

ДИСКУССИЯ

ОТ РЕДАКЦИИ

Последние десятилетия отмечены определенным кризисным состоянием методов фонетических исследований. Связано это отчасти с тем, что в конце XX в. в инструментальной фонетике произошла резкая смена технической базы: аналоговые приборы (осциллографы, спектрометры, интонографы и т.п.) уступили место компьютерным программам, которые позволяют получать сведения об акустических характеристиках звуков речи в режиме реального времени.

Повсеместное распространение персональных компьютеров и их широкое внедрение в практику экспериментально-фонетических исследований, как это ни парадоксально, привело к снижению уровня метрологической культуры работ. Вера в абсолютную надежность результатов, полученных с помощью компьютера, как показала практика, оказалась свойственной не только обывателю, но и многим фонетистам.

В итоге исследователь, мало искушенный в акустике, использует фонетические программы вслепую, и выводы, сделанные на основании полученных таким образом данных, могут оказаться весьма далекими от реальности.

Непосредственным поводом для написания нижеприведенной статьи А.В. Венцова – сотрудника Института физиологии им. И.П. Павлова РАН (Санкт-Петербург) послужила целая серия работ С.В. Кодзасова, посвященных так называемой “скрытой” фонетике. В этих работах автор доказывает существование некоторой системы просодических признаков, присущих, по его мнению, каждому слову русского лексикона, которые можно выявить только с помощью акустического анализа. Редакция сочла полезной публикацию этой статьи вместе с ответной статьей С.В. Кодзасова.

ЧТО ТАКОЕ “СКРЫТАЯ ФОНЕТИКА”?

© 2003 г. А. В. Венцов

В статье анализируется метрологическая надежность экспериментальных данных, якобы говорящих о существовании в фонетике русского языка ранее неизвестных исследователям явлений (“скрытая фонетика”). Показано, что из-за использования автором гипотезы неадекватных психоакустических методов исследования восприятия и неаккуратного применения компьютерных программ при анализе речевых сигналов аргументы в пользу высказанных гипотез нельзя признать убедительными.

The aim of the paper was to test the metrological reliability of the set of experimental data, that was supposed to prove a ‘hidden phonetics’ hypothesis, the latter being an alias for some phonetic phenomena in Russian, which are proclaimed to be almost unstudied for the present. It is shown that the mentioned data have been obtained both with inadequate psychoacoustic methods and with not accurate enough use of computer programs in speech signal analysis. It makes all arguments for the ‘hidden phonetics’ doubtful.

За последние полтора десятка лет С.В. Кодзасов опубликовал множество работ, в которых постулирует существование в русском языке и речи его носителей явлений, ранее неизвестных исследователям и потому не изученных. Для определения этих явлений он вводит термин “скрытая фонетика”, а их исследование считает настолько важным, что видит в этом генеральный путь развития российской экспериментальной фонетики в XXI в. [1].

Поскольку я без малого 40 лет связан с исследованиями механизмов речеобразования и процессов восприятия речи, мне было важно понять, что собой представляет “скрытая фонетика” и

действительно ли все эти годы мы исследовали не то и не так. И прежде всего, оценить методы и технику исследования, позволившие С.В. Кодзасову обнаружить эти явления.

Основным источником данных об этих скрытых явлениях, по признанию самого автора, была интроспекция. Говоря русским языком, это – самонаблюдение, т.е. все свои выводы (заметим, достаточно безапелляционные) автор построил на собственных субъективных ощущениях. Если же рассматривать этот метод с точки зрения современной экспериментальной психологии, то это – психофизический эксперимент, в котором участвовал единственный испытуемый.

Известно, что качественные и количественные оценки явлений и объектов внешнего мира, которые делает человек (испытуемый) на основании своих субъективных ощущений, характеризуются сильной изменчивостью. Поэтому для получения сколько-нибудь надежных результатов на основании данных, полученных в психофизическом эксперименте, используются методы статистического анализа, иногда весьма изощренные. Считается, что для этого требуются оценки, полученные на группе в 15–20 человек минимум. Очевидно, что данных единственного испытуемого будет совершенно недостаточно. Тем более, что в нашем случае испытуемый и экспериментатор – одно и то же лицо и чрезвычайно велика опасность влияния самовнушения, когда наблюдатель подсознательно оценивает свои ощущения или даже артикулирует в соответствии с априорными внутренними установками (гипотезами, ожиданиями). Знаю это по собственному печальному опыту.

Ссылки на большой практический опыт автора в сфере интроспективных наблюдений вряд ли можно признать весомым аргументом, поскольку это его субъективная оценка своих же способностей и возможностей. И она может оказаться завышенной.

В психофизическом эксперименте испытуемый предстает в качестве своеобразного измерительного прибора, показаниями которого экспериментатор пользуется при изучении явлений окружающего мира, недоступных непосредственному измерению с помощью технических средств. Но даже технические измерительные приборы подвергаются периодическим поверкам, на основании которых судят о допустимости использования данных, полученных с их помощью. Поэтому с выводами автора, сделанными на основании интроспективных (психофизических) наблюдений, можно было бы согласиться только при условии, что они подтверждаются аналогичными результатами других исследователей или данными инструментального анализа.

Об одной из таких попыток непосредственно сопоставить свои субъективные оценки с результатами инструментальных измерений С.В. Кодзасов вскользь упоминает в одной из статей. Результат был отрицательный: "Однако в акустическом сигнале ожидаемых коррелятов этих фонационных изменений (речь идет о собственных просодических характеристиках слова, не связанных с фразовой интонацией. – А.В.) не обнаружено, что свидетельствует о неточности интуитивной интерпретации" [2, с. 74].

Ниже я попытаюсь показать, правда, с использованием косвенных оценок, что суждение о "неточности интуитивной интерпретации" можно

отнести не только к фонационным характеристикам речевого сигнала.

И тем не менее, на основании неточной интуитивной интерпретации, никак не подтвержденной инструментальным анализом, формулируется **закон**: "Следующий закон касается нередуцированных словоформ с характеристикой + А. В них обнаруживается жесткая просодическая мотивация разделения ударения: при отсутствии фонационного перелома на стыке слогов акцент стоит на основе, при наличии – на окончании" [2, с. 82].

