



О возможности определения эмоционального состояния говорящего по речи

В.И. Галунов,

доктор биологических наук, профессор

Список авторов, занимавшихся выражением эмоционального состояния в речи, очень велик. В качестве примера можно привести классическое высказывание Ч. Дарвина (1896 г.): «При страхе — голос дрожащий и охрипший от сухости рта. При ярости парализованные губы отказываются повиноваться, и голос обрывается или становится глухим, резким, нестройным. При страдании или жалобе почти всегда говорят высоким голосом. Любовь, соперничество, торжество сопровождаются музыкальными звуками».

Интересно указать на ещё одну принципиальную позицию, высказанную в начале 20 века У. Джеймсом. Он считал, что только с появлением речи связано настоящее проявление эмоций. У животных эмоции заменяет физиологическая, поведенческая реакция. Из более или менее недавних работ по психофизиологии эмоций можно рекомендовать R. Cornelius «The Science of Emotions». N-Y., 1996.

Работы по изучению эмоциональной речи велись в НИИ «Дальняя связь» с 1968 по 1990 гг. Последний этап их выполнения связан с реализацией в 1981–1989 гг. Всесоюзной программы по развитию речевых технологий для спеццелей. Изучение проблемы «Речь и эмоции» до 1990 г. продвигалось довольно активно. В 1978, 1981, 1988 гг. трижды проводились специализированные Всесоюзные конференции по указанной тематике. Кроме того, тема «речь и эмоции» регулярно поднималась на Всесоюзной конференции АРСО, проходившей не реже одного раза в 2 года. Все материалы этих конференций опубликованы. Специализированная конференция за рубежом прошла только в 1990 году прошлого столетия.

Подробный отчёт о старых работах по проблеме можно найти в докторской диссертации Галунова В.И. «Исследование вариативности речевого поведения человека», 1975 г. (следует отметить, что принципиально новых результатов за последние десятилетия практически не появилось).

Количественная составляющая эмоциональных состояний

При определении количества эмоций, отражаемых в поведении человека (прежде всего в речи) существует два подхода. Первый из них отталкивается именно от характеристик речи, которые могут наблюдаться объективно. Обширный материал был опубликован в довольно старой монографии Оствальда:

Печаль — медленная, вялая мелодия, вздохи.

Счастье — оживлённая мелодия.

Страх — прерывистая, нисходящая мелодия.

Гнев — резкая артикуляция.

Наиболее методически тщательно выполнено исследование классика речевых исследований К. Стивенса:

Гнев — увеличение частоты основного тона, диапазона основного тона, частоты первой форманты, темпа.

Страх — частота основного тона несколько выше нормы, пики на мелодической кривой, темп ещё медленнее, чем в норме.

Печаль — частота основного тона ниже чем в норме, малый диапазон изменения основного тона, увеличение длительности звука, увеличение пауз.

Видно, что объективный анализ не обнаруживает больших отклонений от точки зрения Ч. Дарвина. Последние работы также не выявляют резкого отличия от перечня эмоциональных состояний. Например, в работах 2003 г. это — неудовольствие, удивление, радость, страх, гнев, печаль.

Второй возможный подход к анализу списка эмоций основывается на позициях У. Джеймса: только речь организует истинные эмоции. Анализ русского слова на основании проведённых в НИИ «Дальняя связь» (Ленинград) экспериментов дал число эмоций (слов, выражающих эмоции) в количестве 500, организованных в 22 гнезда. Следует сразу сказать, что русский язык очень эмоционален — он содержит около 40% эмоционально окрашенных слов (см. работы А. Вежицкой). В других языках этот процент заметно меньше. Перевод слов (эмоций) часто не даёт прямого соответствия. Это различие было замечено более 100 лет назад выдающимся психологом В. Вундтом. Вместо термина «эмоция» он использует понятие «чувственный тон». Им была предположена концепция, согласно которой все слова, связанные с эмоциями, являются точками в некотором психологическом пространстве:

Хороший — плохой.

Активный — пассивный.

Слабый — сильный.

