

**ЗАРИПОВ Р. Р., ЗАХАРОВА О. И.**  
**ИНТЕГРАЦИЯ ФОРМАЛЬНЫХ ЛИНГВИСТИЧЕСКИХ ЗНАНИЙ**  
**(СИНТАКСИЧЕСКИХ ДЕРЕВЬЕВ, СЕМАНТИЧЕСКИХ РОЛЕЙ)**  
**В ПРОЦЕСС ОБУЧЕНИЯ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ**

УДК 004.93:81'322, ГРНТИ 28.23.37

Статья поступила в редакцию 14.01.2026

Интеграция формальных лингвистических знаний (синтаксических деревьев, семантических ролей) в процесс обучения нейросетевых моделей

Integration of formal linguistic knowledge (syntactic trees, semantic roles) into the training process of neural network models

**Р. Р. Зарипов, О. И. Захарова**

**R. R. Zaripov, O. I. Zakharova**

Поволжский государственный университет телекоммуникаций и информатики, г. Самара

Povolzhskiy State University of Telecommunications and Informatics, Samara

*В статье рассматриваются гибридные нейросетевые подходы, объединяющие статистические методы машинного обучения с формальными лингвистическими знаниями, такими как синтаксические структуры и семантические роли. Показано, что их интеграция повышает качество и интерпретируемость моделей обработки естественного языка.*

*This article examines hybrid neural network approaches that combine statistical machine learning methods with formal linguistic knowledge, such as syntactic structures and semantic roles. It is shown that their integration improves the quality and interpretability of natural language processing models.*

**Ключевые слова:** формальные лингвистические знания, синтаксическое дерево, семантические роли, гибридные модели, нейронные сети

**Keywords:** formal linguistic knowledge, syntax tree, semantic roles, hybrid models, neural networks

### **Введение**

На протяжении истории обработки естественного языка (ОЕЯ) существовали два основных подхода: символичный (правила, формальные грамматики, логика) и статистический (машинное обучение, нейронные сети). Традиционные методы строились на грамматических правилах и детально аннотированных корпусах. Современные нейросетевые модели позволяют обучаться «с нуля» на больших данных, однако сами по себе они представляют собой «чёрный ящик»: сложно объяснить, какие именно закономерности языка

они выучивают. Как отмечают Кулмизев и Нивре, из-за этого исследователи зачастую прибегают к лингвистической теории для интерпретации поведения моделей [1]. В частности, синтаксические и семантические структуры традиционно оказываются важными при анализе и улучшении нейросетевых систем.

С одной стороны, современные нейросети могут аппроксимировать множество явлений на основе статистических шаблонов. С другой – без явной языковой структуры они обычно игнорируют грамматическую информацию. Например, системы машинного перевода часто обрабатывают текст как последовательность токенов, «не учитывая никакой явной синтаксической структуры» [2] (следует отметить, что это наблюдение базируется на исследованиях экспертов Нейронного Машинного Перевода). Отсутствие языковых признаков снижает интерпретируемость результатов и может ухудшать обобщение модели при ограниченных данных. Поэтому гибридные подходы пытаются совместить статистическую мощь нейросетей и строгость формальных лингвистических правил. Цель нашей статьи – обзор таких подходов: как формальная информация о синтаксисе и семантике может быть встроена в обучение современных нейронных архитектур, и какую практическую пользу это приносит в различных NLP-задачах (машинный перевод, парсинг, семантический анализ, извлечение информации и другие).

### **Гибридные модели с учетом синтаксиса**

Рассмотрим синтаксическую информацию и нейросети. Синтаксис традиционно играет ключевую роль во многих задачах. Например, в семантической ролевой разметке (Semantic Role Labeling SRL) с момента зарождения задачи синтаксические структуры считаются критичным фактором качества решения: «с самого начала было показано, что синтаксис критически важен для производительности систем SRL. Большинство моделей использует синтаксически-выведенные признаки и синтаксическое отсечение кандидатов» [3]. Даже сегодня исследователи изучают, необходимо ли явно моделировать синтаксис в нейронных системах или нейросеть «учится» синтаксическим паттернам самостоятельно. В работе Ши и других показано, что три синтаксических шаблона описывают 98% семантических отношений в корпусах PropBank для английского и китайского языков. Авторы свели задачу SRL к синтаксическому парсингу, конвертируя семантические рёбра в узлы и показывая, что обычный парсер даёт конкурентоспособные результаты для SRL [3]. Такое исследование подчёркивает близость синтаксиса и семантики: синтаксическое представление оказывается информативно и само по себе для решения семантических задач.

