с хорошо разработанным аппаратом исследования: были уже сформулированы все фундаментальные постулаты фонологического учения и было определено его отношение к фонетическому ярусу. Но всё, что было сделано в фонетике и фонологии, как у нас, так и за рубежом, основывалось на материале хотя и базовых, но элементарных единиц — гласных и согласных сегментов речи. Исследование более сложных единиц звучания: слога, слова, фразы, текста — и интерпретация их признаков в функциональном плане, по существу, в мире только начинались. Причём если чисто фонологические исследования сегментных единиц могли, в принципе, производиться «в уме», а работа с физическими коррелятами этих единиц требовала лишь метаязыка тех или иных артикуляций, то более сложные единицы, их просодическое устройство в реальном звучании нельзя уже было «взять» без инструмента, без психолингвистических экспериментов. Здесь исследователь имел дело уже не с артикуляциями, а с тем акустическим эффектом, который появлялся на выходе тех или иных артикуляций.

* * *

Все исследования ЛЭФ посвящены **просодическому устройству слога, слова, фразы и текста**. Это диктовалось императивом самого развития лингвистики 50-х — 60-х гг., её интересом к звучащей речи — интересом как теоретическим, так и (в не меньшей степени) прикладным. Во всём мире начинались работы по машинному восприятию и синтезу речи.

Большое количество работ, посвящённых слогу и слову, было продиктовано тем, что просодическое устройство слога и слова в любой человеческой речи (на любом языке) лежит в центре всего просодического яруса языка. Слог является наименьшей произносительной и наименьшей просодической единицей. С него, собственно, и начинается просодия. Гласные (сонорные) сегменты слога в речи становятся носителями самой разной лингвистической, социолингвистической и экстралингвистической информации. На уровне слога выявляются кардинальные типологические особенности языков как фонологического плана, так и нормативно-произносительного. Именно просодическое устройство слога делит языки мира на слоговые — тоновые и неслоговые — нетоновые.

Объединение слогов в фонетическое или морфологическое слово образует более сложную, чем слог, просодическую единицу, в которой проявляются базовые для каждого языка ритмико-акцентуационные отношения. Типы этих отношений проявляют себя и во всех просодиях, которые выше слова, и создают тот национальный колорит речи, по которому она узнаётся как принадлежащая языку такой-то просодической типологии. Сравните, например, речь на языке со слоговыми тонами, на языке со словесным ударением, на языке с преимущественно временной организацией слова. Любые интонации фразы —

коммуникативные, модальные, эмоциональные, стилевые — образуются на изначальном слоговом и словном просодическом субстрате.

Знание просодического устройства слога и слова в каждом данном языке, с одной стороны, позволяет прогнозировать просодическое поведение слова в тех или иных синтагматических и фразовых интонациях (и вообще в речевом тексте), а с другой стороны, по изменениям в просодической структуре слога и слова под влиянием тех или иных интонаций можно судить о самих этих интонациях, их акустических признаках.

Сложность функциональной интерпретации лингвистически разнонаправленных акустических признаков речевого сигнала и состоит в том, что эти признаки локализуются в звучании слогов и слов по принципу взаимного проникновения друг в друга. Иерархически старшие признаки образуются на основе признаков младших.

* * *

Важной особенностью исследований Лаборатории является то, что они основывались не только на анализе естественной речи, но и в значительной своей части на **машинном синтезе** таких единиц, как слог, слово, фраза.

Работа с синтезированными единицами речи показала, что их искусственное машинное продуцирование может быть мощным исследовательским средством, позволяющим более надёжно объективировать данные анализа естественной речи. Программируя речевой сигнал, исследователь может намеренно варьировать или заменять его акустические параметры и таким образом выяснять их функциональную нагрузку в том или ином звучании.

В этом плане большое значение имеет проблема общения человека с машиной. Её надо научить «понимать» человеческую речь и «говорить» с человеком по-человечески.

* * *

Что же, однако, для нас означало решить **проблему рабочего общения человека с машиной**? Какие лингвистические аспекты проблемы подлежали решению?

