

Значимые признаки диагностирования недостоверной информации в речи

Патракова Ева Павловна

НИУ «Высшая школа экономики», Нижний Новгород

Образовательная программа «Современные филологические практики: поэтика, интерпретация, комментарий», магистратура, 2 курс, eva.patrakova@list.ru

Микулинский Артём Денисович

НИУ «Высшая школа экономики», Нижний Новгород

Аспирант, 2-й год обучения, amikulinskii@hse.ru

Ключевые слова: подготовленная и неподготовленная речь, мультимодальность, лингвистическая модель, недостоверная информация в речи, экспериментальная лингвистика.

*Выражаем благодарность научно-учебной группе «КомпЛинг» и отдельно Хоменко Анне Юрьевне за возможность объединиться в поиске новых истин.

Ложь является неотъемлемой частью коммуникации и используется не только в обыденной речи, но и в рамках судебного и досудебного дознания. Для её распознавания, в том числе в ряде криминалистических процедур, были разработаны инструментальные методы, недостаток которых заключался в указании этих признаков и на другие изменения в психоэмоциональном состоянии человека. Именно поэтому многие модели учитывают и лингвистические особенности звучащей речи, в том числе её невербальные характеристики (жесты, позы, мимика и др.). Одним из первых лингвистический метод при определении лжи начал использовать А. Р. Лурия, оказавший большое влияние на лингвистическую криминологию и одним из первых начавший применять лингвистический метод определения лжи в речи [Лурия 1998]. На данный момент существует большое количество методик, направленных на определение недостоверной информации, но отсутствует единая модель, включающая в себя признаки лжи на разных уровнях.

Задача данного исследования – создать модель, включающую пул разнородных параметров и позволяющую определить, является ли произносимая субъектом информация достоверной или недостоверной. Для проверки созданной модели применялся экспериментальный метод, во время которого были записаны устные подготовленные и неподготовленные высказывания 20 человек, включающие как достоверную, так и ложную информацию. Выборка разделена по полу – по 10 человек для мужского и женского – и возрасту (выбраны респонденты от 18 до 75 лет). Длительность записанного материала составила 232 минуты. Распределение речевых отрезков по другим признакам представлено в таблице 1.

Таблица 1. Распределение высказываний в выборке

Типы речи	Подготовленная речь		Неподготовленная речь	
	Достоверн.	Недостоверн.	Достоверн.	Недостоверн.
Муж.	10	10	10	10
Жен.	10	10	10	10

Материал проанализирован на предмет наличия 32 признаков недостоверной информации, взятых из следующих методик (в скобках указаны признаки лжи, включенные в модели, и их примеры):

1. методика В. Ф. Енгальчева (оценка вербальных и невербальных признаков недостоверной информации, например, малое количество деталей, наличие неопределённых местоимений, стабильный темп речи) [Енгальчев 2016];
2. методика Е. И. Галяшиной (вербальные и невербальные признаки, такие как высокая частотность дословных повторов, отсутствие пауз хезитации) [Галяшина 2003];
3. исследование Н. М. Романовой (невербальные признаки лжи, например, увеличение высоты и скорости речи) [Романова 2008];
4. работа О. В. Лыковой (невербальные характеристики, например, понижение беглости речи, большая величина частоты основного тона) [Лыкова 2016].

Поскольку приведенные выше методы не противоречат друг другу и подтверждаются экспериментально, становится возможной их аккумуляция в рамках единой модели. Веса параметров в представленной модели установлены как равные.

Эксперимент проводился в три этапа. Для подготовки к первой и второй части эксперимента респонденты ознакомились с темой и пунктами плана монолога. Затем они были проинформированы о необходимости говорить неправду и правду во время записи первой и второй частей, а также о том, что после рассказа им будут заданы уточняющие вопросы. Таким образом были собраны спонтанные ответы с недостоверной и достоверной информацией. В третьей части эксперимента было предложено рассказать о любом событии – или о произошедшем с участниками, или о выдуманном, – чтобы исключить предвзятость при обработке данных. Во время проведения эксперимента был поднят ряд проблемных вопросов. Во-первых, заметным ограничением являлось участие только одного экспериментатора, не позволившее ранее осуществить качественную проверку модели. Во-вторых, в исследовании применялся полуанкетный формат взаимодействия. Для последующих исследований принято решение его изменить. Так, игровой формат способствует большей мотивации респондентов и образованию более естественной среды.

Для определения уровня значимости признаков использован ансамблевый алгоритм Random Forest.

В результате тестирования были сделаны следующие выводы:

1. Модель, учитывающая разноразмерные вербальные и невербальные маркеры достоверности/недостоверности информации, показывает надёжность при обнаружении лжи на тренировочной выборке (до 92%), что указывает на оправданность комплексного подхода к решению задачи.

2. Ансамблевый метод также подтверждает значимость комплексного подхода, указывая, однако, что наибольший вклад в решение задачи вносят такие признаки, как лексика с семантикой движения, движение лицевых мышц и количество деталей в рассказе.

Перспективой исследования можно считать пополнение исследовательского материала и создание автоматизированной системы обработки речи с функционалом для поиска признаков недостоверной информации.

Список литературы

1. Галяшина 2003 — Галяшина Е. И. Основы судебного речеведения. М.: СТЭНСИ, 2003.

2. Енгальчев 2016 — Енгальчев В. Ф., Кравцова Г. К., Холопова Е. Н. Судебная психологическая экспертиза по выявлению признаков достоверности/недостоверности информации, сообщаемой участниками уголовного судопроизводства (по видеозаписям следственных действий и оперативно-разыскных мероприятий). М.: Юрлитинформ, 2016.
3. Лурия 1998 — Лурия А. Р. Язык и сознание. Ростов-на-Дону: Феникс, 1998.
4. Лыкова 2016 — Лыкова О. В. Акустические признаки наличия лжи в высказываниях // Вестник Московского государственного лингвистического университета. Гуманитарные науки. 2016. №15. С. 144–152.
5. Романова 2008 — Романова Н. М., Усанов Д. А. и др. Диагностика ложности сообщений по параметрам невербального поведения человека: коллективная монография. Саратов: Изд-во Саратов. ун-та. 2008.