

ЛИНГВИСТИКА

Т.Н. Синеокова¹

*Нижегородский государственный лингвистический университет им.
Н.А. Добролюбова*

ОСНОВНЫЕ ПРИЧИНЫ ОШИБОК ЭКСПЕРТОВ ПРИ ИДЕНТИФИКАЦИИ КЛАССИФИКАЦИОННЫХ ПРИЗНАКОВ

В статье рассматривается роль однозначности признаков объекта исследования как инструмента проверки качества лингвистической классификации. Особое внимание уделяется анализу причин и типов ошибок экспертов при идентификации признаков, выявление которых позволяет разработать типологию правок, вносимых в классификационное описание с целью его оптимизации в нужном направлении.

Ключевые слова: классификация, классификационный признак, однозначность, эксперт, ошибка в идентификации.

T.N. Sineokova

Nizhny Novgorod State Linguistics University named after N.A. Dobrolubov

MAIN CAUSES OF EXPERTS' ERRORS IN IDENTIFICATION OF CLASSIFICATION FEATURES

Abstract of the article. The article discusses the role of the classified objects features unambiguity for testing the quality of linguistic classifications. Particular attention is paid to the analysis of causes and types of experts' identification errors which makes it possible to work out the typology of desirable corrections in the classification description in order to optimize it in the right direction.

Key words: classification, classification feature, unambiguity, expert, identification error.

При изучении объекта действительности основной целью исследователя является создание классификации его элементов, т.е. деление предметов, явлений, объектов на виды, классы или типы на основе их общих признаков с образованием определённой системы классов данной совокупности, фиксирующей закономерные связи между ними [Ермак, 1999]. Основными требованиями к выделяемой системе признаков являются ее *полнота, распознаваемость (ясность) и прогностическая полезность*.

Полная система признаков позволяет определить принадлежность любого элемента к объекту исследования и однозначно выделить те элементы, которые разработчик рассматривает как эквивалентные (см. об отношениях принадлежности и эквивалентности элементов в [Синеокова, 2009, с. 10-11]).

Распознаваемость (ясность) признака предполагает «знание его *смысла* и его *денотации*, т.е. того, что он обозначает. Если смысл признака определён отчётливо и

¹ Синеокова Татьяна Николаевна, доктор филологических наук, профессор кафедры английского языка переводческого факультета Нижегородского государственного лингвистического университета им. Н.А. Добролюбова.

однозначно, признак называют *содержательно ясным* или просто *ясным*; если обозначаемое признаком чётко очерчено, говорят, что признак *точен*» [Ермак, 1999]. В.Д. Ермак подчеркивает также, что «к сожалению, очень часто в процессе классификации к характеристикам признаков относятся небрежно <...> удовлетворяются “рыхлыми”, семантически расплывчатыми признаками, использование которых в процессе классификации (типирования, диагностики, типопределения и т.п.) приводит к грубейшим ошибкам» (там же).

Распознаваемость признаков связана с имманентным свойством любой классификации – ее *однозначностью* [Синеокова, 2010]. Однозначность признаков предполагает наличие таких правил (инструкций) их идентификации, руководствуясь которыми любой наблюдатель (эксперт) мог бы правильно идентифицировать их наличие или отсутствие в любом элементе объекта исследования. Свойство распознаваемости признака трудно отделить от свойства однозначности классификации, однако однозначная классификация всегда предполагает хорошую распознаваемость, а хорошо распознаваемые признаки всегда ведут к высокой однозначности.

Наконец, *прогностическая полезность* (*прогностический потенциал, информативность*) классификационных признаков означает, что на их основе (т.е. на основе явно наблюдаемых свойств) можно прогнозировать свойства ненаблюдаемые. В противном случае нет оснований считать, что классификация действительно отображает глубинные особенности исследуемого объекта. Отметим также, что классификация, обладающая хорошим прогностическим потенциалом относительно какого-либо важного аспекта и правильно отражающая существенные особенности исследуемого объекта, скорее всего, окажется полезной и при прогнозировании других аспектов [Бейли, 1970].

