

УДК 81`342.1
Doi 10.26456/vtfilol/2025.4.114

ВИЗУАЛИЗАЦИЯ ЗВУЧАЩЕЙ РЕЧИ: МЕТОДЫ И ПЕРСПЕКТИВЫ

Н.Ю. Сороколетова

Волгоградский государственный университет, г. Волгоград

На примере программы акустического анализа речи Speech Analyzer представлен функциональный потенциал визуализации речевого сигнала для решения теоретических и прикладных задач в области экспериментальной фонетики, прагма- и социолингвистики, синтеза и распознавания речи, фonoскопической экспертизы.

Ключевые слова: экспериментальная фонетика, визуализация, акустический анализ, график спектrogramмы, график основного тона, график интенсивности.

Исследование акустических параметров звучащей речи находится в фокусе внимания специалистов в области экспериментальной фонетики и смежных лингвистических и нелингвистических дисциплин (см., например: [2; 3; 4; 7; 8]). Экспериментальная фонетика развивается как самостоятельный раздел общей фонетики с середины XIX века, и ее основная цель – представить объективные данные, подтверждающие или опровергающие гипотезы, разработки, исследования, связанные с анализом звучащей речи.

Среди основных методов экспериментальной фонетики можно выделить следующие:

- 1) самонаблюдение без помощи инструментов;
- 2) эксперименты с применением приспособлений, приборов и аппаратов;
- 3) акустические методы, позволяющие частично визуализировать звучащую речь.

Метод самонаблюдения чаще всего применяется при изучении артикуляционных особенностей различных фонем, при необходимости сопоставительного анализа работы речевого аппарата при произнесении разных звуков одного языка или схожих звуков разных языков. При этом экспериментатор может наблюдать как за речью других людей, так и за своей собственной речью.

Эксперименты с применением приспособлений, приборов, аппаратов получили наибольшее распространение в конце XIX–нач. XX века, в связи с новыми научно-техническими достижениями в различных сферах человеческой деятельности. Как правило, данные методики направлены на изучение и фиксацию артикуляционных особенностей произнесения отдельных звуков или их сочетания. Среди наиболее

значимых для экспериментальной фонетики методик можно выделить следующие: ларингоскопический метод (наблюдение за положением органов речи при помощи ларингоскопа); метод палатограмм – отпечатков, фиксирующих области контакта языка с твердым небом; метод фото- и киносъемки (цель – описание особенностей положения наружных и внутренних органов артикуляции при произнесении отдельных звуков или их сочетаний); метод рентгенографии и рентгенокиносъемки позволяет получить комплексное изображение артикуляционной картины звука, как в статике, так и в динамике.

Существенным недостатком представленных выше методов является создание определенных помех, нарушающих естественную артикуляцию при произнесении звуков, а также потенциальная опасность для здоровья участников эксперимента.

Звуки речи можно слышать, но их нельзя видеть. Для того, чтобы исследовать акустические характеристики речевого сигнала, необходимо вначале преобразовать звуковые колебания в электрические. На их основе делается спектральное представление речевого сигнала и производится его осциллографическая запись, которая дает возможность визуализировать звуковой сигнал. Акустические методы изучения речи позволяют:

- фиксировать отрезки речи практически любой длины, хранить их;
- получать визуализацию звуковых отрезков;
- получать цифровые показатели акустических параметров;
- сравнивать результаты перцептивного эксперимента с объективными акустическими характеристиками и др.

Одной из первых подобных методик во второй половине XX века была кимографическая методика. Кимограф – специальное устройство для регистрации движений органов речи, позволяющее фиксировать (в виде линий различной формы) понижение и повышение давления воздуха, работу носового резонатора (определять степень назализации), работу полости рта, работу гортани (дрожание голосовых связок, интенсивность дрожания), создавать линию отсчета времени в виде регулярной волнистой черты и т.д.

Благодаря созданию осциллографа стало возможным разработка метода осциллографии. Осцилограмма речевого сигнала представляет собой график, который показывает, как изменяется амплитуда звукового давления во времени при произнесении звука.

Метод спектрографии получил широкое применение в середине XX века. В основе этого метода лежит применение прибора спектографа, который обеспечивает зрительное представление звука. Получаемая на нем спектрограмма представляет все три акустических измерения звука: время представлено на одной оси двумерного изображения, частота – на другой, а интенсивность звука по каждой частоте и в каждый момент времени представлена затемнением [1].