В нейросетевых моделях синтаксис часто интегрируют с помощью структурных слоёв или дополнительных входных признаков. Например, Маркеджани и Титов используют графовые сверточные сети (GCN), которые работают поверх дерева зависимостей предложения. Они кодируют дерево зависимостей в скрытые представления слов, а затем совмещают эти слои GCN с рекуррентными LSTM. Эксперименты показали, что такая «двухслойная»

модель (GCN+LSTM) существенно превосходит чистую LSTM-модель в задаче SRL: при включении синтаксических GCN-слоёв F1-мера SRL заметно возросла, и были получены лучшие результаты на CoNLL-2009 [4]. Это иллюстрирует, что в задаче семантической разметки явное учёта синтаксиса приносит существенное улучшение.

Другой подход – встраивать деревья зависимостей или составляющих через специальные нейронные операции. Так называемые Tree-LSTM модели (Тай и другие) «распускают» LSTM по дереву, а не по последовательности. Недавняя модификация – «Typed Dependency Tree-LSTM» – учитывает не только структуру дерева, но и типы зависимостей: модель строит векторное представление предложения, опираясь на дерево зависимостей и метки отношений. В эксперименте на задачах оценки семантической близости предложений и сентимент-анализа такая модель превзошла аналог без учёта типов зависимостей. Авторы сделали вывод о «существенном влиянии грамматических отношений на моделирование предложений» [5]. Это свидетельствует, что обогащение нейросети информацией о связях в дереве (в том числе метками семантических ролей или другими функциями) усиливает её способность захватывать смысловые нюансы.

Более того, реализуются гибридные модели, которые совместно обучаются на нескольких задачах. Например, Strubell и другие предложили модель LISA (Linguistically-Informed Self-Attention), в которой архитектура Transformer дополняется отдельным механизмом внимания, обучающимся выделять синтаксического «родителя» для каждого токена. Модель одновременно решает задачу синтаксического парсинга, разметки частей речи и семантической разметки. Эффект оказался заметным: при наличии предсказанной синтаксической разметки LISA улучшила F1 по SRL на 2.5% (новостной домен) и более чем на 2.0% (разнородный домен) [6]. Авторы подчёркивают, что «имитация» синтаксических отношений через самообучение внимания даёт заметный прирост точности. Это наглядный пример совмещения «жесткой» лингвистической структуры (деревьев зависимостей) с глубоким обучением.

Наконец, в последние годы популярны графовые нейронные сети (GNN) для кодирования синтаксических деревьев. Такие модели могут обрабатывать дерево зависимостей или AMR-граф как входной граф, передавая сообщения по ребрам. Современные исследования показывают, что гибридные нейросети, объединяющие классические LLM (например, Transformer) с GNN над синтаксическими деревьями, дают стабильный прирост на задачах, требующих структурного понимания (семантический парсинг, многопереходные запросы, умозаключения). По словам авторов, «ядро NLP-задач – это структурированные данные... например, AMR улучшает семантический парсинг, а графы знаний улучшают рассуждения – однако до сих пор недостаточно работ, объединяющих эти подходы с современными LLM. Мы показываем, что такое слияние через GNN позволяет повысить интерпретируемость и точность в задачах, требующих и лингвистической структуры, и мировых знаний» [7]. Данный тренд указывает на перспективность дальнейшей интеграции формальных графовых представлений синтаксиса в нейронные архитектуры.

## Применение синтаксических знаний в задачах

Рассмотрим машинный перевод. В системах НМТ исторически были попытки добавить лингвистические факторы (факторы Лангера и другие). Недавние работы предлагают «лексифицированные» представления входных текстов. Например, Casas и соавторы предложили гибридные словари, в которых слово кодируется не только под словами, но и леммой и морфологическими признаками. В их эксперименте такие лингвистически обусловленные словари улучшили перевод особенно для морфологически богатых языков и на неродной доменной тематике [8]. Аналогично, Casas и другие разработали схему *sparsely factored NMT*, где каждый фактор (напр., «ед.ч», «им.п.» и т.д.) кодируется отдельно. Авторы добились значительного улучшения качества перевода «вне домена» при малых ресурсах на примерах баскского и немецкого языков [9]. Таким образом, явное включение грамматических или лексико-морфологических признаков в словарь нейросети улучшает устойчивость моделей при дефиците данных и богатой морфологии.

Кроме того, в машинном переводе начало интегрировать полную структурную информацию: последние исследования показывают, что семантические графы предложений (AMR) могут служить дополнительным источником знания. Дай и другие используют графовые рекуррентные сети для кодирования AMR и получают заметный прирост качества NMT. Они отмечают, что структурная информация из AMR-графов несёт знание, ортогональное тексту, и позволяет компенсировать малый объём данных [10]. Подобные результаты наблюдались также у других авторов (Damonte и Cohen) – AMR действительно дополняет перевод, фиксируя семантические отношения (например, «Джон-жена»), которые сложно уловить по видимой последовательности слов. Таким образом, в задачах перевода гибридные модели с AMR- или синтаксическими фичами могут повысить качество, особенно при переводе с/на сложные языки и в условиях доменного сдвига.