Рассмотрим для иллюстрации хорошо понятный бытовой пример. Составители многочисленных справочников по съедобным и несъедобным грибам стремятся к тому, чтобы каждый читатель нашел в них полный набор признаков, по которому любой элемент описываемого объекта (гриб) был отнесен к классу грибов в целом и к конкретному виду грибов в частности; чтобы описания признаков, на основе которых проводится объединение элементов объекта в эквивалентные группы, были ясными и однозначными; чтобы на основании этих признаков (цвет, размер, место произрастания и т.п.) любой человек мог правильно идентифицировать съедобные и несъедобные грибы.

Заметим, однако, что, несмотря на кажущуюся ясность вопроса, данные понятия очень сложно определить по отдельности; они работают в комплексе, подобно понятиям *масса, сила, ускорение и др.* в законах Ньютона. Само «содержание» полноты, ясности, прогностической ценности варьирует в зависимости от границ объекта изучения и ставящихся идентификационных задач.

В данной статье рассматривается лишь один, достаточно узкий (но важный для каждого исследователя) вопрос – вопрос проверки созданного разработчиком классификационного описания признаков на однозначность, точнее – систематизация основных причин ошибок экспертов при идентификации признаков с целью последующей корректировки инструкций.

Проблемы, связанным с трудностями вербализации представлений разработчика об объекте в виде инструкций, рассчитанных на «среднего эксперта», хорошо известны всем, кто занимается теорией классификации, в том числе лингвистам [Фрумкина, 2003; Болдырев, 2006; Беляева, 2009]. Дело в том, что исходные представления разработчика о классификационных признаках постепенно трансформируются в базовые для его сознания категории, для которых характерны гештальтность восприятия и более легкая воспроизводимость. Соответственно, определяя наличие или отсутствие какого-либо классификационного признака в некоем элементе объекта исследования, разработчик оперирует целым комплексом представлений и интуитивных знаний о признаке, то есть воспринимает признаковое описание как гештальт, как нечто, неразложимое на части и

несводимое к сумме частей [Фрумкина, 2003]. Скорее, разработчик работает по принципу «распознавания образа» [Синеокова, 2010, с. 191].

Именно поэтому однозначное, понятное любому пользователю вербальное описание признаков, т.е. составление инструкции по их идентификации, является достаточно сложным процессом, требующим особых процедур тестирования классификации, количественной оценки ее однозначности и, наконец, корректировки с этих позиций.

В статье делается попытка типологизировать причины ошибок экспертов, допущенных ими в ходе тестирования классификации синтаксических признаков речи, реализуемой в состоянии эмоционального напряжения. Данная классификация разрабатывалась с целью выявления корреляционных связей между типами психологического состояния говорящего, определяющими особенности речемыслительной деятельности, и их структурными экспонентами. Прогностический потенциал данной классификации относительно выбранного экстралингвистического коррелята, подтвержденный статистически, позволил сделать вывод о том, что она отражает глубинные закономерности речеобразования и может быть использована как инструмент исследования других форм деформированной речи и выявления широкого спектра корреляционных связей с другими экстралингвистическими аспектами. Так, например, был подтвержден прогностический потенциал данной классификации при совместном прогнозе психологического состояния на основе специфики вербального и невербального поведения [Синеокова, 2008], относительно пола говорящего [Лисенкова, 2007]. Классификация применялась в исследованиях псевдоаффективной (имитационной, воздействующей) речи (политической [Лаврова, 2010] и судебной [Антонова, 2010]), при оценке адекватности переводов художественных текстов [Добровольская, Синеокова, 2010], а также при решении ряда других задач. По-видимому, спектр применения классификации может быть расширен и далее.

Как было сказано выше, решение задач, предполагающих применение данной классификации другими исследователями, связано с обеспечением *однозначного* выделения (идентификации) классификационных признаков любым исследователем (наблюдателем, экспертом).