Современные методы исследования звучащей речи наглядно демонстрируют возможности информационных технологий для проведения акустических исследований; описанные выше процедуры осуществляются с помощью специальных программ, которые позволяют проводить высококачественный анализ речи. Комплексный акустический анализ речевых характеристик позволяет делать выводы не только о лингвистических особенностях говорящего, но и о его социокультурном и эмоциональном статусе (месте проживания, уровне формирования речевых навыков, поле, возрасте, эмоциональном и психическом состоянии, уровне образования и т.д.). Данные акустического анализа востребованы в таких стремительно развивающихся сферах, как распознавание и синтез речи, искусственный интеллект. Все это делает подобное программное обеспечение инструментом исследования в области социофонетики, психофонетики, лингвопрагматики, информационных технологий, фоноскопической экспертизы, лингводидактики и некоторых других.

Speech Analyzer (SA) – компьютерная программа, позволяющая проводить акустический анализ звучащей речи. Благодаря свободному доступу, широким возможностям и достаточно интуитивному интерфейсу SA (наряду с программой Praat) является одной из самых популярных и востребованных программ, обеспечивающих высококачественный анализ звучащей речи [9]. Функционал SA включает визуализацию звучащей речи в виде графиков, гибкую настройку системы, многоуровневую фонетическую разметку и многое другое.

Проведение экспериментально-фонетического исследования состоит из нескольких этапов, основными из которых являются аудиторский (перцептивный или слуховой) и акустический анализ. Верификация данных, полученных на этапе перцептивного анализа, является частью акустического анализа и может происходить с помощью визуализации звукового сигнала, которая реализуется в виде графиков акустических параметров (Graph), созданных в программе. Исследователь выбирает те инструменты, которые релевантны для целей проводимого фонетического эксперимента. Рассмотрим исследовательский потенциал некоторых функций программы SA для решения теоретических и прикладных задач.

Как правило, на акустическом уровне анализу подвергаются цифровые параметры частоты основного тона (ЧОТ), интенсивности, длительности, спектра звуков. Важным инструментом для анализа и манипулирования звуковыми сигналами является спектрограмма. График спектрограммы представляет собой изображение спектра звуков, частот звука с их амплитудами (см. рис. 1).



Рис. 1. Пример графика спектрограммы

Этот график является облигаторным методом исследования формантных характеристик ударных и безударных гласных, модификации звуков в потоке речи, перцептивных характеристик фонем и др. На основе спектрограммы возможна дифференциация эмоционального состояния человека, выявление и последующая корректировка различных патологий речевого поведения, составление речевого портрета говорящего. К последнему относится выявление не только гендерной принадлежности говорящего, но и ряда других социально-демографических характеристик, которые используются для идентификации говорящего, например, в рамках фоноскопической экспертизы. Так, ряд спектральных показателей, в частности, показатель сниженной скорости речи как признак пожилого возраста, позволяет делать предположения о возрасте говорящего, а также синтезировать типичный и даже стереотипный эталон каждого возраста на основании фонетической информации, систематизированной в результате обобщения больших массивов фонетических данных, собранных у людей одного возраста [5: 27].

Как правило, частота основного тона, интенсивность и скорость речи являются основными голосовыми свойствами при анализе интонационных характеристик просодических единиц, например, для распознавания эмоций, психологических характеристик личности, выявления прагматической направленности говорящего, его истинных интенций [6]. Построение специальных графиков (график частоты и интенсивности) для выявления этих акустических параметров также является частью функционала программы SA.

Кривая, отражающая изменение основного тона во времени при произнесении речевого отрезка, называется контуром частоты основного тона (ЧОТ) или интонограммой. В программе SA контур ЧОТ соответствует графику Pitch.

По мнению Л.П. Блохиной и Р.К. Потаповой, частота основного тона является одним из самых универсальных супрасегментных

параметров, который принимает участие в передаче почти всех видов просодической и интонационной информации [1: 4].

Частота основного тона – это частота вибрации голосовых связок. Данный акустический параметр измеряется в герцах (Гц). У каждого говорящего базовая ЧОТ индивидуальна и обусловлена особенностями строения гортани. Именно график высоты тона (Auto Pitch) (см. рис. 2) является фундаментальным методом для определения основных параметров, связанных с описанием мелодического контура фрагмента звучащей речи: среднего уровня ЧОТ, частотного диапазона, частотного интервала и некоторых других.

