

МЕТОДОЛОГИЯ ФОНЕТИЧЕСКИХ ИССЛЕДОВАНИЙ

А. В. Венцов

Санкт-Петербургский государственный университет

(Россия, Санкт-Петербург)

av.ventsov@gmail.com

ЭКСПЕРИМЕНТАЛЬНАЯ ФОНЕТИКА И МОДЕЛИРОВАНИЕ СИСТЕМЫ ВОСПРИЯТИЯ РЕЧИ

Цель экспериментально-фонетических исследований — получить исчерпывающее признаковое описание звуков речи, анализируя акустический сигнал. При этом фонетист заранее знает «лингвистическую» структуру анализируемого высказывания и измеряет соответствующие физические параметры сигнала. К тому же он анализирует речевой сигнал *post factum* и в случае необходимости имеет возможность многократно обращаться к анализу одного и того же отрезка.

Цель создания функциональной модели системы восприятия речи — установить набор признаков, необходимых и достаточных для адекватного понимания высказывания, и создать алгоритмы, предположительно воспроизводящие процедуры обработки речевого сигнала слуховой системой человека. Речевой сигнал анализируется слушающим он-лайн и «снизу вверх», т.е. вся информация для принятия лингвистических решений извлекается из исходного акустического сигнала, а сами решения принимаются на основании выделенных параметров и с учетом всех имеющихся у слушающего знаний о языке и текущей коммуникативной ситуации.

В статье приведены некоторые примеры, демонстрирующие способность системы восприятия при принятии «лингвистических» решений обходиться без некоторых признаков из всего многообразия собранных на сегодня экспериментальной фонетикой особенностей акустической реализации звуков речи.

Применяемые в перцептивной фонетике и психолингвистике методики экспериментов не дают возможности использовать полученные результаты при моделировании процессов восприятия речи.

Ключевые слова: экспериментальная фонетика, признаки, восприятие речи, перцептивный эксперимент, функциональное моделирование

Моделирование процессов порождения и восприятия речи следует понимать как создание лишь **функциональных моделей** этих речевых систем, т.е. алгоритмов обработки информации, обеспечивающих «преобразование» говорящим

некоторого смысла в акустический речевой сигнал и извлечение некоего смысла из речевого сигнала слушающим.

Во второй половине XX в. считалось, что когнитивная система носителя языка использует при восприятии речи те параметры речевого сигнала, которые экспериментальная фонетика полагает «ответственными» за лингвистически важные явления в этом сигнале. Поэтому перед физиологами и психологами ставилась ограниченная задача — изучить процессы выделения требуемых параметров, а последующее использование полученных данных предоставить лингвистам.

Сейчас становится очевидным, что это было глубокое заблуждение: за долгие годы исследований такой подход не привел к созданию работоспособной модели системы восприятия речи [Венцов 2015], а фонетика и моделирование процесса восприятия речи имеют дело с физиологически и функционально разными объектами исследования.

Л. В. Щерба считал, что на долю фонетиста «при исследовании того или другого конкретного языка выпадают следующие задачи: 1) выяснить фонетический состав данного языка, иначе, определить *различаемые* им фонемы; 2) в соответствии с имеющимися средствами дать их полное описание; 3) констатировать наблюдаемые дивергенции или, что почти то же, зарегистрировать *по возможности* все оттенки фонем, появляющиеся в данном языке, и дать посильное их описание; 4) определить условия этих дивергенций, и 5) объяснить причины их появления» [Щерба 1912: 20].

Поэтому, заранее «зная», как должно было бы звучать то или иное высказывание, фонетисты измеряют физические параметры соответствующего речевого сигнала, т. е. анализируют речь «сверху вниз» — от значения к форме. При этом фонетист-исследователь анализирует речевой сигнал *post factum* и при необходимости может делать это многократно.

Обычно анализируется речь небольших групп носителей языка, а результатом исследований являются, как правило, параметры речи «усредненного» представителя каждой из групп, учитывающие все нюансы акустической реализации звуков речи, связанные с их произнесением в целостном речевом потоке.

