

# МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА

Авторы: А. В. Гладкий

МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА, математическая дисциплина, разрабатывающая формальный аппарат для описания строения естественных и некоторых [искусственных языков](#). Возникла в 1950-х гг., когда в языкознании назрела потребность уточнения его осн. понятий. В М. л. используются по преимуществу идеи и методы [алгебры](#), [алгоритмов теории](#) и [автоматов теории](#). Не являясь частью лингвистики, М. л. развивается в тесном взаимодействии с ней. Иногда термин «М. л.» используют также для обозначения любых лингвистич. исследований, в которых применяется к.-л. математич. аппарат.



Рис. 1.

Лошади ← кушают → овёс

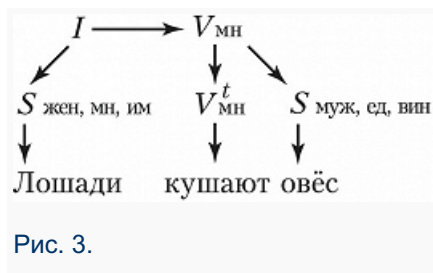
Рис. 2.

Математич. описание языка основано на восходящем к Ф. де [Соссюру](#) представлении о языке как о механизме, функционирование которого проявляется в речевой деятельности его носителей; её результатом являются «правильные тексты» — последовательности речевых единиц, подчиняющиеся определённым закономерностям, многие из которых допускают математич. описание. Изучение способов математич. описания правильных текстов (в первую очередь предложений) составляет содержание одного из разделов М. л. — теории способов описания

синтаксической структуры. Для описания строения (синтаксич. структуры) предложения можно либо выделить в нём [составляющие](#) — группы слов, функционирующие как цельные синтаксич. единицы, либо указать для каждого слова те слова, которые от него непосредственно зависят (если такие есть). Так, в предложении «Лошади кушают овёс» при описании по 1-му способу составляющими будут: всё предложение I, каждое отд. слово и словосочетание C = «кушают овёс» (рис. 1; стрелки означают «непосредственное вложение»); описание по 2-му способу даёт схему, показанную на рис. 2. Математич. объекты, возникающие при таком описании структуры предложения, называются деревом составляющих (1-й способ) и деревом синтаксич. подчинения (2-й способ).

Другой раздел М. л., занимающий в ней центр. место, — теория формальных грамматик, возникшая гл. обр. благодаря работам Н. [Хомского](#). Она изучает способы описания закономерностей, которые характеризуют уже не отд. текст, а всю совокупность правильных текстов того или иного языка. Эти закономерности описываются путём построения «формальной грамматики» — абстрактного «механизма», позволяющего с помощью единообразной процедуры получать правильные тексты данного языка вместе с описаниями их структуры. Наиболее широко используемый тип формальной грамматики — т. н. порождающая грамматика, или грамматика Хомского, — упорядоченная система  $\Gamma = \langle V, W, I, R \rangle$ , где  $V$  и  $W$  — непересекающиеся конечные множества;  $I$  — элемент  $W$ ;  $R$  — конечное множество  $p \rightarrow \psi$ , где  $\phi$  и  $\psi$  — цепочки (конечные последовательности) элементов  $V$  и  $W$ . Если  $\phi \rightarrow \psi$  — правило грамматики  $\Gamma$  и  $\omega_1, \omega_2$  — цепочки из элементов  $V$  и  $W$ , то говорят, что цепочка  $\omega_1 \psi \omega_2$   $\vdash \omega_1 \phi \omega_2$  в  $\Gamma$  из  $\omega_1 \phi \omega_2$ . Если  $\xi_0, \xi_1, \dots, \xi_n$  — цепочки и для каждого  $i=1, \dots, n$  цепочка  $\xi_i$  непосредственно выводима из  $\xi_{i-1}$ , то говорят, что  $\xi_n$

$\forall \langle \lambda_1, \dots, \lambda_n \rangle \in \Sigma_0^n$  в  $\Gamma$ . Множество цепочек из элементов  $V$ , выводимых в  $\Gamma$  из  $I$ , называют языком, порождаемым грамматикой  $\Gamma$ . Если все правила грамматики  $\Gamma$  имеют вид  $A \rightarrow \psi$ , где  $A$  – элемент  $W$ ,  $\Gamma$  называется бесконтекстной, или контекстно-свободной. В лингвистич. интерпретации элементы  $V$  чаще всего представляют собой слова, элементы  $W$  – символы грамматич. категорий,  $I$  – символ категории «предложение». В бесконтекстной грамматике вывод предложения даёт для него дерево составляющих, в котором каждая составляющая состоит из слов, «происходящих» от одного элемента  $W$ , так что для каждой составляющей указывается её грамматич. категория. Так, если грамматика имеет в числе прочих правила  $I \rightarrow S_{\langle x, y, \text{им} \rangle} V_y$ ,  $V_y \rightarrow V^t_y S_{\langle x, y, \text{вин} \rangle}$ ,  $S_{\langle \text{муж, ед, вин} \rangle} \rightarrow \text{овёс}$ ,  $S_{\langle \text{жен, мн, им} \rangle} \rightarrow \text{лошади}$



$V^t_{\langle \text{мн} \rangle} \rightarrow \text{кушают}$ , где  $V_y$  означает категорию «группа глагола в числе  $y$ »,  $V^t_y$  – «переходный глагол в числе  $y$ »,  $S_{\langle x, y, z \rangle}$  – «существительное рода  $x$  в числе  $y$  и падеже  $z$ », то приведённое выше предложение имеет вывод, показанный на рис. 3, где стрелки идут из левых частей применяемых правил к элементам соответствующих правых частей. Порождающие грамматики используются для описания не только естественных, но и искусств. языков, в особенности [языков программирования](#). Другой тип

формальных грамматик – доминационные грамматики, которые порождают цепочки, интерпретируемые обычно как предложения, вместе с деревьями синтаксич. подчинения.

М. л. изучает также аналитич. модели языка, в которых на основе тех или иных данных о речи, считающихся известными (напр., множества правильных предложений), производятся формальные построения, дающие некоторые сведения о структуре языка. Приложение методов М. л. к конкретным языкам относится к области лингвистики (см. [Языкознание](#)).

## Литература

Лит.: Хомский Н. Синтаксические структуры // Новое в лингвистике. М., 1962. Вып. 2; Гладкий А. В., Мельчук И. А. Элементы математической лингвистики. М., 1969 (библ.); Маркус С. Теоретико-множественные модели языков. М., 1970; Гладкий А. В. Формальные грамматики и языки. М., 1973 (библ.); он же. Попытка формального определения понятий падежа и рода существительного // Проблемы грамматического моделирования. М., 1973; он же. набросок формальной теории падежа // Вопросы языкознания. 1999. № 5; он же. Синтаксические структуры естественного языка. 2-е изд. М., 2007 (библ.); Montague R. Formal philosophy. New Haven; L., 1974; Сухотин Б. В. Оптимизационные методы исследования языка. М., 1976 (библ.); он же. Исследование грамматики числовыми методами. М., 1990; Севбо И. П. Графическое представление синтаксических структур и стилистическая диагностика. К., 1981; Мельчук И. А. Опыт теории лингвистических моделей «Смысл  $\Leftrightarrow$  Текст». М., 1999; Парти Б. Х. Грамматика Монтегю, мысленные представления и реальность // Семиотика. 2-е изд. М.; Екатеринбург, 2001; Dekhtyar M., Dikovskiy A. Generalized categorial dependency grammars // Pillars of computer science / Ed. A. Avron a. o. B., 2008.