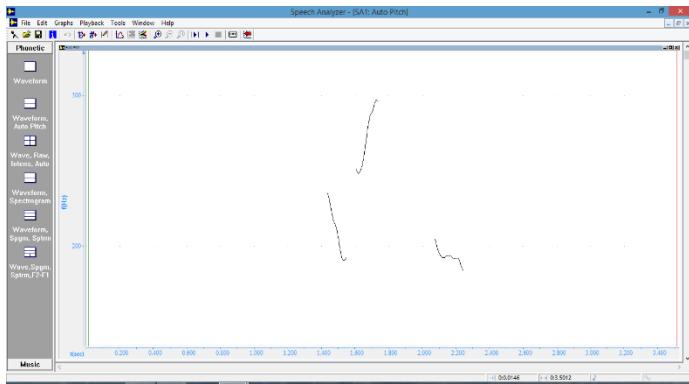


Рис. 2. График высоты тона

В акустической фонетике существуют различные способы определения среднего частотного уровня. Однако чаще всего средний уровень ЧОТ выявляется по безударным слогам и равен соотношению суммы значений ЧОТ безударных слогов в пределах анализируемого фрагмента к количеству безударных слогов в этом фрагменте. Программа SA определяет его автоматически.

Средняя ЧОТ исследуемого фрагмента может быть использована в качестве самостоятельного параметра, либо выступать в роли нормирующего параметра при оценке изменений частотного контура. В этом случае данный показатель выполняет вспомогательную роль [1: 17].

Акустические параметры интенсивности (и/или их изменения), характерные для речевого поведения личности или отдельной социальной группы, также зависят от ряда социолингвистических и экстралингвистических факторов и находятся в фокусе внимания исследований по акустической фонетике и смежным дисциплинам. Анализ графика интенсивности (Intensity) (отдельно или во взаимодействии с графиками осциллограммы и/или графиком Pitch) позволяет выявлять следующие акустические параметры: средняя интенсивность, мин/макс интенсивность, диапазон и интервал интенсивности и некоторые другие (рис. 3).

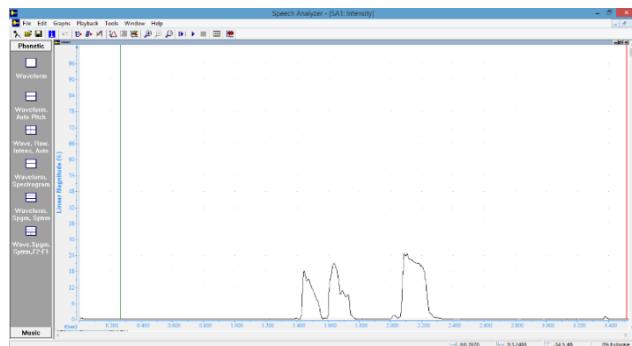


Рис. 3 График интенсивности

Для более детального анализа полученных визуализаций в программе SA предусмотрена ряд дополнительных функциональных возможностей: прослушивание всего файла или его фрагмента, запись звучащей речи (с возможностью регулирования уровня громкости записи, уровня громкости проигрывания, частоты дискретизации, фильтра верхних частот и некоторых других технических характеристики), приближение графика (например, для определения более точных границ исследуемых отрезков) и т.д.

Кроме графиков, которые позволяют визуализировать речь и проводить комплексное экспериментально-фонетическое исследование, функционал программы SA позволяет более наглядно представлять полученные цифровые данные, верифицировать результаты исследования. Рассмотрим некоторые из этих возможностей.

При сравнительно-сопоставительном анализе фрагментов звучащей речи возможно многоуровневое изображение двух и более графиков основных акустических параметров (команда *Tile Horizontal/Vertical*). Данные графики могут наглядно демонстрировать совпадения или расхождения в графических рисунках исследуемых характеристик. На рисунке 4 представлен пример графиков ЧОТ дикторов, произносящих идентичный вербальный материал в различных коммуникативно-прагматических контекстах.