К сожалению, в литературе нет убедительных свидетельств того, что такие параметры действительно важны для надежного и правильного восприятия речевого высказывания в условиях **естественного** речевого общения, когда слушатель «обрабатывает» сигнал он-лайн и лишен возможности делать это многократно. Данные перцептивной фонетики свидетельствуют скорее об обратном.

«Сильное артикуляторно-акустическое различие между лабиализованными и нелабиализованными аллофонами согласных “бесполезно” при восприятии, ибо огубленные гласные (согласные? — А. В.) надежно отличаются от неогубленных по своим собственным характеристикам. Именно поэтому разные по огубленности аллофоны согласных в перцептивном отношении как бы одинаковые» [Бондарко и др. 1991: 92].

«Русские гораздо хуже опознают не только французские носовые, но и русские назализованные гласные (при этом случаев, когда носовое качество гласного

оценивается русскими как присутствие носового согласного, очень мало)» [Бондарко 1981: 86].

Фонетисты рассматривают оглушенные звонкие согласные и озвонченные глухие в качестве самостоятельных единиц описания речевой волны. Физически же оглушенные звонкие согласные представляют собой соответствующие парные глухие, и эксперимент с частотным словарем словоформ показал, что «омофоны», образующиеся в результате глобального оглушения звонких согласных, составляют мизерную часть словаря, предположительно используемого системой восприятия речи [Венцов 2015а]. Таким образом, выделение признака «глухость–звонкость» может оказаться при восприятии русской речи необязательным. Об этом же говорит и наша способность понимать шепотную речь, в которой звонкие звуки отсутствуют.

В естественной и особенно в спонтанной речи вместо согласного /j/ «в конце ударного или безударного слога внутри слова перед согласным, а также в конце слова, например: *бой-ко, вой-лок, вой-на, кай-ма, май, рой, пей, <...>* произносится так называемое неслоговое *и*» [Матусевич 1976: 157], а в некоторых случаях — «ноль звука» [Касаткин 2003: 151]. В описании сигнала фонетист на месте редуцированного согласного /j/ помещает соответствующие символы, и в дальнейшем такая транскрипция легко читается лингвистом, знающим соответствующие правила.

Система восприятия речи не «умеет» отличать неслоговые гласные от слоговых, а тем более — интерпретировать отсутствие звука. Однако инструментальный анализ речи свидетельствует о том, что даже при редуцированном до «ноля» согласном на соседних гласных сохраняются характерные *i*-образные переходы формант и этим обеспечивается восприятие согласного /j/ [Блохина 2013].

Далее. Существует мнение, что «русские гласные под ударением являются не только более сильными, но и более длительными, нежели *соответствующие* (выделено мною, — *А. В.*) безударные гласные. Можно даже сказать, что в русском языке все слоги под ударением являются долгими, а все безударные слоги — краткими» [Трубецкой 1960: 60].

Для фонетики достаточно простой констатации этого факта. При моделировании системы восприятия речи важно иметь правила использования длительности гласных для идентификации ударности. В экспериментальной фонетике эти правила не исследуются, поскольку это не входит в ее задачи: фонетику интересует, **что** может использоваться, но не интересует, **как**.

Например, при восприятии изолированного двусложного слова слуховой системы достаточно просто сравнить длительности гласных, если это гласные одного качества. Правда, при этом следует учитывать, что какую-то разницу длительностей человеческий слух вообще не способен обнаружить.

Однако задача усложнится применительно к словам с большим числом слогов и станет практически неразрешимой при сравнении гласных разного качества: так, гласный [а] воспринимается равноударным с соседним гласным [i], будучи в полтора раза более длительным [Апушкина 2013]. Кроме того, вариативность

длительностей гласных в естественной речи настолько велика, что «абсолютная длительность (т. е. время звучания данного гласного, измеренное в мсек) не может служить достаточным критерием для определения ударности–безударности» [Бондарко и др. 1973: 144].

При инструментальном анализе речи исследователи ищут выход, используя различные способы нормализации. Беда, однако, в том, что системе восприятия речи недоступны используемые для этого математические операции с результатами измерений.

Не следует ли предположить, что длительность гласных сама по себе не является параметром, определяющим ударность / безударность гласных (слов)?