Здесь можно выделить два аспекта, две фундаментальные трудности. Одна из них связана с восприятием, или распознаванием, естественной человеческой речи машиной, другая — с производством машиной искусственной речи, понятной человеку и не отличающейся или мало отличающейся от речи естественной.

Решить проблему восприятия естественной человеческой речи машиной означало, прежде всего, научить машину сегментировать человеческую речь, т.е. выделять из потока речи разные её единицы не только по физическим их характеристикам, но и по языковому контексту, по системно-семиотическим противопоставлениям.

Человек воспринимает звучащую речь не только по её физическим параметрам. Машина же в принципе до сих пор ориентировалась только на «физику» звучания. Разработчики опираются на так называемые «островки надёжности», т.е. на такие участки речевой цепи, физическая информация которых представлена надёжно. До сих пор машинам явно не доставало умения «лингвистически» оперировать речевым материалом, такими его частками, в которых «физика» звуков смазана, нивелирована либо вообще представлена нулями звучания: в каких-то позициях и темпах речи того или иного звука речи вообще нет, а человеческое сознание, однако, легко

его восстанавливает, знает, что там должно быть. Вся человеческая речь в своём непрерывном звучании физически настолько нестационарна, что в лингвистике возникла гипотеза о её недискретности — принципиальной невозможности её сегментации (понимания) на чисто физикалистских основах.

Следует признать, что при решении проблемы, о которой идёт речь, современная лингвистика была достаточно консервативной. Машину же надо было научить производить «лингвистические» операции, что недоступно для инженера, поскольку он не знает, как это делает человек. Круг, таким образом, замыкался, и инженер оставил бы проблему машинного восприятия речи до лучших времён, если бы внял чисто лингвистическим рассуждениям. К счастью, однако, этого не произошло. Во всём мире были созданы устройства, в принципе работающие (воспринимающие речь) на чисто физикалистских и статистических параметрах.

В целом, сама возможность машинного понимания речи человека только по физике её звучания, продемонстрированная инженерами, даёт богатый материал для размышления лингвистам. Кибернетический опыт убедительно показывает, что лингвистика слишком самонадеянно недооценила фактор физической дискретности самих звуков человеческой речи и свои постулаты формулировала слишком категорично, абсолютизировала их. Между тем сегментация речи при её восприятии человеком осуществляется всё же с большой опорой на сегментацию физическую (акустическую). Можно, по-видимому, даже утверждать, что сегментация речи является результатом взаимодействия двух начал — дискретности физической и дискретности, привносимой языковым сознанием человека, которое по-своему квантует физику звуков. Чисто лингвистический характер, не подкреплённый субстанционально, такая сегментация носит лишь в тех случаях, когда акустические признаки сегмента смазаны или вообще выходят за пределы данного звукотипа. К тому же фонетически надёжные участки в речи вряд ли можно уподобить «островкам» в океане фонетически ненадёжного звучания.

Речь, её континуум можно представить в виде синусоиды с более или менее регулярными чередованиями фонетически надёжных реализаций её единиц. На каждом данном отрезке речевой цепи фонетика может нивелироваться лишь там, где рядом уже есть надёжный участок, и общий морфолого-синтаксический и семантический контекст позволяет правильно понять говоримое. Можно, вероятно, сказать и так: не будь дискретными сами звуки, не было бы и дискретности системной, в лингвистическом смысле.

В связи с изложенным следует подчеркнуть, что «понимание» человеческой речи машиной, конечно, не сводится только к способности сегментировать человеческую речь, опознавать слова. Если иметь в виду «понимание» связной звучащей речи, текстов, то проблема становится неизмеримо более сложной. Это уже другой этап её решения, требующий специального рассмотрения всех её аспектов.

* * *

Важнейшим этапом машинного моделирования является проблема соответствия искусственной речи своему естественному прототипу — речи человеческой. В принципе, **искусственная речь** не должна отличаться от речи человеческой; только в этом случае она будет вполне удовлетворять тем требованиям, которые к ней предъявляются.