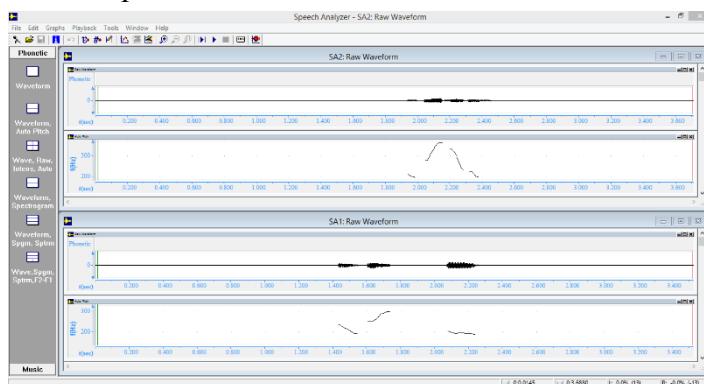


Рис. 4. Многоуровневое изображение графиков осциллограммы и ЧОТ

В ряде исследований, например, в области социофонетики, психо- и прагмафонетики определяющими являются не столько цифровые показатели акустических параметров, сколько их корреляция с социальными характеристиками говорящего. В программе Speech Analyzer предусмотрена полезная функция, позволяющая вносить и сохранять информацию о дикторе / участнике эксперимента. Для каждого файла в меню *Information* (диалоговое окно File Information), хранится информация о технических особенностях данного файла (акустические характеристики). В соответствующие графы можно добавить информацию о дикторе (имя, пол, место рождения, особенности произношения и т.д.) и комментарии к файлу (его содержание, описание и т.д.) (рис. 5).

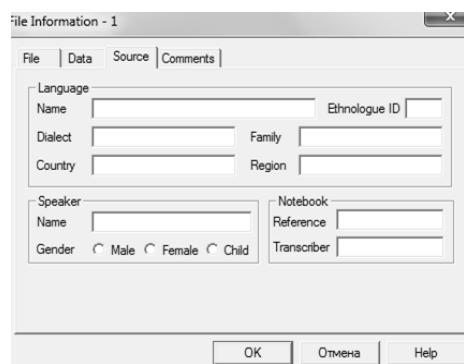


Рис. 5. Диалоговое окно File information

Программа Speech Analyzer также позволяет сохранять снимки графиков и сами графики в отдельных файлах, копировать и вставлять их в любые текстовые редакторы (File→Screen Capture→Save; Window→Copy Window→Save Graphs→Copy Graphs). Также есть возможность копировать результаты акустических измерений и сохранять их в виде таблиц в другие программы (Open→Graphs→Types; Edit→Copy Measurements) (см. Таб. 1).

Таблица 1

Пример сохранения результатов акустического анализа в виде таблицы

File Name	Speaker	Gender	Start (sec)	Duration (sec)	f0 (Hz)	f0 (st)	Intensity (dB)
BBC	American	Male	0.4719	0.418	148.0	50.1	22.29

Полученные рисунки графиков и таблиц могут являться наглядным материалом, верифицирующим результаты перцептивного анализа.

Еще одной особенностью программы Speech Analyzer является возможность наложения (Overlay) одной визуализации звукового отрезка поверх другой (рис. 6). Данная функция, безусловно, очень полезна как для

изучающих иностранный язык (можно сравнить визуализацию своего произнесения звука/слова/фразы с визуализацией носителя языка), так и для экспериментатора-фонетиста (например, при исследовании особенностей и причин акцента в речи и проблем интерференции).

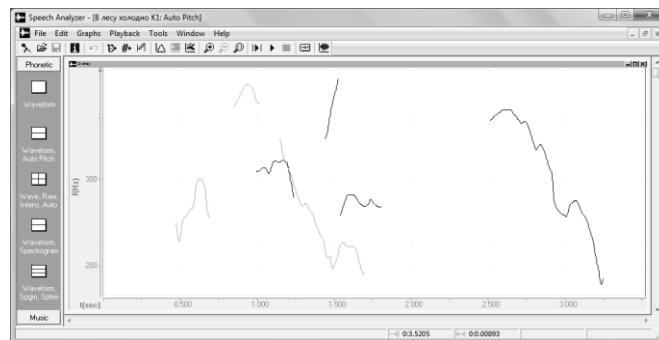


Рис. 6. Пример наложения графиков высоты основного тона двух дикторов

Неоспоримым достоинством функции наложения программы SA является автоматическое отсечение тишины перед началом фразы (это происходит, например, если участник эксперимента делает незапланированную паузу перед произнесением фразы). Это серьезно облегчает работу экспериментатора.