«Если носитель русского языка слышит звучащую речь, он должен представить ее в виде последовательности фонем» [Бондарко 1977: 115], «...не подлежит сомнению, что мы отличаем слово *стол* от слова *стул* и опознаем эти слова вне всякого контекста именно потому, что первое имеет гласный /о/, а второе — гласный /у/. В этом смысле можно говорить о “словоразличительной” и “словоопознавательной” функциях фонем» [Зиндер 1979: 41].

Однако в естественном речевом сигнале, а особенно в спонтанной речи частотным оказывается гласный (шва, ер), который, не будучи русской фонемой, способен быть реализацией нескольких фонем (/а/, /у/, /е/, /ы/). Видимо, по этой причине фонетисты приходят к выводу, что «окончательная фонемная интерпретация слова происходит только на основе смысловой интерпретации целого словосочетания» [Бондарко 1979: 96].

Ситуация для лингвистики оказывается неразрешимой: лексическая идентификация невозможна без фонемного описания, но последнее возможно лишь после лексической интерпретации. Удивительно, но система восприятия речи затруднений подобного рода не испытывает!

Возможно, дело в том, что радикальной редукции в речи подвергаются только гласные, согласные же не имеют аллофонов и образуют «каркас», обеспечивающий идентификацию лексических единиц. Тогда на этапе лексического поиска система сможет обходиться без точной оценки фонемного качества гласных, ограничиваясь лишь констатацией их позиции в речевой цепочке. При этом очевидно, что при поиске в ментальном лексиконе «всплывет» сразу несколько «омофонов».

В эксперименте с частотным словарем акцентуированных словоформ [Венцов, Грудева 2008] все гласные в словоформах словаря были заменены единым символом. Была учтена палатализация согласных перед «я, е, ё, ю, и», а также их реализация в виде слога с согласным «й» в абсолютном начале словоформы и перед «ъ, ь»; были также удалены символы основного и вторичного словесного ударения.

Всего было трансформировано 122 292 словоформы. Затем из полученного «словаря» были выбраны единицы с двумя и более «омофонами»; таковых оказалось 38 815 (31,7%). Однако омофония зачастую разрешается благодаря контексту (грамматическому и семантическому). К тому же «искусственная» омофония

в ряде случаев ничем не отличается от естественной, возникающей в речи при редукации окончаний.

Из приведенных в таблице данных также следует, что при лексическом поиске в таком словаре только 5640 единиц (4,6%) требуют для надежной идентификации использования ритмической структуры словоформы.

Число омофонов	Количество омофонов (словоформ)	Ритмическая структура	
		Одинаковая	Разная
1	83 477	–	–
2	12 465 (24 930)	11 403 (22 806)	1062 (2124)
3	2860 (8580)	2411 (7233)	449 (1347)
4	948 (3792)	715 (2860)	233 (932)
5	164 (820)	44 (220)	120 (600)
6	43 (258)	7 (42)	36 (216)
7	29 (203)	2 (14)	27 (189)
8	8 (64)	0	8 (64)
9	7 (63)	0	7 (63)
10	8 (80)	0	8 (80)
11	1 (11)	0	1 (11)
14	1 (14)	0	1 (14)
Всего	–	–	(5640)

Если система восприятия речи действительно способна обходиться без установления точного фонемного качества гласных, то при функциональном моделировании это избавляет от необходимости решать две весьма серьезные проблемы: «нормализацию» дикторов (различия мужских, женских и детских голосов) и качественную редукацию гласных (даже ударных) в естественной речи.

Приведенные выше примеры свидетельствуют, что из всего многообразия акустических признаков звуков речи, установленного экспериментальной фонетикой, система восприятия речи может использовать только необходимый и достаточный набор признаков, позволяющий слушающему понимать звучащую речь, используя для этого все имеющиеся у него знания о языке.

Для функционального моделирования важно отыскать эти признаки и сформулировать алгоритмы их обработки когнитивной системой в режиме он-лайн.