По моему мнению, результаты только интроспективных наблюдений никак не могут быть приняты в качестве доказательства существования "скрытой фонетики", хотя использование интроспекции (самонаблюдения) вполне допустимо, и может оказаться весьма продуктивным, на стадии формулирования соответствующих гипотез.

К данным интроспективных наблюдений можно было бы отнести по-другому, если бы они подтверждались результатами инструментального анализа. Но степень доказательности результатов такого анализа существенно зависит от того, насколько метрологически корректно он был выполнен, для чего, на мой взгляд, должны быть соблюдены несколько совершенно необходимых условий: выбор техники и средств, обеспечивающих требуемую точность измерения исследуемых параметров, использование адекватного поставленной задаче экспериментального материала и применение четких количественных критериев при статистическом анализе результатов измерения и оценке наблюдаемых различий.

Здесь я с сожалением должен заметить, что повсеместное распространение персональных компьютеров и их широкое внедрение в практику экспериментально-фонетических исследований привело к снижению уровня метрологической культуры работ.

Прежде экспериментатор-фонетист при необходимости исследовать контур основного тона или спектральные характеристики речевого сигнала обращался к техникам-профессионалам, поскольку его знаний и практического опыта (по вполне объяснимым причинам) было недостаточно для настройки и обслуживания сложной аналоговой аппаратуры. В результате, к подавляющему большинству исследований не было претензий по части их метрологической добротности.

Ныне внедрение в исследовательскую практику программных "пакетов" породило у многих экспериментаторов безоглядную, и ни на чем не основанную, веру в абсолютную надежность результатов, полученных с помощью ЭВМ. При этом не обращается внимания на то обстоятельство, что подробности об используемых в "пакетах" алгоритмах в документации, как правило, отсутствуют, а тонкости их "настройки" примени-

тельно к особенностям анализируемых сигналов, а также возможные ограничения их применения в конкретных экспериментальных ситуациях не описываются. В результате, неискушенный в особенностях применения таких алгоритмов исследователь использует их "вслепую", и выводы, сделанные на основании полученных таким образом результатов, могут оказаться весьма далекими от реальности.

Судя по виду графиков, представленных в его статьях, для исследования интонационных характеристик речи С.В. Кодзасов использовал программу CECIL для DOS (WinCECIL или SpeechAnalyzer в более поздних версиях программы, рассчитанных на Windows 3.1 и 95). Эта программа первоначально была ориентирована на применение в полевых, экспедиционных условиях для оперативного анализа и отбора речевого материала. По этой причине в нее заложены некие усредненные количественные параметры алгоритмов анализа, видимо, в расчете на то, что при последующем лабораторном исследовании для точного и детального анализа сигналов могут быть применены другие программы.

Период основного тона в ней оценивается автокорреляционным методом. Он базируется на том обстоятельстве, что форма речевой волны оказывается похожей на двух соседних периодах основного тона. Именно степень этой похожести определяется величиной коэффициента автокорреляции: для полностью совпадающих форм он равен 1.0, для совершенно не совпадающих может стать меньше нуля. Поскольку в реальном речевом сигнале два соседних периода практически никогда не совпадают полностью по форме звуковой волны, коэффициент корреляции не становится равным единице и приходится вводить некоторое "пороговое" значение этого коэффициента, при превышении которого форма речевой волны на соседних периодах основного тона может считаться одинаковой.

Оценивание периода основного тона при использовании этого метода производится следующим образом. Выбирается некоторый отрезок речевого сигнала ("окно"), затем последовательно вычисляются коэффициенты корреляции сигнала на этом отрезке с сигналом такой же длительности, но со сдвигом во времени от исходного на некоторое число шагов дискретизации (имеется в виду, естественно, цифровая обработка сигнала). Величина сдвига, при котором этот коэффициент превышает упомянутый "порог", принимается за период основного тона.

При этом результат работы алгоритма существенно зависит от нескольких количественных его параметров, подбирая которые можно получить оптимальные результаты измерения для конкретного речевого сигнала. В первую очередь, это, конечно, ширина "окна" (если задать ее

меньше предполагаемого периода основного тона, велика вероятность, что алгоритм "зацепится" за вторую, гармонику), величина приращения сдвига между двумя "окнами" (чем она меньше, тем точнее оценивается период основного тона) и "порог" для коэффициента корреляции. Ни один из этих параметров в программе CECIL исследователь задать не может, и даже неизвестно, какие именно их значения заложены в программе ее разработчиками. Поэтому хорошо "настроить" алгоритм для адекватной работы с конкретным сигналом практически невозможно. Следствием же неточной настройки может быть нестабильная работа программы как измерителя параметров основного тона, а при неудачном выборе величины "порога" для коэффициента корреляции "измеренная" частота основного тона может на октаву отличаться от истинной. В документации к программе эти обстоятельства не отмечены и соответственно нет указаний на необходимость осторожного отношения к полученным данным, предполагающего возможность проверки результата с помощью других процедур. Исследователь же, полностью доверившийся программе, может интерпретировать резкие "скачки" в интонограммах как парадоксальные изменения в интонации, хотя они могли быть лишь следствием неточной или неправильной настройки алгоритма.

По-моему, с чем-то подобным столкнулся С.В. Кодзасов в своей работе [3, рисунки на с. 169–180]. Большинство интонограмм, приведенных на рисунках, характеризуются резкими "скачками" основного тона на гласных (в том числе – на ударных) слова *Потанов*, т.е. на тех участках сигнала, где следовало бы ожидать монотонного изменения частоты основного тона. Могу объяснить это только нестабильностью работы алгоритма оценки частоты основного тона вследствие неадекватной настройки его параметров применительно к особенностям анализируемых сигналов.

Не столь очевидна роль сбоя в работе алгоритма в сдвиге вверх контура основного тона на последнем гласном (после глухого взрывного согласного) относительно ожидаемого продолжения контура на этом участке сигнала. Но, думаю, что и здесь дело именно в неадекватной настройке, поскольку процесс управления голосовыми связками, обеспечивающий интонационные изменения, достаточно инерционен и не прерывается на участках глухих согласных, потому подобный резкий "скачок" частоты основного тона на интервале подчас меньшем 50 мс крайне маловероятен (см. рис. 1)¹.