Рассмотрим синтаксический разбор. Сам по себе синтаксический парсинг часто решается нейросетями (LSTM-парсеры, трансформеры и т.п.), но даже в таких моделях иногда вводят формальные ограничения. Примером являются *Recurrent Neural Network Grammars (RNNG)* – нейронные генеративные модели, которые включают в себя явную фразовую структуру генерируемого предложения. Duerg и соавторы показали, что RNNG-архитектура, в которой грамматические операции «*shift-reduce*» интегрированы в рекуррентный генератор, даёт лучшие результаты парсинга и языкового моделирования, чем простые последовательные RNN [11]. Это демонстрирует, что жёсткая грамматика, заложенная в структуру модели, усиливает её способности по сравнению с нейросетями, которые «читают» текст как линейную цепочку.

Сентимент-анализ, вопрос-ответ, другие задачи. Инъекция синтаксиса показана эффективной и в задачах классификации и QA. К примеру, в задаче извлечения событий и отношений исследования выявляют, что деревья зависимостей позволяют находить связи между отдалёнными сущностями (например, причины и следствия в медицинских текстах) [12]. В анализе

тональности также показаны выгоды от учёта структуры (например, модели с графовыми слоями на основе дерева зависимостей дают лучшие результаты на тональных данных с сарказмом). Фактически синтаксические признаки часто позволяют обойтись с менее параметризованными моделями, усиливая их лингвистический «фон».

### **Интеграция семантики и семантических ролей**

Семантические роли (SMR, тематические роли, FrameNet, PropBank) – формализованные признаки «кто делает что, с кем, когда, как». Нейросети также пытаются учитывать их. В классическом подходе семантические роли выводились после или вместе с синтаксическим анализом (либо через специальные SRL-системы, либо через правила). В нейронной парадигме часто ставят вопросы: нужны ли явные семантические представления или модели их «сами находят»?

Семантическая ролевая разметка (SRL). Это самостоятельная задача, но часто она сочетается с синтаксисом. Например, в Coling показано, что совместное использование разнообразных синтаксических представлений (гетерогенных деревьев) и их кодирование через GCN улучшает нейронный SRL [13]. Авторы экспериментировали с разными типами деревьев (зависимости разной теории) и показали, что их агрегация приносит значимое улучшение метрики. Это подчёркивает, что семантика ролей тесно связана с грамматикой.

С другой стороны, некоторые современные подходы к SRL вовсе не требуют парсера. Например, недавно предложен энд-то-энд SRL-сегментатор, который обходит синтаксис, но это скорее исключение. Тем не менее, даже в полностью нейросетевых SRL-архитектурах добавление структурной информации оправдано: Xia и другие экспериментально подтвердили, что «Syntax-aware» модели SRL (с явными признаками синтаксиса) превосходят синтаксически-независимые. На CoNLL-2005 их лучшая модель достигла  $F1 \approx 85.6$ , опередив сильные безсинтаксические baselines [14]. Авторы делают вывод, что дополнительная синтаксическая информация всё ещё полезна, даже несмотря на мощные контекстные эмбединги (например, ELMo).

Рассмотрим фреймовые семантики и семантические графы. Помимо SRL, формальные семантические представления также начинают использоваться. Например, выделение смысловых рамок (FrameNet) или графов смысла (AMR) может быть внедрено в нейросети. Недавние работы рассматривают привязку фактов из KB или семантических графов к нейронной модели через специальные слои/слои внимания. Опыт показывает, что модели, обученные одновременно решать задачи SRL/парсинга и цели основного NLP-задания, чаще дают более интерпретируемые результаты. Связывание ролей и грамматики также помогает при извлечении информации: например, метод TakeFive для построения онтологий в клинических текстах сначала делает глубокий синтаксический и семантический анализ (включая SRL), а потом конвертирует всё в факты (парадигма семантически ориентированного извлечения).

Выводы по семантике: в совокупности результаты работ показывают, что формальные семантические признаки могут служить дополнительными

ориентирами при обучении нейросетей. Особенно это актуально в низко ресурсных условиях или при необходимости строгой семантической точности (например, в юридических или медицинских текстах, где роли аргументов критичны).

### **Применение гибридных подходов**

В разных приложениях интеграция формальных знаний даёт разные преимущества:

- **Машинный перевод:** как уже отмечено, добавление синтаксических и морфологических признаков преимущественно улучшает перевод в условиях «низких ресурсов» и при переводе неродных доменов [8, 9]. Впрочем, при наличии огромных обобщенных корпусов выигрыш может быть мал. Автор (в исследовании Noe Casas) подчёркивает, что в классических правилах переводимости прослеживаются явные лингвистические трансформации, которые «нынешние черные ящики» НМТ не объясняют, хотя статистически обучаются на соответствиях. Поэтому лингвистические особенности служат «мостом» между данными и человеческой интерпретацией.