Принято считать, что этим занимается перцептивная фонетика (а в последние годы — психолингвистика), поскольку «одной из наиболее общих задач перцептивной фонетики является определение того, какие звуковые характеристики существенны для воспринимающего речь человека» [Бондарко и др. 1991: 89]. Однако в любом перцептивном эксперименте поведение испытуемого определяется не только свойствами предъявляемых ему сигналов, но и всей структурой эксперимента (количеством стимулов, способом их предъявления, инструкцией и т. п.) [Венцов, Касевич 1994], и если условия эксперимента значительно отличаются от условий естественного речевого общения, результаты эксперимента невозможно использовать при создании функциональной модели [Венцов 2015]. Именно это происходит с результатами перцептивно-фонетических исследований: по существу,

их результаты говорят только о том, какие звуковые признаки слуховая система человека способна различать или идентифицировать в условиях конкретного эксперимента.

Литература

Апушкина И. Е. Собственная длительность гласных и восприятие ударения в русской спонтанной речи // Известия вузов. Сер. Гуманитарные науки. 2013. Т. 4, вып. 1. С. 34–40.

Блохина М. Э. Особенности реализации согласного /j/ в русской спонтанной речи. Выпускная квалификационная работа / СПбГУ. СПб., 2013. 50 с.

Бондарко Л. В. «Знание языка» и восприятие речевых единиц // Сенсорные системы. Л.: Наука, 1979. С. 90–103.

Бондарко Л. В. Звуковой строй современного русского языка. М.: Просвещение, 1977. 175 с.

Бондарко Л. В. Фонетическое описание языка и фонологическое описание речи. Л.: Изд-во ЛГУ, 1981. 199 с.

Бондарко Л. В., Вербицкая Л. А., Гордина М. В. Основы общей фонетики. Учебное пособие. СПб.: Изд-во СПбГУ, 1991. 152 с.

Венцов А. В. От психоакустики к психолингвистике // Биологические и социальные основы коммуникации. Сборник трудов. Вып. 1. СПб., 2015. С. 14–23. Цит. по: URL: www.infran.ru/Educ_Methodic_Department/Book-BSOK-2015.pdf

Венцов А. В. Важен ли признак «звонкости» согласных для восприятия речи? // Проблемы порождения и восприятия речи: Материалы XIII выездной школы-семинара / Отв. ред. Е. В. Грудева. Череповец: Изд-во ЧГУ, 2015а. С. 6–11.

Венцов А. В., Грудева Е. В. Частотный словарь словоформ русского языка (проект). Череповец: Изд-во ЧГУ, 2008. 204 с.

Венцов А. В., Касевич В. Б. Проблемы восприятия речи. СПб.: Изд-во СПбГУ, 1994. 232 с.

Зиндер Л. Р. Общая фонетика: Учеб. пособие. М.: Высшая школа, 1979. 312 с.

Касаткин Л. Л. Фонетика современного русского литературного языка. М.: Изд-во МГУ, 2003. 224 с.

Матусевич М. И. Современный русский язык. Фонетика. М.: Просвещение, 1976. 288 с.

Трубецкой Н. С. Основы фонологии. М.: Изд-во иностранной лит-ры, 1960. 373 с.

Щерба Л. В. Русские гласные в качественном и количественном отношении / Ред. Л. Р. Зиндер, Л. В. Бондарко. Л.: Наука, ЛО, 1983. 321 с.

Anatoliy V. Ventsov
St. Petersburg State University
(Russia, St. Petersburg)
av.ventsov@gmail.com

EXPERIMENTAL PHONETICS AND MODELING OF THE SPEECH PERCEPTION SYSTEM

The goal of experimental-phonetic research is to get an exhaustive indicative description of the sounds of speech, analysing the acoustic signal. In this case, the phonetician knows in advance the “linguistic” structure of the analysed statement and measures the corresponding physical parameters of the signal. In addition, he analyses the speech signal *post factum* and, if necessary, has the opportunity to repeatedly access the analysis of the same segment.

The purpose of creating a functional model of the speech perception system is to establish a set of features that are necessary and sufficient for an adequate understanding of the utterance, and to create algorithms presumably reproducing the procedures for processing the speech signal by the human auditory system. The speech signal is analysed by the listener on-line and “bottom-up”, i.e. all the information for making linguistic decisions is extracted from the original acoustic signal, and the decisions themselves are made on the basis of the selected parameters and taking into account the knowledge of the language and the current communicative situation available to the listener.

The article gives some examples demonstrating the ability of the perception system when making “linguistic” decisions to dispense with some features from the whole variety of the features of the acoustic realisation of speech sounds gathered today by experimental phonetics.