¹ Должен извиниться за низкое качество рисунков: по понятным причинам, их пришлось копировать из рассматриваемых работ С.В. Кодзасова, что резко снижает качество копии даже при использовании ЭВМ. Чтобы хоть немного повысить "читаемость" рисунков, я позволил себе заново сделать текстовые надписи на них.

Наконец, во всех версиях программы CECIL для получения "красивых" графиков изменения основного тона предусмотрена возможность "сглаживания" контура (Fsmooth), и именно этот режим использовался в рассматриваемой статье. Алгоритм сглаживания в документации не описан, но в любом случае сглаживание может привести к искажению реального контура и значительнее всего эти искажения будут проявляться именно после глухих взрывных согласных: за взрывом обычно следует сильно укороченный период основного тона, что в терминах частоты означает резкий "скачок" вверх, который при сглаживании может "потянуть" за собой остальную часть контура.

Все это должно было бы заставить исследователя усомниться в адекватности полученных результатов и попытаться проверить их другим способом. Но, боюсь, именно эти сомнительные результаты, принятые за истинные, поскольку их выдал компьютер, привели автора к выводу об отсутствии "...конечного падения тона как в вопросах, так и в утверждениях. В качестве завершителя реплики скорее используется конечный восходящий тон как показатель готовности к продолжению диалога" [3].

К сожалению, автор не уточнил, что он понимает под "конечным падением тона" и что такое "конечный восходящий тон", а без такого определения его вывод выглядит достаточно странно, поскольку на многих интонограммах, приведенных в статье, отчетливо выделяется понижение тона на конечном гласном слова, являющееся к тому же продолжением понижения на ударном гласном, особенно если ввести коррекцию контура на предполагаемый октавный скачок, вызванный неудачной настройкой алгоритма (см. рис. 1).

Пользуясь спектральным анализатором программы CECIL, следует помнить, что при 8-разрядном АЦП (возможность работать с сигналами при большей разрядности АЦП появилась только в SpeechAnalyzer'e для Windows 95), используемом этой программой, динамический диапазон анализатора составляет чуть более 40 дБ. И то лишь при оптимальных условиях ввода речевого сигнала в ЭВМ (оцифровки), когда амплитуда вводимого сигнала в точности равна допустимому максимуму входного напряжения преобразователя. В этих условиях высокие форманты (третья и четвертая) могут оказаться частично или полностью замаскированными шумом, а количественные оценки их уровня – недостоверными.

В документации к программе CECIL режим работы спектрального анализатора не описывается и пользователь полностью лишен возможности изменять его параметры, но во многих подобных пакетах предусмотрена предобработка сигнала, когда интенсивность высокочастотных составля-

ющих спектра речевого сигнала искусственно повышается специальным фильтром с крутизной 6 дБ / окт. Это делается для того, чтобы на динамических спектрограммах отчетливее были представлены высокие частоты, существенные для анализа шумных согласных. После такой предварительной обработки высокочастотные форманты (третья, четвертая, пятая) оказываются "поднятыми" по интенсивности. Процедура эта искусственная, и в слуховой системе человека соотношения интенсивностей формант будут совсем иными. Не исключено, что анализатор CECIL "по умолчанию" работал именно в таком режиме.

Доказывая существование в русском языке редуцированных (стертых) слов, С.В. Кодзасов опирается именно на соотношение уровней высокочастотных и низкочастотных формант в спектрах входящих в эти слова гласных: "Ударные гласные стертых слов имеют более равномерное, менее дифференцированное заполнение частотного диапазона, чем гласные полных слов. Это проявляется прежде всего в том, что высокие форманты (Ф3, Ф4, Ф5) не столь сильно отличаются по интенсивности от первых двух формант ..." [4]. Но при полном отсутствии в его работе сведений о режиме анализатора велика вероятность того, что упомянутые спектральные особенности ударных гласных характеризуют скорее собственно анализатор, чем стертые / полные слова. К тому же, учитывая естественную вариативность параметров речевого сигнала, использование для количественной оценки наблюдаемых различий критерия "не столь сильно" представляется мне недопустимым и ничего не доказывающим.

Ширина формант (добротность резонаторов в речевом тракте) действительно могла бы быть показателем напряженности артикуляторов при произнесении соответствующих гласных. Беда только в том, что оценить ширину форманты на основании данных спектрального анализа речи – задача трудновыполнимая. Ширина форманты должна определяться как ширина спектральной кривой на уровне -3 дБ (иногда -10 дБ) относительно соответствующего спектрального максимума. Очевидно, что по динамической спектрограмме типа "Видимая речь" сделать это невозможно. Для этого надо получить спектральный срез. Но тогда возникает проблема, связанная со способом получения такового: если в качестве анализатора используется набор полосовых фильтров, то ширина спектральных пиков будет в значительной степени определяться шириной полосы этих фильтров. Если же используется Фурье-анализ (а именно этот алгоритм заложен в программу CECIL), то его результатом является линейчатый спектр и для оценки частоты и уровня форманты, а также для вычисления ее ширины требуется предварительная аппроксимация полученного спектра некоторой плавной функцией