Современные исследования показали, что существует ещё одна ось, которая соответствует хорошо определяемому состоянию в исследованиях эмоций: Удивление, т.е. неожиданное изменение ситуации (см. Галунов В.И. Психологический журнал. №6, 1980 г.).

Эмоциональные размерности появляются не сразу, а развиваются из более простой структуры (например, из простого возбуждения. Затем происходит расщепление, например, *distress* — *delight*, а затем *fear* — *anger*. Соответственно развитие эмоций, аффекта идёт параллельно развитию интеллекта).

Следует иметь в виду, что в каждом языке слова, соответствующие разным эмоциям, могут смещаться в указанном пространстве. Так, например, русскому «тоска» практически нет аналогов в других языках.

Недостаточно ясным продолжает оставаться следующий вопрос. Модель Вундта предполагает, что пространство эмоции является евклидовым. Однако есть предположения, что в нём бифуркации. Хорошим примером может служить поведение собаки, которая злобно и агрессивно приближается к противнику, а потом неожиданно в страхе убегает. Такого поведения не может быть в евклидовом пространстве. Аналогично у человека возможен переход от смеха к слезам. Фактически данная проблема совершенно не исследована, хотя в процессе выполнения работ исследователи НИИ ДС сталкивались с возможными вариантами бифуркации.

В проводимых работах использовалась несколько расширенная модель Вундта с учётом особенностей русского языка и специфики русской культуры:

1. Тоска — радость.
2. Спокойствие — страх.
3. Отвращение — томление.
4. Смущение — исступление.
5. Радость — гнев.

Выбор модели эмоциональных состояний

Исходя из целей и условий проведения эксперимента к модели эмоциональных состояний предъявляются следующие требования, согласно которым она должна:

1. Быть легко воспроизводима в смысле возможности получения достаточно большого количества звукозаписей на достаточно большом количестве испытуемых.
2. Позволять получать от испытуемых речевые реакции, необходимые экспериментатору.
3. Позволять получать эмоции широкого спектра.
4. Допускать текущую оценку (или контроль) состояния информанта.

Практически все возможные варианты методик были опробованы в работах специалистов НИИ ДС:

- *Актёрская имитация эмоциональных состояний.*

Здесь сразу же следует указать на некоторые недостатки, которые не позволяют применить её в качестве единственной (или хотя бы основной) модели. Прежде всего следует отметить то обстоятельство, что любые, даже самые отрицательные, состояния при актёрском воплощении реализуются на положительно эмоциональном фоне актёрского вдохновения. Кроме того, актёр не столько переживает эмоцию, сколько имитирует некоторые проявления её в обычной жизни или использует определённый штамп. Начинающие актёры действительно переживают эмоции и их работа с преподавателем близка к гипнозу (см. далее).

- *Эмоциональные состояния в условиях психиатрической клиники.*

Наиболее доступной базой получения «естественных» эмоций является именно психиатрическая клиника, которая даёт богатый материал для наблюдения чётко проявленных и устойчивых эмоций. При этом в ряде случа-

ев удаётся от одного и того же испытуемого (пациента) получить несколько состояний. Очевидным недостатком рассматриваемой модели является то, что всегда имеется трудно устранимое побочное влияние медикаментозных средств. Последнее особенно касается снятия у больных отрицательных эмоций.

- *Получение реальных состояний.*

Реальные состояния были получены двумя способами. Первый состоял в записи речи лётчиков, проходивших пробное (наземное) катапультирование. Второй — запись речи детей 4–12 лет на 1–3-й день госпитализации и через 1–2 недели после госпитализации.

- *Воспроизведение эмоций в гипнозе.*

Эта модель оказалась наиболее приемлемой. Она хорошо контролируется по электрофизиологическим характеристикам. Трудность представлял отбор достаточно гипнабельных и не страдающих побочными психическими заболеваниями людей. (И конечно, нельзя устраивать массовые сеансы, подобные сеансам Кашпировского, что, конечно, замедляет получение результатов эксперимента.)

Объективные характеристики эмоциональной речи

Исходя из положений ныне принятой теории речеобразования, можно предположить, каким образом могут проявляться объективные признаки эмоций в акустическом сигнале.