The methods of experiments used in perceptual phonetics and psycholinguistics do not allow us to use their results in the modelling of speech perception processes.

Key words: Experimental-phonetics, features, speech perception, perceptual experiment, functional modeling

References

Apushkina I. E. [The intrinsic duration of vowels and the perception of stress in Russian spontaneous speech]. *Izvestiia vuzov. Ser. Gumanitarnye nauki*. 2013. Vol. 4, Iss. 1. P. 34–40. (In Russ.)

Blokhina M. E. *Osobennosti realizatsii soglasnogo /j/ v russkoi spontannoi rechi*. Vypusknaya kvalifikatsionnaya rabota [Peculiarities of the consonant /j/ realization in Russian spontaneous speech. Graduation qualification work]. St. Petersburg, 2013. 50 p.

Bondarko L. V. [“Knowledge of the language” and perception of speech units]. *Sen-sornye sistemy*. Leningrad, Nauka Publ., 1979. P. 90–103. (In Russ.)

Bondarko L. V. *Zvukovoi stroi sovremennogo russkogo yazyka* [Sound system of the modern Russian language]. Moscow, Prosveshchenie Publ., 1977. 175 p. (In Russ.)

Bondarko L.V. *Foneticheskoe opisanie yazyka i fonologicheskoe opisanie rechi* [Phonetic description of language and phonological description of speech]. Leningrad, Leningrad St. Univ. Publ., 1981. 199 p. (In Russ.)

Bondarko L.V., Verbitskaya L.A., Gordina M.V. *Osnovy obshchei fonetiki. Uchebnoe posobie* [Fundamentals of general phonetics. Tutorial]. St. Petersburg, St. Petersburg St. Univ. Publ., 1991. 152 p. (In Russ.)

Kasatkin L.L. *Fonetika sovremennogo russkogo literaturnogo yazyka* [Phonetics of the modern Russian literary language]. Moscow, Moscow St. Univ. Publ., 2003. 224 p. (In Russ.)

Matusevich M.I. *Sovremennyi russkii yazyk. Fonetika* [The modern Russian language. Phonetics]. Moscow, Prosveshchenie Publ., 1976. 288 p. (In Russ.)

Shcherba L.V. *Russkie glasnye v kachestvennom i kolichestvennom otnoshenii* [Russian vowels in qualitative and quantitative terms]. Eds. L.R. Zinder, L.V. Bondarko. Leningrad, Nauka Publ., Leningr. Dept., 1983. 321 p. (In Russ.)

Trubetskoi N.S. *Osnovy fonologii* [Fundamentals of phonology]. Moscow, Foreign Literature Publ., 1960. 373 p. (In Russ.)

Ventsov A.V. [From psychoacoustics to psycholinguistics] *Biologicheskie i sotsialnye osnovy kommunikatsii. Sbornik trudov* [Biological and social bases of communication. Collection of Works]. Iss. 1. St. Petersburg, 2015. P. 14–23. (In Russ.) URL: www.infran.ru/Educ_Methodic_Department/Book-BSOK-2015.pdf

Ventsov A.V. [Is the consonant feature “voiced” necessary for the perception of speech?] *Problemy porozhdeniya i vospriyatiya rechi: Materialy XIII vyezdnoid shkoly-seminara* [Proc. of the 13th seminar “Problems of production and perception of speech”]. Ed. E.V. Grudeva. Cherepovets, Cherepovets St. Univ. Publ., 2015a. P. 6–11. (In Russ.)

Ventsov A.V., Grudeva E.V. *Chastotnyi slovar' slovoform russkogo yazyka (proekt)* [Frequency dictionary of Russian word forms (project)]. Cherepovets, Cherepovets St. Univ. Publ., 2008. 204 p. (In Russ.)

Ventsov A.V., Kasevich V.B. *Problemy vospriyatiya rechi* [Problems of speech perception]. St. Petersburg, St. Petersburg St. Univ. Publ., 1994. 232 p. (In Russ.)

Zinder L.R. *Obshchaya fonetika: Ucheb. Posobie* [General phonetics: Textbook]. Moscow, Vyssh. Shkola Publ., 1979. 312 p. (In Russ.)