Часто синтезированная речь сразу узнаётся как машинная. В чём же трудность? Почему синтезированная речь отличается от естественной? Что делает человеческую речь человеческой, какие её признаки? Теперь эту «человечность» речи надо физически и лингвистически обнаружить: определить, какие именно характеристики её продуцируют, запрограммировать их и искусственно воссоздать.

В самом общем плане известно, что натуральность (человечность) речи связана со свойствами речевого аппарата человека, с тем, что органы речи устроены из живой ткани: голосовые связки — мышцы, резонирующие полости покрыты слизистыми оболочками. Всё это придаёт звукам человеческой речи живые тембральные оттенки. Подобно тому,

как одна и та же нота, взятая на скрипке и на рояле, звучит по-разному, неодинаково звучит и одна и та же частота, поданная с физического прибора — генератора (искусственная частота), и возникшая от колебаний голосовых связок, получившая тембральное наполнение в сложных резонирующих полостях речевого тракта человека. Проблема заключается в том, как при анализе естественной человеческой речи физически и лингвистически вычленили характеристики, ответственные за её натуральность, с тем чтобы их запрограммировать, т.е. воссоздать.

Одна из гипотез связывает натуральность речи с нестационарной реализацией её параметров, с наличием тонкоструктурных микровариационных процессов в её сигналах, которые должны быть воспроизведены в искусственной речи. Предполагается при этом, что микровариации затрагивают все параметры речи, все её спектральные составляющие.

Спрашивается, какие задачи здесь выдвигаются на первый план? Технические, лингвистические, психолингвистические? К технике может быть, видимо, предъявлено только одно требование: она должна обладать такой разрешающей способностью, чтобы можно было задать любую микровариацию по любому параметру. Какие именно вариации задавать, каких параметров, в каких структурно-количественных соотношениях — всё это проблемы не технические, а лингвистические и психолингвистические. Собственно лингвистическими и психолингвистическими являются и задачи получения в синтезе правильной речи на данном языке, приемлемой с точки зрения её орфоэпических норм. А такая речь создаётся правильностью, точностью и уместностью речевых параметров в каждой данной позиции речевой цепи, их структурным взаимопроникновением, наложением одной характеристики на другую, соблюдением норм их варьирования и правил перехода одной характеристики в другую.

Например, представим себе какой-то звук речи, скажем, «а» или какой-либо другой. У этого звука, изолированно произнесённого как слог, — свои качества, своя физическая (формантная) структура. В пределах же слога, состоящего из трёх или четырёх элементов, этот звук уже не реализуется так, как реализовался изолированно: он приспосабливается к другим звукам, несколько меняет свою качественную и количественную структуру.

Когда же слог с этим звуком попадает в слово, состоящее из двух или нескольких слогов, звук этот может оказаться в ударном или безударном слоге, причём вхождение в слово, в ударный или безударный слог его тоже видоизменяет. Слово, в свою очередь, окажется в сочетании с другим словом, может выступить в речи как предикат какого-то сообщения при завершении или незавершении высказывания, логически подчёркнутый или нейтральный.

На основе этой лингвистической единицы может быть образован вопрос, побуждение, восклицание. Говорящий обязательно прибавит ко всему этому какую-то свою эмоцию, с которой будет произнесено это слово во фразе, какую-то модальность. И все эти метаморфозы живой человеческой речи оставляют свои отпечатки, свои отметины на отдельном исходном звуке. Человеческое ухо слышит всё это, замечая всё, что сделано не по правилам данной системы.

Вся человеческая речь и состоит из этих «тонкостей», а ведь их надо запрограммировать в физических величинах, в определённых структурных отношениях, с тем чтобы в каждой данной точке речевого пространства (континуума) получить нужное, приемлемое в данной языковой системе качество звучания.