Первое, что можно отметить, — это нарушение нормального контроля за процессом речеобразования, что обнаруживается при сдвиге эмоционального состояния. Это особенно касается контроля за просодическими характеристиками, управляемыми правым полушарием головного мозга.

Изменение объёмов артикуляторного тракта и изменение упругости стенок (см., например, уже указанный сдвиг 1-й форманты, Stevens R.). Однако следует указать, что заметных и не вызывающих сомнения работ в последнее время не появлялось.

В последнее время был опубликован ряд исследований по механизмам голосового источника (см., например, монографию I.R. Titz «Vocal Fold Physiology» San Diego, 1993, в которой предложены новые подходы и которая повлекла за собой целую серию работ, указавших, в свою очередь, на возможные механизмы определения эмоций по речи).

Возможно, классические воззрения на процессы речеобразования придётся пересматривать (В.И. Галунов «Актуальные проблемы речевой акустики», XI сессия Акустического общества; Галунов В.И., Уваров В.К. «Ещё раз о механизмах речеобразования», XI сессия Акустического общества.) Следует обратить внимание на некоторые эффекты, обнаруженные уже Титцом (в частности бифуркация основного тона) и некоторые данные, обнаруженные при речеобразовании птиц, что заставляет задуматься над модуляционными механизмами, участвующими в речеобразовании.

Попытаемся кратко перечислить уже достаточно хорошо исследованные признаки речевого акустического сигнала, связанные с эмоциональным состоянием говорящего:

- *Громкость речевого сигнала.*

Обычно этот признак связан со стеничностью эмоции. Различия же между состояниями в пределах одной степени стеничности несущественны.

Со степенью стеничности связана так же форма импульса подсвязочного давления: чем более позитивный характер носит состояние говорящего, тем ближе форма импульсов к треугольной. Это явление может быть объяснено влиянием двух факторов: во-первых, увеличением упругости мышц голосовых складок при возрастании положительного настроения переживаемой эмоции и, во-вторых, возрастанием подсвязочного давления, что также приближает форму импульсов голосового источника к треугольной.

- Изменение интегрального спектра.

В литературе имеются сведения о таком изменении. В исследованиях НИИ ДС подобный результат не был подтвержден. Не были подтверждены и данные по сдвигу формант. При этом необходимо учитывать сложность проблемы их выделения.

- Изменение частоты основного тона.

Этот признак отчётливо выделяется при использовании во всех методиках. Экспериментальная база: психиатрическая клиника, гипноз, имитация. При этом заметное повышение основного тона происходит по шкалам:

1. Депрессия, тоска, печаль.
2. Норма.
3. Радость, страх, ярость.
4. Изменение выраженности мелодической кривой.

5. Обычно используются два параметра: среднеквадратичный разброс основного тона на фразе или величина изрезанности мелодического контура (сумма по модулям изменений его частоты на протяжении фиксированного текста. Для эмоциональных состояний, степень стеничности которых может варьироваться, не изменяя качественной характеристики состояния, наблюдается сближение характеристик изрезанности мелодического контура различных по качеству эмоций, но близких по уровню стеничности. При этом наблюдается слияние по с нормой активной печали и пассивной радости. Страх, близкий к оцепенению, хотя на слух и индицируется как страх, но по степени изрезанности близок к норме. Следует также указать, что этот параметр не работает для фраз с простой синтаксической структурой. Вероятно, в этом случае сказывается определённый автоматизм.

6. Изменение темпа речи является хорошим индикатором отклонения от нормы.
7. Джиттер и шиммер (т.е. модуляция основного тона по частоте и интенсивности) дают также хорошие результаты.

Восприятие эмоционального состояния говорящего

Исследование механизмов восприятия эмоционального восприятия говорящего представляет интерес по двум причинам — *предоставление* дополнительных я к исследованию тех признаков акустического речевого сигнала, которые могут лечь в основу объективного анализа этого сигнала, а также их *использование* при анализе эмоционального состояния говорящего.

Работ в указанном направлении существует довольно много. В основном они направлены на прямую оценку эмоционального состояния. Достаточно хорошо оцениваются следующие состояния:

Снижение настроения. — Апатия. — Тоска. — Тревога. — Злоба. —
Страх. — Радость.