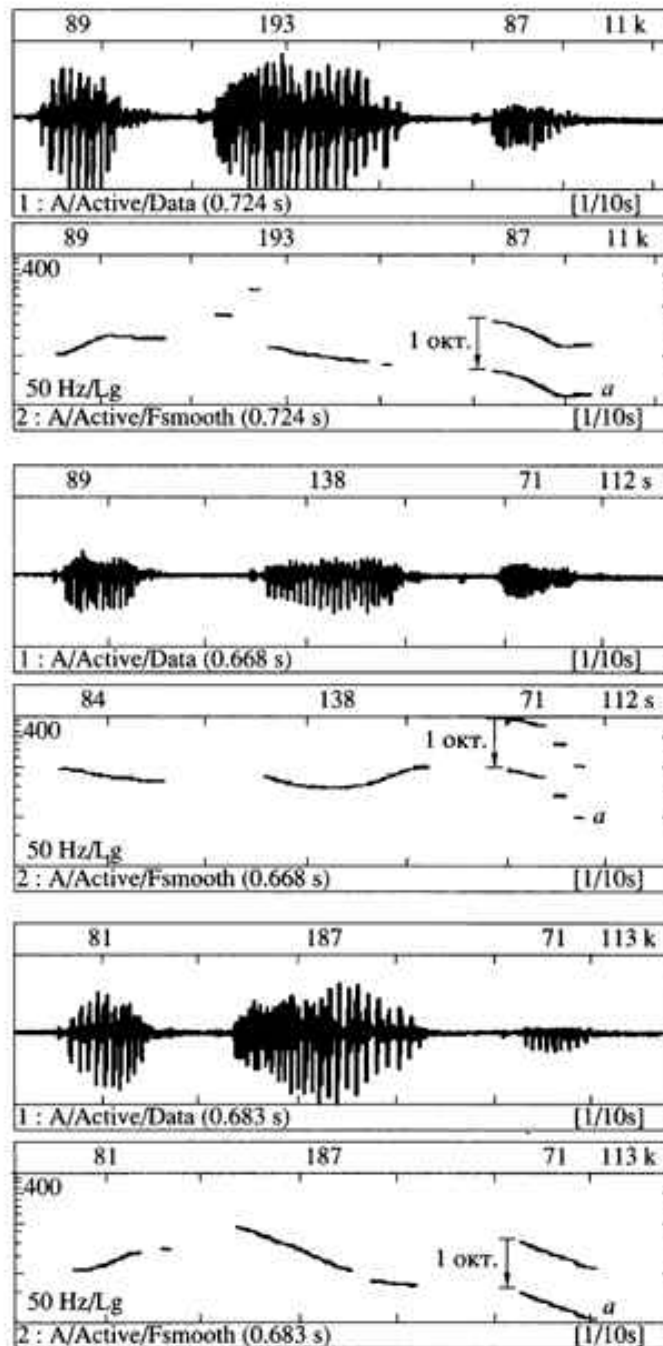


Рис. 1. Осциллограммы и интонограммы слова *Лотанов*, произнесенного в разных модально-ситуационных контекстах [3, с. 169]: *a* – положение контура основного тона после коррекции предполагаемого октавного сдвига.

(точность получаемых параметров формант в таком случае в значительной степени определяется качеством аппроксимации). Возможность управлять параметрами аппроксимации появилась только в пакете SpeechAnalyzer – доработанной версии программы CECIL, предназначенной для

32-разрядной операционной системы Windows 95. Но этим пакетом С.В. Кодзасов, судя по всему, не пользовался.

По Фанту [5] ширина первых двух формант не превышает 125 Гц. Из этого следует, что уже при

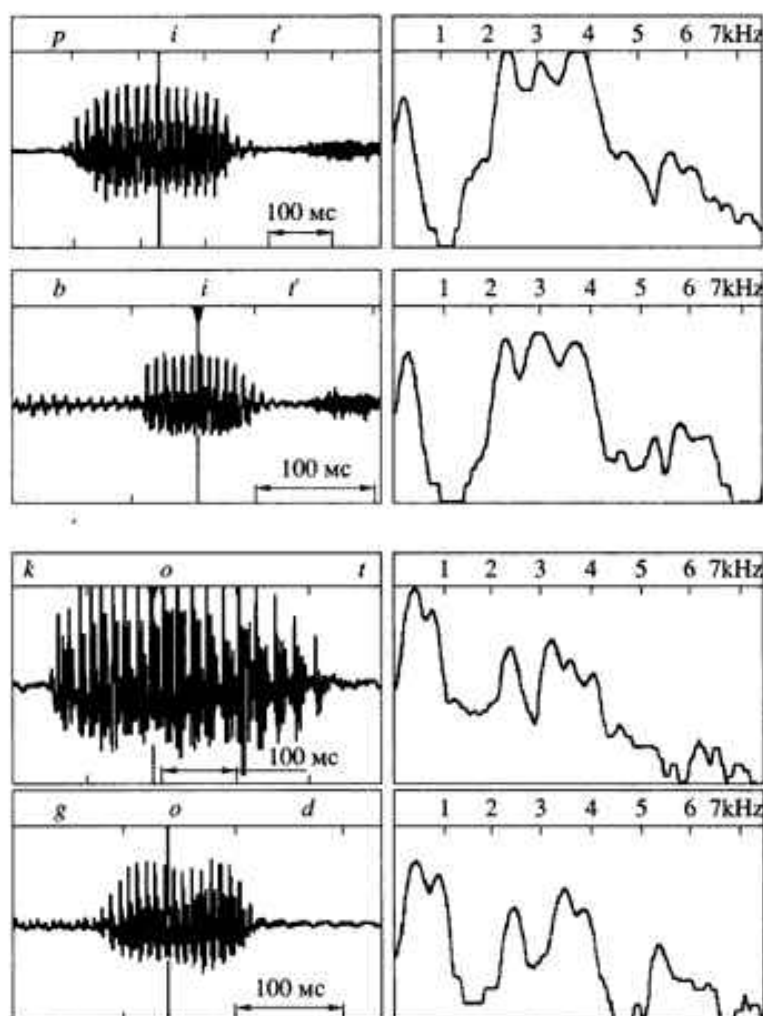


Рис. 2. Сопоставление осциллограмм и спектральных срезов на ударных гласных полных и редуцированных слов [4, с. 168]. Масштаб оси времени (100 мс) добавлен мною.

частоте основного тона большей 60–70 Гц расстояние между двумя соседними гармониками окажется больше половины ширины форманты и измерение последней окажется практически невозможным. Насколько можно судить по графикам в работах С.В. Кодзасова, подобная ситуация была типичной во всех его исследованиях.

Вероятно, измерять ширину формант было бы надежнее на спектральных срезах, полученных методом линейного предсказания, но ни в одной из версий программы CECIL подобная возможность не предусмотрена.

Из текстов рассматриваемых статей не понять, проводил ли вообще их автор измерения ширины формантных максимумов. Поэтому говорить об использовании этого показателя для оценки напряженности / ненапряженности гласных можно

только гипотетически. До практического его применения пока еще далеко.