Таким образом, визуализация звучащей речи является единственным инструментом для решения широкого круга теоретических и прикладных задач в различных областях современных исследований, связанных с анализом звучащей речи. Функциональные возможности современных программ акустического анализа позволяют не только визуализировать речь (строить мгновенные графики осциллограмм, спектрограмм, интенсивности, частоты основного тона), но и получать цифровые значения акустических параметров, модифицировать и адаптировать их в соответствии с целями конкретного эксперимента. Все это сочетается с возможностями системы фиксировать полученные данные, оформлять репрезентативный материал, проводить сравнительно-сопоставительный анализ сегментных и сверхсегментных единиц звучащей речи. Результаты акустического анализа, полученные с помощью таким программ, могут иметь прикладное значение не только для фонетического исследования, но и исследований, проводимых на стыке фонетики и других лингвистических и нелингвистических дисциплин.

Список литературы

1. Блохина Л.П., Потапова Р.К. Методика анализа просодических характеристик речи. М.: МГИИЯ, 1982. 75 с.
2. Грачев А.М. Лингвистические подходы к автоматическому распознаванию речи // Вестник ННГУ. 2013. № 6 (2). С. 61–63.
3. Ежкова Н.А., Елецкая О.В. Использование спектрального анализа при оценке звукопроизношения у дошкольников с речевой патологией // Концепт. URL: <https://cyberleninka.ru/article/n/ispolzovanie-spektralnogo->

- analiza-pri-otsenke-zvukoproiznosheniya-u-doshkolnikov-s-rechevoy-patologiey (дата обращения: 03.08.2025).
4. Каганов А.Ш. Криминалистическая идентификация личности по голосу и звучащей речи как интегральное научно-экспертное исследование // Филологические науки. Вопросы теории и практики. 2019. №6. С. 246–250.
 5. Плотникова С.Н. Лекции по судебной фонетике. Иркутск: ИГЛУ, 2006. 58 с.
 6. Рафикова А.С., Валуева Е.А., Панфилова А.С. Голос и психологические свойства человека: обзор современных исследований // Психология. Журнал ВШЭ. 2022. №1. С. 195–250.
 7. Соломенник А.И. Технология синтеза речи: история и методология исследований // Вестник Московского университета. Серия 9. Филология. 2013. №6. URL: <https://cyberleninka.ru/article/n/tehnologiya-sinteza-rechi-istoriya-i-metodologiya-issledovaniy> (дата обращения: 06.07.2025).
 8. Сороколетова Н.Ю. Вариативность мелодических паттернов обращения в англоязычном интервью // Вестник Волгоградского государственного университета. Серия 2, Языкоznание. 2018. Т. 17, № 3. С. 70–79.
 9. Сороколетова Н.Ю. Функциональные возможности компьютерного приложения PRAAT // Иностранные языки: лингвистические и методические аспекты 2014. № 27. С. 127–131.

Об авторе:

СОРОКОЛЕТОВА Наталья Юрьевна – канд. филол. наук, доцент кафедры теории и практики перевода и лингвистики, Волгоградский государственный университет (400062, г. Волгоград, пр. Университетский, 100); e-mail: sorokoletovanat@volsu.ru

SPEECH VISUALIZATION: METHODS AND PROSPECTS

N.Yu. Sorokoletova

Volgograd State University, Volgograd

The functional potential of speech visualization for solving theoretical and applied problems in the field of experimental phonetics, pragma- and sociolinguistics, speech synthesis and recognition, phonoscopic expertise is revealed. The review is based on Speech Analyzer as one of the tools of acoustic Speech Analysis Software.

Keywords: *experimental phonetics, visualization, acoustic analysis, spectrogram, pitch graph, intensity graph.*

About the author:

SOROKOLETOVA Natalya Yurievna – Ph.D. in Philology, Associate Professor, Department of Theory and Practice of Translation and Linguistics, Volgograd State University (400062, Volgograd, pr. Universitetsky, 100); e-mail: sorokoletovanat@volsu.ru

Статья поступила в редакцию 15.09.25
Подписана в печать 13.10.25

© Сороколетова Н.Ю., 2025