То, что классификационное описание (инструкция по идентификации признаков) может быть недостаточно однозначной, с самого начала было понятно его разработчику – автору данной статьи. Безусловно, не последнюю роль здесь сыграла основанная на интуиции и опыте и отмечаемая Р.М. Фрумкиной гештальтность восприятия. Кроме того, ряд ошибок идентификации мог быть спровоцирован особенностями объекта и задачами, которые решал автор.

В основе изучения объекта исследования лежало представление о деформированном (относительно ядерного предложения) характере синтаксических форм, реализуемых в состоянии эмоционального напряжения; данные деформированные формы являются информационными элементами, позволяющими обнаружить корреляционные связи между структурной спецификой и типом психологического состояния говорящего [Синеокова, 2003]. Хотя основной целью работы являлось установление корреляционных связей между психологическими состояниями и их синтаксическими экспонентами, логика работы потребовала рассмотрения всех возможных отклонений от нейтральной области языка. С одной стороны, появилась сложная иерархическая классификация признаков деформированной речи. С другой стороны, относительно узкая направленность исследования привела к тому, что не все описания признаков были представлены достаточно полно: признакам, играющим определяющую роль при идентификации психологического состояния, уделялось больше внимания.

Кроме того, в качестве основной цели исследования была декларирована демонстрации самого факта наличия корреляционных связей деформированной речи с особыми психологическими состояниями – измененными состояниями сознания (которые на тот момент не были очевидны), а вовсе не полное выявление этих связей. Поэтому усилия,

направленные на обработку инструкций идентификации классификационных признаков, тоже были ограничены. Способность разработчика классификации в разное время одинаково идентифицировать признак(и) одного и того же высказывания рассматривалась как гарантия потенциальной возможности достичь однозначной классификации также и разными экспертами, но специальные усилия для реализации этой потенциальной возможности на том этапе работы не прилагались.

Впрочем, было проведено пробное тестирование экспертов, которое позволило получить, с одной стороны, достаточно удовлетворительные результаты, но и обнаружило ряд связанных с однозначностью проблем.

Решение этих проблем стало целью специальных исследований, связанных с тестированием классификационных описаний. Сама процедура тестирования (методика работы с экспертами, описание параметров оценки качества признаков описаний и т.д.) рассмотрены в [Синеокова, Лаврова, Чернышова, 2006; Беляева, 2009]. В ходе тестирования удалось выявить три типа ошибок (пропуск признака – ПП, ложное обнаружение признака – ЛО, замещение признака – ЗП), количественно оценить «размеры» ошибок и т.д. Но для оптимизации инструкции необходимо было понять, почему конкретно в том или ином случае возникают сбои.

Причины ошибок экспертов были условно разделены на две группы: психологические и лингвистические.

Не останавливаясь на таких хорошо известных психологических факторах, влияющих на решение экспертов, как обстановка, самочувствие и т.п. [Сулейманова, 2010], назовем еще две причины, которые, как представляется, можно отнести к психологическим. Поскольку в ходе тестирования экспертам предлагался ряд высказываний, в которых они на основе инструкций должны были выделить признаки, можно было предположить, что *частотность реализации признаков* и *число входящих в структуру высказывания признаков* (длина структуры высказывания) также являются психологическими факторами, затрудняющими запоминание и распознавание признаков.

Частотность реализации признака определяется отношением числа реализаций признака к общему числу высказываний в тестируемой базе данных. В соответствии с частотностью реализации классификационные признаки были разделены на следующие группы: частотные признаки (условно с частотностью $F > 0,1$, т.е. признак реализуется в 10% и более высказываниях опорной БД), признаки со средней частотностью ($0,02 < F < 0,1$) и нечастотные признаки ($F < 0,01$).

Исходя из психологических особенностей восприятия частотных и нечастотных признаков, было необходимо проверить: а) не обнаруживают ли частотные признаки ошибки типа «ложное обнаружение», связанные с формированием у экспертов устойчивых ожиданий их обнаружения в большинстве анализируемых высказываний; б) не характерны ли для нечастотных признаков ошибки типа «пропуск признака», связанные с тем, что эксперты с трудом идентифицируют редко встречающийся тип деформации структуры высказывания.