Выше я уже говорил о сильной вариативности речевого сигнала. На практике различаться будут два последовательных произнесения одной фразы одним и тем же диктором. Тем более – разными дикторами. С этим явлением столкнулся и С.В. Кодзасов, исследуя модальные интонации [3]. Чтобы в этих условиях обнаружить какие-то закономерности, надо иметь большой объем экспериментального материала, достаточный для добротного статистического анализа результатов инструментальных измерений. Двух-трех дикторов для этого совершенно недостаточно. Поэтому вызывает удивление та уверенность, с которой автор дает развернутые интерпретации на основании столь малого числа наблюдений: "Падение

тона происходит уже на предупредном гласном, а ударный произносится на ровном нейтральном. Это один из способов показать самоочевидность события. Смещение тонального акцента на предупредную часть как бы компенсируется дополнительной морой ударного гласного. Скрипучий голос иконически выражает негативную реакцию" [3, с. 152]. Помимо всего прочего, при внимательном анализе интонограммы, из которой сделан такой вывод, читатель может самостоятельно обнаружить, что никакого падения тона на предупредном гласном нет, а то, что автор выдает за падение, есть не более, чем ошибка измерителя основного тона (рис. 2). Длительности первых двух периодов основного тона предупредного гласного, измеренные мной по осциллограмме, составили 17 и 13 мс соответственно, что соответствует примерно 60 и 77 герцам (скорректированный участок *a* интонационного контура на рис. 2). Даже если каждый из измеренных мной интервалов на самом деле соответствовал двум периодам основного тона (вполне вероятная ошибка вследствие мелкости масштаба осциллограммы), то и тогда скорректированный интонационный контур (*b* — на рис. 2) только сдвигается на октаву вверх, оставаясь тем не менее восходящим.

К этому надо добавить следующее. Говоря о просодических средствах реализации тех или иных модальных значений, автор соотносит их не с теми модальными значениями, которые на самом деле выразил диктор и воспринял слушатель, а с теми, которые диктор должен был выразить в соответствии с заданной инструкцией. Ведь аудиторской проверки успешности реализации дикторских намерений никто не проводил.

В работе "О просодии русского слова" [6] утверждается, что каждое слово русского языка может быть описано присущим только ему набором просодических признаков (словесный тон, специфический темп, раствор(?) и т.п.), не замечавшихся ранее фонетистами. Как ни странно, в статье не делается ни малейшей попытки объяснить, каким образом эти признаки включены в систему языка и какова их функция. Не заботит автора и вопрос о том, почему эти признаки так долго оставались незамеченными.

Не будучи филологом, не могу ничего сказать в ответ на первый вопрос, но кажется, смогу дать ответ на второй.

По-моему, дело в том, что фонетисты всегда стремились к получению достоверных сведений, используя для этого инструментальные средства анализа и наблюдения и соответствующие методики эксперимента. Кодзасов же опирается на результаты самонаблюдения, и, как я говорил об этом выше, никак нельзя исключить вероятность того, что сделанные им выводы являются следствием самовнушения, чем опасен используемый

им метод. Призрачность этих выводов я попытаюсь показать ниже на нескольких примерах.

Тональная просодия слова. Если допустить ее существование, придется признать, что интонационный контур любой фразы, даже когда это изолированно произнесенное слово, будет результатом одновременного действия уже не двух, а трех независимых процессов: общего понижения тона от начала к концу фразы (чисто физиологический компонент), собственный тональный контур слова и фразовая интонация. Это признает и сам С.В. Кодзасов.

Если мы попытаемся из этого суммарного контура выделить любую из трех составляющих, нам придется, говоря языком математики, решить одно уравнение с тремя неизвестными. Из курса алгебры средней школы известно, что это теоретически невозможно, если к имеющемуся уравнению не добавить еще два. В нашем случае это должны были бы быть точные количественные правила реализации фразовой интонации при произнесении какого-либо слова, учитывающие к тому же индивидуальные особенности говорящего, и уравнение линии деклинации (понижения) тона на интервале фразы. На сегодняшний день подобные правила и уравнения фонетистам не известны. С учетом сказанного остается только удивляться, как автору удалось разделить на слух то, что невозможно разделить теоретически.

В рассматриваемой статье упоминается о предварительных экспериментальных попытках решить эту задачу и говорится, что получившиеся "... словесные тоны проявляются как сравнительно небольшие модификации фразовых акцентов" [6, с. 28]. Однако в статье ни словом не упомянуты те экспериментальные процедуры, с помощью которых автору удалось решить это уравнение, а без их детального описания читатель не в состоянии проделать подобную работу самостоятельно и убедиться в правильности или ошибочности выводов автора. Я же, как читатель, склонен отнести эти "сравнительно небольшие модификации" к естественным индивидуальным вариациям контуров фразовой интонации.

Темп по Кодзасову "определяет не только скорость произнесения данного слога, но и тип перехода от гласного к последующему согласному" [6, с. 30].

Прежде всего надо отметить, что автор никак не определяет смысл термина "скорость произнесения", видимо, считая его самоочевидным. На мой взгляд, возможны две достаточно различающиеся интерпретации этого термина.

Во-первых, под "скоростью произнесения" можно понимать скорость артикуляции, т.е. скорость перемещения органов речевого тракта, участвующих в процессе речеобразования. Она зависит от усилий, развиваемых соответствующими мышечными группами, времени их действия и

инерционности конкретного артикулятора: очевидно, что при прочих равных условиях нижняя челюсть будет перемещаться с меньшей скоростью, чем губы или небная занавеска. Конечно, слова с различающимся фонемным составом будут различаться по "скорости произнесения" входящих в них слогов, поскольку в их произнесении будут участвовать разные артикуляторы. В этом случае окажется, что понятие "темпа", обычно используемое для описания явлений надсегментного уровня, т.е. действительно просодических, применено для описания процессов сегментного уровня, давным-давно известных фонетистам, более или менее изученных (см., например, [7]), но, правда, не названных одним словом. Что же тогда понимает под просодией автор?

Во-вторых, это может быть скорость (правильнее – частота) следования друг за другом в речевом потоке слогов (гласных), задаваемая темпом речи в его общепринятом употреблении. Этот параметр характеризует стиль речи (полный, разговорный), а его изменение может использоваться для выделения некоторых составляющих речевого сообщения (вводные предложения, логически выделенные слова). В этом его качестве темп является параметром надсегментным, относящимся к уровню высказывания, т.е. просодическим в общепринятом понимании. Очевидно, что одно и то же слово может быть произнесено в разном темпе, определяемом контекстом фразы или даже целого звучащего текста, и следовательно, не может быть охарактеризовано некоторым присущим данному слову значением "собственного" темпа. Каким бы странным это ни показалось, но при подобных изменениях темпа скорость движения отдельных артикуляторов может оставаться постоянной, поскольку частота следования гласных в высказывании может изменяться только за счет сокращения или prolongации стационарных участков, на которых артикуляторы находятся в состоянии относительного покоя и скорость их движения близка к нулевой.

