

С.Ю. Куликов (ООО «2Толк»)

S.Yu. Kulikov (ООО «2Talk»)

Рецензия на книгу Перебийніс В.І. Математична лінгвістика.

К.: Вид. центр КНЛУ, 2014. — 125 с.

Рецензируемая работа классика советской и украинской прикладной лингвистики Валентины Исидоровны Перебийнос посвящена одному из ключевых направлений прикладной лингвистики — применению математических методов в языкознании. В основу учебного пособия лег курс «Математической лингвистики», читающийся на отделении прикладной (компьютерной) лингвистики Киевского национального лингвистического университета с 2008/2009 учебного года.

Пособие состоит из предисловия, восьми разделов и послесловия. Каждый из разделов завершается списком практических заданий и списком рекомендуемой литературы.

Первый раздел («Причины возникновения математической лингвистики») посвящен общетеоретическим вопросам, связанным с математической лингвистикой. В.И. Перебийнос указывает на различные интерпретации понятия «математическая лингвистика» в математике и прикладном языкознании (стр. 4—5). Данное разграничение, на наш взгляд, способствует формированию у студентов понимания терминологических особенностей междисциплинарных наук, к которым относится прикладная лингвистика. Следующий поднимаемый в пособии вопрос касается возможности применения математических методов к языку. Сама возможность использования математического аппарата убедительно объясняется через понятие «знака», которым оперируют и математика, и языкознание. Далее автор (вслед за Р.Г. Пиотровским) рассматривает различия между «математическим знаком» и «лингвистическим знаком». При этом автор не упоминает другие модели языкового знака (например, *трехстороннюю модель знака Огдена-Ричардса*), с помощью которых

можно так же описать нечеткие семантические множества. В этой связи несколько теряется особая значимость четырехкомпонентной модели знака Р.Г. Пиотровского для моделирования прагматики. Заключительный параграф данного раздела посвящен выбору конкретного математического метода для решения той или иной лингвистической задачи. Список математических методов традиционен и не включает относительно новые математические области — такие, как теория игр, которая активно применяется для моделирования дискурсивных и прагматических стратегий, например, в политической лингвистике.

Второй раздел посвящен количественным и статистическим методам в языкознании. В.И. Перебийнос проводит четкое разграничение между количественными и статистическими методами. Возможности применения статистического моделирования демонстрируются на примере статистической стилистики (моделирование порождения текста) (стр. 15—16). В параграфе «Правила организации статистического исследования» вводятся ключевые статистические понятия «единица анализа», «генеральная совокупность», «выборка», «подвыборка», «репрезентативность» и др. Следующие параграфы данного раздела рассматривают этапы развития лингвостатистики, а также различные статистические закономерности, характерные для текстов.

Третий раздел учебного пособия посвящен частотным словарям, их структуре и построению. Следует отметить, что В.И. Перебийнос является признанным специалистом в области статистической лексикографии, поэтому данный раздел в значительной степени опирается на собственный опыт автора. В разделе рассматриваются как пионерские работы Кэдинга и Торндайка и Лорджа, так и недавние работы под руководством В.И. Перебийнос. Также в разделе описывается закон Ципфа, являющийся основополагающим для статистической лексикографии.

Четвертый раздел посвящен методам теории множеств. В нем приводятся различия между комбинаторной и квантитативной лингвистикой, а также детализируются ключевые области применения комбинаторных методов к

лингвистическому материалу. В заключительном параграфе раздела описывается статистико-комбинаторный метод (подходы Б.В. Сухотина и А.Я. Шайкевича).

Пятый раздел пособия рассматривает вопросы анализа синтаксической структуры предложения. Приведены варианты разбора по *непосредственно составляющим* и *грамматике зависимостей*. Особое внимание уделено анализу непроективных конструкций.

Шестой раздел посвящен теории алгоритмов. В данном разделе наряду с теоретическими сведениями (определения, формы записи и др.) приведены формальные описания реализованных алгоритмов автоматической обработки естественного языка.

Седьмой раздел посвящен лингвистическим моделям и лингвистическому моделированию. Помимо классификаций лингвистических моделей (подходы Ю.Д. Апресяна и Е.А. Карпиловской) В.И. Перебийнос приводит примеры лингвистических моделей, описывающих различные уровни языка. В частности, рассмотрены фонетические, словообразовательная и словоизменительная модели. Далее автор обобщает требования к моделям, которые были сформулированы Р.Г. Пиотровским и Ю.Д. Апресяном. Следует заметить, что в данном разделе описаны ранние (и во многом забытые) модели буквенных классов Б.В. Сухотина и аппликативной порождающей грамматики С.К. Шаумяна.

Заключительный раздел пособия посвящен теории графов. В качестве примеров различных типов графов, а также в контрольных заданиях №№2 и 3, В.И. Перебийнос использует данные украинского языка, что должно способствовать закреплению навыков, полученных студентами в ходе изучения пятого раздела пособия.

В послесловии автор констатирует, что данное учебное пособие является вводным курсом в проблематику математической и компьютерной лингвистики. Предполагается, что каждый из разделов пособия будет охвачен отдельным курсом (стр. 123).

Жанр введения в проблематику позволяет делать отсылки к другим пособиям (прежде всего, к работам Е.А. Карпиловской и Н.П. Дарчук). Особый интерес

представляет и историческая составляющая учебного пособия: в нем приведены некоторые классические определения, модели и подходы, которые часто отсутствуют в современных учебных изданиях по компьютерной лингвистике. Большое значение имеет раздел, посвященный частотным словарям, т.к. он опирается на многолетний опыт составителя целого ряда подобных словарей.

Как указывалось выше, некоторые важные, на наш взгляд, моменты не нашли отражения в учебном пособии. В частности, в пособии не отражена роль корпусной лингвистики в современных лингвистических исследованиях. Также отсутствует информация о «машинном обучении», которое в последние два десятилетия понимается в западной компьютерной лингвистике как синоним статистического моделирования языка. Существенным недостатком, на наш взгляд, является выбор литературы к соответствующим разделам. Нельзя отрицать, что приводимый список литературы включает значимые работы, но большая часть из них относится к 1960-м—1980-м гг. Кроме того, в списке литературы отсутствуют фундаментальные англоязычные работы по рассматриваемым проблемам, например, [1; 2].

В заключение необходимо отметить, что рецензируемое пособие вносит значительный вклад в существующую учебную литературу по компьютерной лингвистике на украинском языке.

Литература

1. Jurafsky, M., Martin, J.H. Speech and language processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. 2nd ed. New Jersey, 2009.
2. Partee, B.H., ter Meulen, A.G., Wall, R. Mathematical Methods in Linguistics. Dordrecht, 1990.