

*А. Н. Гордей, д. филол. н.
Минский государственный лингвистический университет, Минск, Беларусь
e-mail: alieks2001@yahoo.com*

*A. Hardzei, Doctor of Sc. (Philology)
Minsk State Linguistic University, Minsk, Belarus
e-mail: alieks2001@yahoo.com*

ЧАСТИ ЯЗЫКА И ЧЛЕНЫ ПРЕДЛОЖЕНИЯ

Пропедевтика к автоматической семантической разметке структурированного и неструктурированного контента для систем обработки и понимания естественного языка.

Ключевые слова: комбинаторная семантика; динамический синтаксис; морфология; части языка; члены предложения; субстантивы (тайгены); предикативы (ёгены); знаки алфавита синтаксиса; лексема; комбинаторный вариант лексемы; стереотип; субъект; акция; объект.

PARTS OF LANGUAGE AND PARTS OF SENTENCE

The presented approach is a pre-description for automatic semantic markup of structured and unstructured content for Natural Language Processing and Natural Language Understanding.

Key words: Combinatory Semantics; Dynamic Syntax; Morphology; Parts of Language; Parts of Sentence; substantives (taigenes); predicatives (yogenes); Signs of Syntax Alphabet; lexeme; combinatory version of a lexeme; pattern; subject; action; object.

1. **Семантика** – лингвистическая дисциплина, изучающая отношение языка к модели мира. Под *семантикой* также понимается *содержание* стереотипов, *значение* знаков и *смысл* предложений. **Комбинаторная семантика** исследует отображение языком динамики ролей индивидов в событии (в отличие от **комбинаторной**, которая изучает совместную встречаемость знаков статистическими методами). Таким образом, семантика как отношение

языка к модели мира проявляется в динамике ролей индивидов в событии, что находит своё отражение в содержании стереотипов, значении знаков и смысле предложений.

2. **Части языка** – подмножества языковой системы, элементами которых являются знаки с общим предельно абстрактным значением, а именно: **субстантивы (тайгены)** и **предикативы (ёгены)**, обозначающие, соответственно, индивидов и признаков индивидов. **Постоянные субстантивы** обозначают множества однородных индивидов (*i-const*), напр. *слесарь*, **переменные** – разнородных (*i-var*), напр. *он*; **постоянные предикативы** обозначают множества свойств индивидов *p(i)*, напр. *слесарный*, **переменные** – процессов *f(i)*, в которых участвуют индивиды, напр: *слесарить*. **Предикативы первой степени** обозначают основной признак индивида: *бежевый* → *цвет*, *сжимать* → *пружину*. **Предикативы второй положительной степени (оценочные)** сужают значения предикативов первой степени: *красивый* → *бежевый*, *быстро* → *сжимать*, а **предикативы третьей положительной степени** – значения предикативов второй положительной степени: *самый* → *красивый*, *очень* → *быстро*. Предикативы первой степени, в свою очередь, сужают значения предикативов **второй отрицательной степени (модальных)**: *рад* ← *видеть*, *любить* ← *бегать*, а предикативы **третьей отрицательной степени** – значения предикативов второй отрицательной степени: *очень* → *рад*, *очень* → *любить*. **Лексема** – субстантив или предикатив конкретного естественного языка; будучи знаком, имеет в аспекте выражения комбинацию фигур, а в аспекте содержания – стереотип; в синтетических языках обладает развитой морфологической парадигмой; является центральной единицей лексикографического описания. **Комбинаторный вариант лексемы** – вариант лексемы в упорядоченном наборе её вариантов. **Номинативная единица** – устойчивая последовательность комбинаторных вариантов лексем, в которой один вариант лексем (модификатор) определяет другой (актуализатор), напр.: *записная книжка*, *бежать галопом*. **Морфология** – лингвистическая дисциплина, изучающая зависимость формы знака от его комбинаторики. Под **морфологией** также понимается вариант формы знака в упорядоченном наборе её вариантов (**морфологической парадигме**). Если у формы знака нет вариантов или они для комбинаторики являются неупорядоченными (нерелевантными), напр.: *резюм/э/* vs. *резюм/е/*, то знак считается **аморфным**. Комбинаторный вариант включается в морфологическую (формоизменяющую) парадигму лексем только при отсутствии изменений в её семантике (качественных, количественных, временных, родовых, видовых и др.). Все диахронические (глоттогонические) и синхронические (семиозисные) генеративные (формообразовательные) лексические парадигмы относятся к синтаксису, т. е. к микросинтаксическим правилам построения и преобразования лексем.

3. **Синтаксис** – лингвистическая дисциплина, изучающая комбинаторику составных частей языковых структур. Под **синтаксисом** также понимается упорядоченность языковых структур и техника (алфавит и правила) их построения и преобразования. **Знаки алфавита синтаксиса** – вспомогательные средства синтаксиса (на макроуровне – частицы, на микроуровне – аффиксы), служащие для соединения составных частей языковых структур и образования морфологических парадигм. Под **динамическим синтаксисом** понимается порождение второстепенных членов предложения из главных, а главных членов – из глоттогонического ядра с целью достижения баланса противоположных коммуникативных стратегий говорящего и слушающего: свёртки высказывания для экономии времени и мускульной энергии (принцип экономии [1]) и развёртки высказывания для облегчения восприятия информации (принцип запаса прочности). **Предложение** – семантически некоммутативная последовательность комбинаторных вариантов знаков, ср.: *Я ем суп* и **Суп ест меня*. Делятся на: а) **повествовательные** и **вопросительные** (по форме); б) **утвердительные** и **побудительные** (по цели); в) **положительные** и **отрицательные** (по оценке); г) **нейтральные** и **восклицательные** (по окраске).

4. **Члены предложения** – роли частей языка в предложении. **Подлежащее** – исходный пункт описания события, выбранный наблюдателем, **прямое дополнение** – конечный пункт описания, **сказуемое** – отображение наблюдателем исходного пункта описания события в конечный. С точки зрения семантики и синтаксиса подлежащее, сказуемое и прямое дополнение являются тремя главными членами предложения, потому что, во-первых, при прямой предикации отображаются в **субъект, акцию** и **объект**, без которых невозможно событие, во-вторых, имеют зависимые члены: подлежащее и прямое дополнение – определения, сказуемое – обстоятельства. **Обстоятельство** – второстепенный член предложения, модифицирующий либо локализирующий сказуемое; обстоятельство степени и обстоятельство образа действия модифицируют сказуемое, обстоятельства места и обстоятельство времени локализуют сказуемое в пространстве и, соответственно, во времени. **Определение** – второстепенный член предложения, модифицирующий подлежащее, дополнение, обстоятельство места и времени.

6. В 2018 г. в базу Web of Science было включено 1,6 млн. статей. В среднем рост числа рецензируемых журналов составляет 4 % в год [2]. **Без автоматического интеллектуального поиска и семантического анализа текстов уже не обойтись**, над чем усиленно трудятся специалисты в рамках проекта Semantic Web и ресурса Schema.org разработчиков Google, Microsoft, Yahoo и Yandex.

ЛИТЕРАТУРА

1. Мартине А. Основы общей лингвистики // Новое в лингвистике. 1963. Вып. 3. С. 366–566.

2. Домнина Т. Н., Хачко О. А. Научные журналы: количество, темпы роста. URL: http://www.benran.ru/SEM/Sb_15/sbornik/83.pdf (дата обращения: 05.08.2021).