Все сказанное выше в равной степени относится и к термину **раствор** (характеристика, по Кодзасову, связанная с перемещением нижней челюсти относительно верхней). Ясно, что это не независимый параметр, поскольку величина перемещения зависит от его скорости ("темпа" по Кодзасову) и продолжительности, и будет изменяться при изменении реального (просодического) темпа речи. И это явление тоже давным-давно обнаружено фонетистами, и проявляется, в частности, в изменении качества гласных при ускорении/замедлении темпа речи.

В заключение хочу сказать, что во всей аргументации автора отсутствует очень важный, на мой взгляд, элемент: напрочь игнорируется вопрос о том, насколько и для чего все описывае-

мые им "просодические" признаки слова используются слушателем при восприятии речи.

Еще один объект исследования для "скрытой" фонетики – редуцированные слова в русском языке [4, с. 8].

Прежде всего, что значит – "в русском языке"? Являются ли такие слова нормой литературного языка, т.е. включены ли они в его систему и в таком виде присутствуют (или должны присутствовать) в речи большинства носителей языка? В этом случае следовало бы описать их роль в языке, правила порождения таких слов и взаимоотношение редуцированных и нередуцированных слов.

И здесь меня больше всего удивляет отсутствие логики в построениях С.В. Кодзасова. Для примера – две пары фраз, в которых присутствуют полные и редуцированные слова (последние выделены курсивом):

Ваш брат прекрасно поет. – Ваш *брат* интеллигент всегда недоволен.

Пришел кот. Прошел *год*.

Очевидно, что слово *брат* в первой паре имеет разный смысл: разумеется, ни о каких родственных отношениях во второй фразе речь не идет. Возможно, слова, теряющие в некоторых контекстах свой исходный смысл, можно было бы назвать редуцированными. Не мне судить: я не лингвист. Но в чем тогда особенность второй пары? Какие законы и правила или оттенки смысла дают основание для отнесения слова *кот* к нередуцированным, а *год* – к редуцированным. Я здесь не вижу никакой логики.

Еще примеры полного отсутствия, на мой взгляд, логики. Кодзасов пишет: "признак Р (редуцированный. – А.В.) обнаруживают слова *голос, душа, мать, отец*. Эти лексемы предполагают неотчуждаемую принадлежность соответствующих свойств некоторому субъекту (*голос Вани, мать Вани* и т.п.)" [9]. Какому субъекту "неотчуждаемо" принадлежат упоминаемые свойства в выражениях *голос разума, мать-и-мачеха, душа общества, отец Сергей*?

И там же: "Характеристику ГП (гиперполный. – А.В.) получают лексемы, указывающие на ингерентную чуждость некоторого компонента или свойства данному объекту: *горб, зоб, бельмо, грыжа, дупло*". Бедные-бедные верблюды, которому *горб*, оказывается, "ингерентно" чужд, и птицы, которым также "ингерентно" чужд *зоб*!

Возможно, "в русском языке" следует понимать как "в речи носителей русского языка"? В этом случае важно уточнить, является ли редуцированное слово полноправным элементом внутреннего (ментального) словаря говорящего, на основе которого формируется артикуляторная программа. Если да, то как в словаре одно и то же

слово представляется одновременно в редуцированном и нередуцированном виде (Там находится бык. – Изгородь сломал *бык*; см. [4, с. 158])? А может быть словарь и артикуляторная программа здесь не при чем, а речь идет о редукции только на уровне акустической реализации слова? Тогда имело бы смысл говорить о **редуцируемых** словах и исследовать вероятность большей или меньшей "склонности" каких-то слов к редукции.

Так о какой редукции идет речь в статье? Имеются ли в виду любые количественные изменения акустических характеристик звуков, образующих слово? В этом случае пришлось бы признать, что в русской речи отсутствуют нередуцированные слова, поскольку в конкретных условиях речь подвергается постоянным естественным вариациям.

Или имеются в виду только те количественные изменения, которые приводят к структурным изменениям: редукция безударных гласных, выпадение одного или нескольких звуков? Очевидно, что изменения такого рода должны выходить за пределы диапазона допустимых естественных вариаций и для их характеристики требуются четкие количественные экспериментальные данные. Здесь невозможно обойтись простой констатацией того, "что высокие форманты не столь сильно отличаются по интенсивности от первых двух формант" [4, с. 158].

Автор предлагает: "Проанализируем произношение ударных гласных следующих пар фонетически близких слов в одинаковых контекстах:

Это барабан	Это сарафан
Он медный	Он бедный
Пить его не надо	Бить его не надо
Пришел кот	Прошел год
Экзамен был устный	Проход был узкий
Это лыжа	Это грыжа

В первой колонке представлены полные слова, во второй – редуцированные (здесь и далее они записаны курсивом). Ударные гласные редуцированных слов не достигают целевой артикуляции, являясь несколько более централизованными, близкими к нейтральному (шва-образному) состоянию".

Какое именно произношение предлагается проанализировать читателю, совершенно непонятно: спектрограммы этих слов в статье отсутствуют, нет и ссылок на материалы, где их можно было бы найти. Правда, для части пар слов в статье приведены спектральные срезы, выполненные на ударных гласных (рис. 1а, б, в – [4]). Но нет в ней количественной оценки ощущаемых автором изменений, а без этого нельзя утверждать, что частоты формант этих ударных гласных не попадают в характерную для этих гласных область, и гласные следует считать редуцированными. Говоря об этих иллюстрациях, мне кажется нужным добавить некоторые детали.