Результаты тестирования показали, частотность признака не влияет на качество его идентификации. Несмотря на то, что частотность различных деформаций структуры ядерного предложения варьируется, причины их ошибочной идентификации следует искать в лингвистической плоскости.

По *длине структуры* все высказывания, содержащиеся в тестируемой базе данных, можно разделить на высказывания с короткой уникальной структурой (1-2 признака) и высказывания с длинной уникальной структурой (3 и более признаков).

Необходимо было проверить, наблюдаются ли ошибки типа «ложное обнаружение» и «пропуск признака», связанные с психологической тенденцией к «усреднению» длины уникальной структуры: поскольку уникальные структуры, состоящие из двух-трех признаков, встречаются чаще, эксперты могли пропускать признаки в более длинных уникальных структурах (особенно состоящих из более, чем 4-х признаков) и ложно обнаруживать «дополнительные» признаки в коротких уникальных структурах.

Как показала проверка, длина структуры высказывания также не влияет на идентификацию признаков. Типичные ошибки идентификации отдельных признаков наблюдаются как в длинных, так и в коротких структурах.

Таким образом, основными причинами, обуславливающими ошибки идентификации классификационных признаков, следует признать лингвистические причины, к которым были отнесены: *новизна признака*, *«яркость» признака*, *сходство признаков* и *неполнота описания признаков*. Данные причины были отнесены к группе лингвистических причин ошибок идентификации, поскольку они непосредственно связаны с качеством признаковых описаний.

Новизна признака в тестируемой классификации – наиболее типичная причина ошибок. Эксперт может пропустить никогда ранее не встречавшийся признак (допустить ошибку типа ПП) или заменить на знакомый или легко узнаваемый признак или сочетание знакомых и/или легко узнаваемых признаков (ошибка типа ЗП). Заметим также, что в качестве причины ошибки в данном случае может рассматриваться не сам признак, а иное («старое») осмысление термина(ов), используемых при его описании [Зимова, 2011].

Яркость признака как причина ошибок идентификации предполагает следующее. Ряд классификационных признаков могут обладать легко запоминающимися и потому узнаваемыми структурными и/или функциональными особенностями, что может провоцировать их ложное обнаружение экспертами на основании общих формальных маркеров (ошибка типа ЛО). Кроме того, «яркие» признаки могут «затенять» в восприятии экспертов другие классификационные признаки, в связи с чем эксперты могут пропустить «неяркие» признаки в высказывании при правильной идентификации «ярких» (ошибка типа ПП).

Что касается такой причины ошибок, как *сходство признаков*, отдельные признаки рассматриваемой структурно-семантической классификации действительно характеризуются сходными, пересекающимися структурными и/или функциональными и семантическими особенностями, наличие которых может вызывать их замещение экспертами при идентификации (ошибка типа ЗП).

Неполнота описания признака означает, что в инструкции содержится недостаточно сведений и признак будет пропущен экспертами (ошибка типа ПП).

Таким образом, выявление причин ошибок идентификации признаков и связанных с ними типов ошибок позволяет разработать типологию правок, вносимых в инструкцию, и оптимизировать признаковые описания в нужном направлении.

СПИСОК ЛИТЕРАТУРЫ

Антонова, Е. С. Особенности функционирования повторов в судебной речи: Дипломная работа / Е. С. Антонова. – Нижний Новгород, 2010. – 65 с.

Бейли, Н. Математика в медицине и биологии / Н. Бейли. – М.: Мир, 1970. – 326 с.

Беляева, Е. И. Способы регистрации изменений восприятия при разработке однозначного классификационного описания // Теоретические и прикладные аспекты изучения речевой деятельности: Сборник научных статей. Вып. 4. – Н. Новгород: Нижегородский государственный лингвистический университет им. Н.А. Добролюбова, 2009. – С. 30-35.

Болдырев, Н. Н. Языковые категории как формат знания / Н. Н. Болдырев // Вопросы когнитивной лингвистики. – 2006. – № 2. – С. 5-22.

