

УДК 004.43

Контрастивное изучение синтаксиса естественных языков и языков программирования

Герасимова Софья Вячеславовна

gerasimovasv@student.bmstu.ru

Крыленкина Антонина Вячеславовна

krilenkina@bmstu.ru

Белов Юрий Сергеевич

ysbelov@bmstu.ru

КФ МГТУ им. Н.Э. Баумана, Калуга, Россия

Представлен обзор синтаксиса естественных языков и языков программирования, рассмотрены их сходства и различия. Естественные языки характеризуются гибкостью, контекстуальной зависимостью и вариативностью, языки программирования обладают строгим и формализованным синтаксисом. Оба типа языков имеют иерархическую структуру и поддерживают рекурсию, что подчеркивает их общие черты. Исследование синтаксиса этих языковых систем имеет важные практические приложения, включая разработку более интуитивных языков программирования и улучшение систем автоматической обработки естественного языка.

Ключевые слова: синтаксис, естественные языки, языки программирования, лингвистика, гибкость, формализация, рекурсия, иерархическая структура, контекстуальная зависимость, неоднозначность, нейронные сети

Сравнительный анализ синтаксиса естественных языков и языков программирования представляет собой междисциплинарную область исследований, которая объединяет лингвистику, компьютерные науки и когнитивную психологию. Несмотря на то, что естественные языки и языки программирования служат разным целям, их синтаксические структуры имеют как сходства, так и различия. В последние десятилетия развитие искусственного интеллекта и нейронных сетей привело к необходимости глубокого анализа этих двух языковых систем.

Естественные языки отчасти обладают высокой степенью вариативности и допускают некоторую свободу в порядке слов, использовании синонимов и грамматических конструкций. Например, в русском языке допустимо сказать как «Я прочитала книгу», так и «Книгу я прочитала», не изменяя смысла высказывания, но придавая ему различные прагматические оттенки. Языки программирования, напротив, имеют строгий синтаксис, где даже небольшое отклонение (например, пропущенная скобка или неверный оператор) ведет к синтаксическим ошибкам. Фразы на естественных языках значительно зависят от контекста и подразумевают смысловую интерпретацию. Языки программирования опираются на формально определенную грамматику, где каждая конструкция имеет четко определенное значение.

Как естественные, так и программные языки имеют иерархическую структуру. В естественных языках предложения строятся из фраз, а фразы состоят из слов. В языках программирования аналогичная иерархия проявля-

ется в виде операторов, выражений и блоков кода. Например, конструкция «while (condition) { statements }» является аналогом сложноподчиненного предложения в обычном языке.

Синтаксис естественных языков характеризуется высокой степенью гибкости и контекстуальной зависимостью. В данном случае синтаксические структуры часто допускают вариативность порядка слов, опущение элементов и использование ссылок на предыдущие элементы в тексте [1]. Например, в естественном языке можно сказать: «Он пошел в магазин, а потом — домой», опуская глагол «пошел». В языках программирования подобное опущение невозможно.

Кроме того, синтаксис естественных языков тесно связан с семантикой и прагматикой. Контекст играет ключевую роль в интерпретации предложений, что делает естественные языки мощным инструментом для передачи сложных идей и эмоций [2]. Однако эта гибкость также приводит к неоднозначности, которая может затруднять автоматическую обработку текста.

В отличие от естественных языков, синтаксис языков программирования строго формализован и однозначен. Программные языки разрабатываются с целью минимизировать неоднозначность и обеспечить точное выполнение инструкций компьютером. Например, в языке Python синтаксис требует строгого соблюдения правил отступов для обозначения блоков кода, а в языке C++ каждая команда должна заканчиваться точкой с запятой. Однозначность синтаксиса позволяет компиляторам и интерпретаторам точно анализировать и выполнять код. Однако эта строгость также ограничивает выразительные возможности языка. В отличие от случая с естественными языками, где контекст может прояснить значение, в программировании контекстная зависимость отсутствует, и каждая конструкция должна быть явно определена. В программировании порядок элементов строго фиксирован. Например, в языке Java вызов метода всегда следует за объектом, к которому он применяется: `object.method()`. Это правило не допускает вариаций [3].

Оба типа языков обладают свойством рекурсивности. В естественных языках это проявляется, например в виде вложенных придаточных предложений: «Я знаю, что он сказал, что он придет завтра». В языках программирования рекурсия используется для вызова функций внутри самих себя, что позволяет эффективно решать задачи с повторяющейся структурой.

Изучение синтаксиса естественных и программных языков также имеет важные когнитивные последствия. Исследования показывают, что освоение синтаксиса языка программирования требует от пользователей развития специфических когнитивных навыков, таких как способность к абстрактному мышлению и точному следованию правилам. В то же время, пользователи естественных языков полагаются на интуитивное понимание контекста и гибкость синтаксиса.

Контрастивное изучение синтаксиса двух типов языков имеет важные практические приложения. Например, понимание сходств и различий между синтаксическими структурами может помочь в разработке интуитивных языков программирования, которые будут легче осваиваться новичками. Напри-

мер, Python во многом напоминает естественный язык, что делает его более доступным для начинающих программистов. Кроме того, результаты таких исследований могут быть использованы для улучшения систем автоматической обработки естественного языка (NLP), которые часто сталкиваются с проблемами неоднозначности и контекстуальной зависимости [4].

Синтаксис естественных языков и языков программирования, несмотря на значительные различия, обладает рядом общих структурных свойств. Контрастивный анализ этих языковых систем дает ценные результаты для лингвистики, информатики и разработки ИИ-моделей. Современные нейросетевые технологии все больше объединяют методы обработки естественного и программного языка, что позволяет создавать более эффективные и адаптивные системы автоматизированного анализа и генерации текста и кода. Исследование и анализ синтаксиса не только углубляет понимание языковых структур, но и способствует развитию технологий, которые делают взаимодействие человека с машиной более естественным и эффективным.

Литература

- [1] Стародумова Е.А. *Синтаксис современного русского языка*. Москва, ФЛИНТА, 2023, 260 с.
- [2] Шاپочкин Д.В., Богданова Л.И., Пчелинцева И.Г. *Прототипические модели в языках и дискурсах*. Тюмень, ТюмГУ, 2018, 148 с.
- [3] Пруцков А.В. *Язык программирования Java. Введение в курс: операторы и типы данных*. Рязань, РГРТУ, 2016, 72 с.
- [4] Ганегедара Т. *Обработка естественного языка с TensorFlow: руководство*. Москва, ДМК Пресс, 2020, 382 с.

Contrastive study of the syntax of natural languages and programming languages

Gerasimova Sofia Vyacheslavovna

gerasimovasv@student.bmstu.ru

Krylenkina Antonina Vyacheslavovna

krilenkina@bmstu.ru

Belov Yuri Sergeevich

ysbelov@bmstu.ru

Kaluga Branch of BMSTU, Kaluga, Russia

This article provides an overview of the syntax of natural languages and programming languages, their similarities and differences. Natural languages are characterized by flexibility, contextual dependence and variability, programming languages have strict and formalized syntax. Both types of languages have a hierarchical structure and support recursion, which highlights their common features. The study of the syntax of these language systems has important practical applications, including the development of more intuitive programming languages and the improvement of automatic natural language processing systems.

Keywords: syntax, natural languages, programming languages, linguistics, flexibility, formalization, recursion, hierarchical structure, contextual dependence, ambiguity, neural networks