На рисунках приведены спектральные срезы, сделанные по утверждению автора, на стационарных участках ударных гласных. Стационарным участком гласного исследователи речи обычно называют участок, характеризующийся стационарным положением формант (максимумов в динамическом спектре). Поскольку на рисунке вместо динамических спектрограмм приведены осциллограммы, нет никакой уверенности в том, что срезы сделаны именно на стационарном участке. Но тогда различия в частотах и уровне формант, на которые автор ссылается как на доказательство редуцированности слов, могут быть простым следствием коартикуляции, если из-за неправильного выбора места спектральные срезы в сравниваемых реализациях приходились на разные участки формантного перехода от согласного к гласному.

К тому же, ни одна из сравниваемых пар сигналов, похоже, не записана в одинаковом темпе. Так, длительность гласного в "полном" слове *кот* составляет 375 мс, а в "редуцированном" слове *год* – всего 167 мс (рис. 3). Это можно определить по отметкам времени, хотя на графиках отсутствует маркировка шкалы абсцисс: в программе SECIL метки времени ставятся через 1/10 секунды. Поскольку слова совпадают по числу слогов, качеству ударного гласного и схожи по сегментному составу, единственной причиной такого различия в длительностях может быть только разный темп произнесения. Совершенно естественно, что при двукратном ускорении темпа целевое положение артикуляторов может не достигаться и гласные при этом оказываются "несколько более централизованными, близкими к нейтральному (шва-образному) состоянию".

Читатель также лишен возможности оценить, насколько упоминаемые автором различия в параметрах формант статистически достоверны, ибо никаких количественных данных ни о частотах формант, ни об их уровне, ни о точности измерения того и другого в работе нет.

К тому же, читателя вынуждают также принять на веру утверждение, что редукция одного лишь ударного гласного (допустим, что она имела место) является достаточным основанием для признания редуцированным целого слова. Почему к редукции слов не приводит редукция безударных гласных? Может быть, дело в том, что в этом случае все слова в речи носителя русского языка окажутся редуцированными, и рассуждения на тему редукции слов станут бессмысленными?

Различия в динамических спектрограммах пары слов *лыжа* – *грыжа* (рис. 2 в [4]) автор приписывает влиянию "редукции", тогда как этому существует более простое и обычное объяснение: динамика формант отражает всего лишь изменение спектра на переходе от согласного к гласно-

му. Неужто все слова, содержащие такие участки, относить к редуцированным? Тогда в эту категорию попадет большинство слов русского языка и придется исследовать феномен появления в языке **нередуцированных** слов.

Об акцентно-интонационных свойствах "редуцированных" слов судить на основании данной статьи невозможно; в ней нет ничего, кроме безапелляционных (и бездоказательных) утверждений автора. Читателя вынуждают просто **поверить** в то, что все сказанное автором – правда и основано на результатах экспериментальных исследований. Непонятно, правда, каких, кем, как и когда выполненных. Но если эти предполагаемые эксперименты были сделаны с той же степенью методической чистоты и тщательности, как и упомянутый выше спектральный анализ, то для веры в выводы автора нет никаких оснований.

Насколько я знаю литературу по экспериментальным исследованиям просодики, последняя считается явлением более высокого уровня, чем сегментный состав речи (слова), и скорее она модифицирует акустическую реализацию слов, чем наоборот. Особенно показательны в этом отношении изменения слоговых тонов (относящихся к уровню слова) под влиянием фразовой интонации в тональных языках.

Таким образом, "экспериментальный материал", описанный в статье С.В. Кодзасова не дает никаких оснований для выводов о существовании (или отсутствии) в русском языке каких-то особых редуцированных слов.

И последнее. В докладе "Фонетика интенсификации" [1] С.В. Кодзасов излагает свой взгляд на будущее экспериментальной фонетики в России. Он убежден, что помимо лексико-грамматических средств подчеркивания (усиления, интенсификации) смысла высказывания (*много-много; как прыгнет!; такой умный*) носители русского языка используют для этой цели еще и специфические фонетические средства.

Соглашаясь, что высказывания такого рода оказываются эмоционально окрашенными ("различия между интенсивами и аффеكتивами не всегда очевидны"), и понимая, что в акустическом сигнале найти признаки, относящиеся только к "аффеكتивам" или только к "интенсивам", почти невозможно, автор предполагает, что в качестве этих специфических средств могут использоваться некие интегральные артикуляционные параметры, которые, по его мнению, "иконически" (однозначно-?) связаны со смыслами: объемы полостей речевого тракта, динамика речевого потока, фонационные типы. Признаки этого типа он относит к надсегментным (просодическим) и абсолютно уверен в их существовании: "... само наличие просодических средств описанного вида отрицать совершенно невозможно" [1, с. 107].

Постулируя способность говорящего контролировать такие артикуляторные параметры в процессе речеобразования, автор считает необходимым выяснить, как эти параметры управляются в неэкспрессивной речи и, судя по тексту доклада, видит в этом основное направление развития российской экспериментальной фонетики.

Возможно, в речи некоторых носителей русского языка действительно удастся наблюдать подобные артикуляторные признаки, но чтобы это стало явлением, свойственным языку, и заслуживало внимания исследователей, необходимо, на мой взгляд, выполнение по меньшей мере двух условий (здесь я отвлекаюсь от обсуждения необходимости использовать фонетические средства усиления смысла в дополнение к имеющимся лексико-грамматическим средствам):

- подобные артикуляторные характеристики должны быть свойственны большинству говорящих на этом языке и;

- слушающий должен уметь обнаруживать эти признаки и использовать их при восприятии речи, причем даже в условиях, когда слушающий не видит говорящего, т.е. пользуясь только акустическими проявлениями этих артикуляций, если таковые имеются.

Последнее особенно важно, поскольку при недоступности этих средств усиления для слушающего они становятся бессмысленными с точки зрения говорящего: в этой ситуации говорящий использует эти средства как бы в разговоре с самим собой.

Итак, утверждается, что определенному смыслу речевого сообщения однозначно соответствует некоторая определенная конфигурация речевого тракта. Например: "Горизонтальная удаленность от говорящего, наряду с удлинением гласных (особенно ударных), кодируется смещением назад базы языка и расширением верхней глотки: – *Далеко-о-о полетел!*; *Во-о-он там!*" [1, с. 106]. Или: "заполненность объема кодируется смещением языка как целого вверх, что приводит к сужению орального прохода: *Полна коробочка! Битком набита!*". Или: "экспрессивная оценка жары маркируется подъемом гортани, напряженной фонацией, сужением верхней глотки и широким раствором ..." [1, с. 106–107].