* * *

Опыт работы по синтезу речи в Лаборатории экспериментальной фонетики Института стран Азии и Африки при МГУ на материале восточных и африканских языков свидетельствует: натуральность (естественность) речи зависит не только от наличия микровариационных процессов в её сигналах, но и от **особых взаимоотношений в структуре этих сигналов** — межформатных взаимоотношений, отношений между частотными составляющими формант, их амплитудными и временными значениями. Натуральность искусственных единиц речи повышается, когда мы удачно моделируем их просодические характеристики.

Но работа с синтезированными единицами речи требовала соответствующей аппаратной базы: нужны были **синтезаторы речи**. В 70-е годы купить заводской синтезатор было невозможно: их либо не было как таковых, либо стоили они баснословно дорого, если иметь в виду зарубежные фирмы. В такой ситуации мы разрабатывали и делали их сами. Были разработаны и изготовлены два синтезатора речи: СППИ-76 и его улучшенная модификация.

Сейчас в нашей лабораторной практике применяется формантный синтез по правилам при помощи компьютерной программы Analysis — Synthesis Laboratory (ASL). Это достаточно дорогостоящая программа, но дающая очень высокое качество синтеза по правилам. Она позволяет делать синтез речевого сигнала после анализа Linear Predictive Coding (LPC, линейное предсказание речи) и манипуляций с LPC-параметрами (графически или численно), а также менять специфические акустические параметры (ЧОТ и форманты).

* * *

Основной смысл нашего участия в конференциях, конгрессах, сессиях заключался в том, чтобы **работать на всероссийском и мировом уровне**, без каких-либо скидок на экзотику материала. Для нас всегда были предпочтительны конференции общелингвистические, которые собирали исследователей по типологически различным языкам, где обсуждались проблемы общие на материале разных языков, в том числе восточных и африканских. Мы осознанно ставили своей целью «вытащить» восточную и африканскую фонетико-фонологическую лингвистику из только своего — восточного и африканского — «соуса». Мы понимали: чем дольше китаистика будет сидеть в китаистике, а арабистика — в арабистике, тем хуже будет для них — можно задохнуться.

В этом ключе выведения исследований экзотических фонетик и фонологий на общелингвистический уровень проходила и наша деятельность по оказанию помощи другим учебным и академическим учреждениям, в том числе и зарубежным (Вьетнам, Египет, Афганистан). Имеется в виду наша работа с аспирантами и стажёрами. Через Лабораторию их прошло немало, и география здесь была и остаётся самая разная: и близкая — Московский институт языкознания, РУДН и др., и далёкая — Владивосток, Сибирь, Казахстан, Средняя Азия.

* * *

В заключение необходимо сказать **ещё об одном аспекте** работы Лаборатории, принципиально важном как для понимания облика современной лингвистики вообще, так и для понимания того, с каким человеческим контингентом мы работаем.

Дело в том, что востоковедение и африканистика до сих пор несут в себе трудно изживаемые черты комплексного, синкретического знания. Такие, например, столпы нашей востоковедной науки, как академики В.М. Алексеев, Н.И. Конрад, а в Чехословакии акад. Я. Прушек в своих работах выступали то как лингвисты, то как литературоведы, историки или этнографы. Во второй половине прошедшего века вся наука — как естественная, так и гуманитарная — характеризовалась ярко выраженной тенденцией к специализации — размежеванию разных отраслей знания.



Румянцев М.К., Капун М.И.

Из истории Лаборатории экспериментальной фонетики ИСАА при МГУ

Ныне в мире существуют такие специальности, как теоретическая физика и экспериментальная физика, существует и фонетика — общая, теоретическая и экспериментальная. Существуют — следовательно, и будут существовать — лаборатории экспериментальной фонетики.

* * *

P.S. Эту статью я начинал писать в соавторстве с Михаилом Кузьмичом Румянцевым, но 28 апреля 2010 года его не стало. В дальнейшем, работая над текстом этой статьи, я старался максимально учесть всё то, что Михаил Кузьмич думал и надеялся воплотить в этой работе.

Литература

Артёмов В.А. Коммуникативная, синтаксическая, логическая и модальная функции интонации // Материалы коллоквиума по экспериментальной фонетике и психологии речи. М.: Изд-во МГУ, 1966. С. 3.