Очевидно, что лучшую оценку дают врачи-психиатры. Но и неподготовленные аудиторы достаточно хорошо справляются с задачей. Ошибки в определении эмоции носят естественный характер: радость — оживление; печаль — тоска. Незначительные эмоции часто принимаются за норму. Часто путаются между собой стенические состояния: раздражительность — оживление; гневливость — радость.

Очень хорошие результаты по определению эмоционального состояния по речи даёт метод семантически противоположных пар. В лингвистике он известен как метод шкал Осгуда (Ch. Osgood «Focus on Meaning», 1978, Paris). На самом деле методика была применена гораздо раньше (в 50-х годах прошлого столетия) в США для обучения акустиков-подводников, которая заключается в следующем. Аудитору предлагается оценить предъявляемые звуки по ряду признаков, имеющих в словесном выражении противоположное значение:

хороший — плохой, тихий — шумный, и т.д.

Практически всегда испытуемые вне зависимости от опыта дают согласованные оценки. Характер оценок носит отчётливо оценочно-эмоциональный характер. Осгуд, который пытался с помощью подобного метода получить для слов единое семантическое пространство, получил для слов не семантическое значение, а так называемый коннотат, т.е. их эмоциональное значение.

В НИИ ДС был проведён ряд исследований, который показал их перспективность. Было выделено 4-мерное пространство уже указанных вундтовских признаков.

Перспективы построения системы оценки эмоционального состояния говорящего

Что предстоит сделать:

- Создать хорошую базу данных. При этом следует учитывать, что выражение эмоций носит строго индивидуальный характер, поэтому необходимо иметь запись нормы. Небольшие базы данных, используемые за рубежом (порядка до 10 дикторов), явно непригодны.
- Создать единую процедуру распознавания эмоций по речевому сигналу, применяя современный аппарат анализа речевых сигналов. Большинство современных исследователей используют явно чужие методики, почерпнутые из Интернета.
- Попытаться развить современную теорию речеобразования (прежде всего на основе исследования механизмов работы гортани).
- Создать современную экспертную (партнёрскую) систему, возможно на основе метода семантически-противоположных пар. Возможным перспективным путём может стать исполь-

зование не собственно акустического сигнала, а его видеограммы. Таким путём шёл V. Zue (МТТ) и Н. Зиновьева (МГУ, теперь в США). Это позволяет более подробно и неспеша знакомиться с речевым материалом.

Под конец — несколько слов об одном «подводном камне», с которым сталкиваются зарубежные исследователи, пытаясь создать универсальные (для нескольких языков) системы определения двух составляющих, выражающих эмоции (прямо по У. Джеймсу):

- интенциональную, т.е. намерения говорящего выразить свое эмоциональное состояние;
- непроизвольную, выражающую неконтролируемые изменения психофизиологического состояния.

Галунов Валерий Иванович (1938 – 2005)

*Выдающийся ученый в области речевых технологий,
доктор биологических наук,
профессор Санкт-Петербургского государственного университета
Опыт работы:*

1961 — 1991 — НИИ «Дальняя связь».

Инженер, старший инженер, начальник лаборатории, начальник сектора, начальник отдела.

Участие и руководство НИР/ОКР, посвященных проблемам автоматического распознавания и синтеза речи, кодирования речи, определение эмоционального состояния человека по речи и т.п.

1968 — 1989 — научный руководитель Всесоюзной программы «Автоматическое распознавание и синтез речи».

1968 — 2000 — Российская академия наук

с 1965-го года — ученый секретарь

с 1976 года — председатель Секции «Автоматическое распознавание и синтез речи».

1991 — 2005 — ООО «Одитек». Научный директор.

1965 — 2005 — Санкт-Петербургский государственный университет.

Профессор кафедр: математической лингвистики, высшей нервной деятельности, психофизиологии.

Публикации

Публикации: около 200 научных публикаций, 5 монографий.

Образование и получение ученых степеней:

1960 — Ленинградский государственный университет.

1968 — защита кандидатской диссертации.

1975 — защита докторской диссертации.