Еще утверждается, что "для интенсификации используются также специфические динамические характеристики речи". Но напрасно читатель ждет описания этих характеристик. Вместо этого речь ведется о том, что "сочетание консонантного напряжения с напряженной фонацией иконически подчеркивает физическую твердость, а сочетание консонантной мягкости с расслабленной фонацией – мягкость объекта: – *Он такой прочный, крепкий!* vs. – *Он такой мягкий!*" [1, с. 107].

Читатель должен сам догадаться (или придумать), какими же параметрами речевого сигнала или артикуляторными позами выражаются все эти "напряженности", "мягкости" и "расслабленности". На каких именно согласных и одинаково ли на всех? Какие вокализованные участки сигнала используются для оценки качества фонации? Каковы единицы измерения этих признаков? Какая количественно должна быть разница для смежных оценок?

Подобным же образом описываются способы реализации в речевом сигнале других смыслов: вертикальное расстояние, объем, оценки погоды и т.д.

Напрашивается вопрос: какими объективными методами и на каком количестве дикторов автор наблюдал описанные изменения "интегральных" артикуляторных параметров при произнесении этих и подобных им фраз? Ссылок на какие-либо работы, подтверждающие справедливость таких безапелляционных утверждений, в тексте доклада нет. Я же осмелюсь утверждать, что никакими и ни на каком, поскольку российским исследователям никогда не была доступна необходимая для этого экспериментальная техника.

Создается впечатление, что в данном случае использовался все тот же метод интроспекции (самонаблюдения), доказательная сила которого ничтожно мала, насколько мне удалось показать это выше.

Таким образом, использование "интегральных" артикуляторных параметров для интенсификации смысла перестает быть очевидным, поскольку не только отсутствуют доказательства использования их большинством говорящих на русском языке, но недоказанным оказывается само утверждение, что определенным смыслом соответствуют специфические артикуляторные жесты или позы.

Посмотрим теперь, способен ли обнаружить описываемые "интегральные" артикуляторные параметры слушатель и, если да, то какими средствами. Для этого зададимся следующими вопросами и оценим возможные ответы на них.

– Влияет ли просодическая структура высказывания на восприятие смысла речевого сообщения?

– Да, влияет.

– Являются ли тембральные характеристики речи просодическим признаком?

– Возможно. Поскольку «отделение просодического тембра, который только и может быть предметом интонационного исследования, от варьирующих в широких пределах спектральных характеристик отдельных звуков речи представляется трудной задачей. Она осложняется необ-

ходимостью разграничивать эмоционально обусловленные особенности спектра и особенности так называемого "качества голоса" (жесткий, мягкий, придыхательный, хриплый и т.п.) – момента индивидуально и отчасти ситуативно обусловленного» [10, с. 54].

– Влияют ли размеры и форма полостей речевого тракта на спектральные характеристики речевого сигнала и, соответственно, его тембральную окраску?

– Да, безусловно.

Может быть, используя изощренные методы спектрального анализа, автору удалось обнаружить те изменения в динамическом спектре фраз, подобных приведенным выше, которые вызываются изменениями размеров и формы речевого тракта, являющимися "интегральными" артикуляторными признаками, "иконическими" связанными со смыслом? В докладе ответа на этот вопрос нет, а я уверен, что это невозможно даже теоретически, поскольку одинаковые изменения спектра могут быть следствием разных изменений в конфигурации тракта. Другими словами, можно оценить, какие изменения в спектре звуков речи вызовет, например, опускание гортани, но обнаружив подобные изменения спектра в реальном сигнале, нельзя утверждать, что они были вызваны именно опусканием гортани (три плюс два всегда равно пяти, но пять – это не обязательно три плюс два).

Кроме того, как показано выше, технические средства спектрального анализа, используемые С.В. Кодзасовым, крайне несовершенны, а методы интерпретации результатов измерения не позволяют судить о достоверности наблюдаемых различий в параметрах сигнала.

Таким образом, создается впечатление, что читателя вольно или невольно вводят в заблуждение, выдавая собственные ожидания и ощущения за научно обоснованный факт.

Как же будет в дальнейшем развиваться российская фонетика, если мы примем полученные таким путем "факты", как научно обоснованные?

Суммируя сказанное выше, на вопрос "что такое "скрытая" фонетика?", по-моему, можно ответить следующим образом: "скрытая" фонетика представляет собой: частично – описание под новыми именами явлений, известных исследователям речи носителей русского языка и в той или иной степени изученных; частично – собрание заблуждений автора, базирующихся на некорректном применении методов экспериментального исследования; частично – логически слабо аргументированные гипотезы.

СПИСОК ЛИТЕРАТУРЫ

1. Кодзасов С.В. Фонетика интенсификации. Материалы международной конференции "100 лет экспериментальной фонетике в России". СПб. 1-4 февраля 2001 г. С. 105-108.
2. Кодзасов С.В. Просодические классы слов и место ударения // Просодический строй русской речи. Глава 3. Отв. ред. Николаева Т.М. М., 1996. С. 70-84.
3. Кодзасов С.В. Исследование модальных интонаций // Просодический строй русской речи. Глава 6. Отв. ред. Николаева Т.М. М., 1996. С. 142-180.
4. Кодзасов С.В. О редуцированных словах в русском языке // Проблемы фонетики. 1995. Вып. II. С. 157-171.
5. Фант Г. Акустическая теория речеобразования. М., Наука: 1964.
6. Кодзасов С.В. О просодии русского слова // Славянское и балканское языкознание / Просодия. Под ред. Зализняка А.А. и др. М.: Наука, 1989. С. 26-40.
7. Чистович Л.А. и др. Речь. Артикуляция и восприятие. М.; Л.: Наука, 1965.
8. Кодзасов С.В. Семантико-фонетическое расщепление русских частиц и просодическая информация в словаре // Словарь. Грамматика. Текст. М., 1996. С. 97-111.
9. Кодзасов С.В. Категория свой/чужой: семантика и фонетика // Проблемы семантического анализа лексики / Тезисы докладов международной конференции "Пятое Шмелевские чтения". Москва, 23-25 февраля 2002 г.
10. Светозарова Н.Д. Интонационная система русского языка. Л., 1982.