Добровольская, Л. В., Синеокова, Т. Н. Особенности перевода конструкций с повторами, реализуемыми в состоянии эмоционального напряжения (на материале русского и английского языков) // Теоретические и прикладные аспекты изучения речевой деятельности: Сборник научных статей. Вып. 5. – Нижний Новгород: НГЛУ им. Н.А. Добролюбова, 2010. – С. 52-65.

Ермак, В. Д. Классификация?.. Типология... Идентификация!.. [Электронный ресурс] / В.Д. Ермак. – Режим доступа: <http://vlp.chat.ru/ermak5.html>. – Заглавие с экрана.

Зимовая, М. В. Многозначность в терминологии: автореф. дис. ... канд. филол. наук / М. В. Зимовая. – Орел, 2011. – 19 с.

Лаврова, А. А. Синтаксические особенности реализации эмоционального компонента в политической речи (на материале американских предвыборных теледебатов): дис. ... канд. филол. наук / А. А. Лаврова. – Нижний Новгород, 2010. – 224 с.

Лисенкова, О. А. Синтаксическая транспозиция в мужской и женской аффективной речи (на материале английского языка): дис. ... канд. филол. наук / О. А. Лисенкова. – Нижний Новгород, 2007. – 140 с.

Синеокова, Т. Н. Парадигматика эмоционального синтаксиса: Монография / Т. Н. Синеокова. – Нижний Новгород: Изд-во ННГУ им. Н.И. Лобачевского, 2003. – 244 с.

Синеокова, Т. Н. К вопросу об идентификации типа измененного состояния сознания литературного персонажа // Язык. Речь. Речевая деятельность: Межвузовский сборник научных трудов. Вып. 10. – Н. Новгород: Нижегородский государственный лингвистический университет им. Н.А. Добролюбова, 2008. – С. 136-147.

Синеокова, Т. Н. Интерактивные лингвистические классификации: статистические методы анализа: Монография / Т. Н. Синеокова. – Н. Новгород: Нижегородский государственный лингвистический университет им. Н.А. Добролюбова, 2009. – 162 с.

Синеокова, Т. Н. Референтная модель создания лингвистической классификации // Методология современного языкознания: Сборник статей / отв. ред. А. Г. Сонин, А. С. Баранов. – М.: АСОУ, 2010. – С. 189-200.

Синеокова, Т. Н., Лаврова, А. А., Чернышова, Е. И. Лингвометодический аспект классифицирования структурных форм ненейтральной речи // Теоретические и прикладные аспекты изучения речевой деятельности: Сборник научных статей. Вып. 1. – Нижний Новгород: НГЛУ им. Н.А. Добролюбова, 2006. – С. 40-58.

Сулейманова, О. А. К вопросу о типологии лингвистического эксперимента // Методология современного языкознания: Сборник статей / отв. ред. А. Г. Сонин, А. С. Баранов. – М.: АСОУ, 2010. – С. 230-241.

Фрумкина, Р. М. Психолингвистика: учеб. пособие / Р. М. Фрумкина. – М.: Академия, 2003. – 316 с.

Н.Н. Кошкарлова¹

Южно-Уральский государственный университет

ЛИНГВИСТИЧЕСКИЕ МЕХАНИЗМЫ РЕЧЕВОЙ АГРЕССИИ В ПОЛИТИЧЕСКОМ ИНТЕРВЬЮ МЕЖКУЛЬТУРНОГО УРОВНЯ

Статья посвящена анализу лингвистических механизмов речевой агрессии в политическом интервью межкультурного уровня. В работе предлагается следующая схема анализа: 1) общая тональность через употребление паремий; 2) грамматические особенности; 3) оценочность через употребление метафоры.

Ключевые слова: конфликтный дискурс, речевая агрессия, коммуникативный конфликт, паремии, грамматические особенности, метафорические модели

N.N. Koshkarova

¹ Кошкарлова Наталья Николаевна, кандидат филологических наук, доцент кафедры международных коммуникаций Южно-Уральского государственного университета (национального исследовательского университета), Челябинск.