

## **ПЛАН СОДЕРЖАНИЯ ЯЗЫКА И СТАТИСТИЧЕСКИЕ МЕТОДЫ ИССЛЕДОВАНИЯ**

План содержания является объектом внимания не только лингвистов, но и философов, психологов, математиков, представителей других наук. Основной причиной этого является признание того факта, что языковые сообщения представляют собой прежде всего обмен информацией, смыслом [1, с. 28–29].

Реальная действительность воспринимается человеком не непосредственно, а через систему знаков. Функционирование и развитие любого знака направляется целым рядом различных по характеру взаимосвязей с другими знаками.

Знак представляет собой материальный факт, служащий для обозначения реалий и используемый людьми для передачи информации. Отличительной особенностью языкового знака является универсальность его использования во всех сферах жизни человека, первичность по отношению к другим знакам, стихийное саморазвитие, эффективность, асимметричное соотношение формы и содержания. Указанными особенностями, а также спецификой упорядочивания элементов языковой системы, т.е. ее структурой, в определенной степени и обусловлено своеобразие языковой системы и ее отличие от других знаковых систем.

В настоящее время язык рассматривается лингвистами не как логическая, а как вероятностная система, ибо язык представляет собой сложную, развивающуюся, открытую, вторичную систему, в которой действуют как детерминированные, так и вероятностные элементы.

Неотъемлемым и обязательным атрибутом языковой системы являются семантические связи. Именно благодаря им план содержания языка представляет собой систему, т.е. совокупность взаимосвязанных и упорядоченных в единое целое языковых элементов [2, с. 123].

Исследование и познание языкового содержания предполагают учет всевозможных факторов существования и функционирования языка. В лингвистике сложилась устойчивая традиция исследовать, прежде всего, те значения, которые выражены и закреплены конкретной языковой формой. Сложные языковые формы, в которых реализуются те или иные виды значения, а также определенные семантические свойства языковых единиц, можно изучать, применяя разнообразные методы исследования.

Каждый метод имеет сферу своего приложения и служит оптимальным средством раскрытия определенных семантических свойств языковых единиц и постижения сущности семантических явлений. Выбор методов анализа и

описания лексических единиц определяется в первую очередь целью исследования и характером имеющегося материала.

Все большее признание и применение в самых разнообразных областях знаний в последние годы находит математическая статистика и статистические методы. Действие статистических законов обнаруживается в очень широкой области различных явлений жизни природы и общества. Все сложные системы (структуры) подчиняются в своем развитии и функционировании статистическим законам (развитие умственной деятельности человека, формирование нравственных качеств личности, развитие речи и т.д.) [3, с. 11].

Закономерно использование математических методов и в различных гуманитарных науках. Точное описание любого языка непременно нуждается в количественных характеристиках всех языковых элементов и явлений (морфемы, части речи, грамматические формы и значения, члены предложения, типы словосочетаний, разновидности лексической, морфологической и синтаксической сочетаемости и т.д.). Основное достоинство статистических методов состоит в возможности получения достоверных результатов. Исследователем создается фундамент из систематизированных фактов, они выстраиваются в упорядоченную систему объективных количественных показателей [4, с. 335].

Статистические методы включают в себя разнообразные формулы, применяемые для выявления правил распределения языковых единиц в речи, для измерения различных формальных и качественных явлений языка, для устранения тех или иных тенденций в развитии и функционировании языка, для изучения различных семантических явлений – определения значимости исследуемых семантических явлений, определения веса и значимости семантических признаков слов, установления сходства семантики слов, силы семантических связей между словами с целью раскрытия структуры лексических подсистем и семантических полей, для определения меры синонимичности и антонимичности исследуемых слов, для определения семантического объема слов и других задач. Неодинаковая интенсивность проявления различных языковых явлений, настоятельно требующая количественных исследований, наиболее зримо видна именно при реализации разнообразных семантических свойств языковых единиц. Количественная оценка и характеристика качественных явлений языка представляет собой составную часть исследования сущности рассматриваемого языкового факта [2, с. 184].

Для учета отношений между языковыми единицами, их распределения в тексте получил применение дистрибутивный метод. Дистрибутивно-статистическая методика была предложена А.Е. Супруном для определения силы семантических связей между членами заданной лексической подсистемы. Суть методики заключается в статистическом анализе слов, сочетающихся с рассматриваемыми лексемами в тексте. В лексической сочетаемости проявляются синтаксические свойства слова, Сравнивая сочетаемость интересующих нас слов, мы можем получить о них

семантическую информацию. В сочетаемости слов отражается значение этих слов, характеризуемое определенной совокупностью признаков, т.е. значение слова проявляется, манифестируется в его сочетаниях с другими словами, в его употреблении. Есть основания предположить, что если мы будем иметь набор всех сочетаний, в которых употребляется данное слово, то мы сможем вполне определенно представить его значение [5, с. 79–80].

Анализ особенности семантики определенной группы слов позволяет сравнить весь спектр сочетаемости исследуемых слов для выделения близости их значений одновременно по всем свойственным данным словам семантическим признакам. Одним из показателей такой близости является коэффициент корреляции, вычисляемый по формуле:

$$r_{xy} = \frac{\sum (x_1 - \bar{x})(y_1 - \bar{y})}{\sqrt{\sum (x_1 - \bar{x})^2 \sum (y_1 - \bar{y})^2}}, \text{ где}$$

$r$  – коэффициент корреляции,

$x_1, y_1$  – размеры признаков, присущих двум сопоставляемым объектам,

$\bar{x}, \bar{y}$  – средние размеры признаков.

При помощи коэффициента корреляции мера связи выражается в пределах от  $-1$  до  $+1$ . Наиболее тесной мерой прямой связи двух объектов по сопоставляемым признакам является  $+1$ , самая высокая мера обратной связи выражается величиной  $-1$  [6, с. 37].

Необходимо отметить, что учет корреляции используется и при изучении спектров сочетаемости семантически родственных слов для установления их отношений между собой, и для анализа изменений в сочетаемости одного и того же слова за некоторое время, и при сравнении выражаемых в сочетании семантических свойств соотносимых слов в родственных языках [5, с. 87–88].

Статистический метод, как и любой другой метод исследования, не может быть признан универсальным для изучения всех аспектов и свойств предмета. Но в сочетании с другими методами лингвистического анализа статистический метод дает более объемную и полную характеристику различным явлениям языка.

## СПИСОК ЛИТЕРАТУРЫ

1. Богданов, В. В. Предпосылки формализации смысла предложения / В. В. Богданов // Вопросы общей и прикладной лингвистики. – Минск : Выш. школа, 1975. – С. 28–32.
2. Плотников, Б. А. Основы семасиологии / Б. А. Плотников. – Минск : Выш. школа. – 1984. – 223 с.
3. Головин, Б. Н. Из курса лекций по лингвистической статистике / Б. Н. Головин. – Горький : ГГУ, 1966. – 96 с.
4. Граудина, Л. К. Норма и статистика / Л. К. Граудина // Актуальные проблемы культуры речи. – М. : Наука, 1970. – С. 326–368.
5. Супрун, А. Е. Лекции по языковедению / А. Е. Супрун. – Минск : Изд-во БГУ, 1978. – 144 с.

6. Плотников, Б. А. Дистрибутивно-статистический анализ лексических значений / Б. А. Плотников. – Минск : Выш. школа, 1978. – 136 с.