- **Синтаксический анализ:** гибридные модели (RNNG, Tree-LSTM и др.) показывают, что знание грамматики может быть заложено непосредственно в нейросеть. Это ведёт не только к повышению точности парсера (как в RNNG [11]), но и к более структурированным представлениям языка, которые полезны далее.

- **Семантический анализ и извлечение информации:** семантические роли и другие метки улучшают идентификацию отношений. Например, в задаче извлечения причинно-следственных пар было показано, что добавление информации о синтаксической структуре значительно повышает качество (модель DepBERT учит встроить дерево зависимостей прямо в Transformer). Отчётность показывает, что «объединение семантики и синтаксиса повышает точность распознавания отношений» [12].

- **Вопросно-ответные системы и QA:** при сложных запросах (особенно требующих переносов значений между частями предложения) важна структурированная семантика. Подходы с участием AMR или фрейм-семантики позволяют моделям выполнять многоходовые выводы и внятно обрабатывать предикаты и аргументы. Исследования с GNN указывают на заметные улучшения в много переходном QA и задачах здравого смысла [7].

Таким образом, гибридные модели проявляют силу главным образом в тех задачах и сценариях, где чисто статистические методы наталкиваются на ограничения: малый объём обучающей выборки, богатая морфология, необходимость формальной интерпретируемости или соблюдения логических связей. Во многих практических приложениях – от медицинского извлечения до диалоговых систем – добавление явных лингвистических знаний повышает надёжность и объяснимость результатов.

## Заключение

Современная практика ОЕЯ постепенно возвращает актуальность лингвистических формализмов, объединяя их с мощью глубоких нейросетей. Как показано в статье, формальная информация – будь то синтаксическое дерево или семантическая ролевая разметка – может быть интегрирована в модели разными способами: через дополнительные слои (graph/LSTM), многоуровневое обучение (multitask с парсингом), расширенные словари (лексические факторы), или как внешние графовые структуры (AMR, фреймы). Это приводит к лучшим результатам там, где требуется учесть грамматическую структуру (SRL, парсинг, IE), и даёт преимущество в условиях ограниченных данных [4, 9]. В высокоресурсных задачах эффект может быть менее заметен, однако семантическая интерпретируемость систем благодаря лингвистическим знаниям только растёт.

В дальнейшем важно развивать методы ещё более глубокого слияния формальных моделей и нейросетей. Например, архитектуры, умеющие одновременно работать со структурными графами (деревьями зависимостей или грамматиками) и с широким контекстом текста, представляют собой перспективное направление. Также необходимо изучать, как хорошо синтаксические и семантические знания переносятся между языками – уже есть попытки многоязычного обучения с учётом структур [7]. Наконец, адаптивные гибридные подходы, которые регулируют вклад правил и статистики в зависимости от домена и объёма данных, представляются многообещающими. Поддерживая интеграцию лингвистической информации в нейронные модели, можно одновременно повысить точность и интерпретируемость систем ОЕЯ, что соответствует призыву к «глубокому включению знаний» в языковые технологии.

## Список использованных источников и литературы:

1. Кулмизев А., Нивре Дж. Дерево Шрёдингера – О синтаксисе и моделях нейронного языка // Границы искусственного интеллекта. – 2022. – Том. 5, статья 796788. – DOI: 10.3389/frai.2022.796788. – URL: <https://arxiv.org/abs/2110.08887> (дата обращения: 13.12.2025).
2. Касас Мансанарес Н. Внедрение лингвистических знаний в модели нейронной генерации текста – Диссертация 20/21 (PhD) / Ноэ Касас Мансанарес; Политехнический университет Каталонии, факультет теории Сеньяла и коммуникаций. – Барселона, Испания, 14 декабря 2020 г. – 133 с. – URL: <https://mt.cs.upc.edu/2021/10/11/injection-of-linguistic-knowledge-into-neural-text-generation-models-thesis-20-21-phd/> (дата обращения: 10.12.2025).
3. Ши Т., Малиутов И., Ирсой О. Семантическая маркировка ролей как синтаксический анализ зависимостей // Труды конференции по эмпирическим методам обработки естественного языка (EMNLP) 2020 года. – 2020. – Pp. 7551–7571. – DOI: 10.18653/v1/2020.emnlp-main.610. – URL: <https://aclanthology.org/2020.emnlp-main.610.pdf> (дата обращения: 16.12.2025).

4. Марчеджиани Д., Титов И. Кодирование предложений с помощью графовых сверточных сетей для семантической маркировки ролей // Труды 55-го ежегодного собрания Ассоциации вычислительной лингвистики (ACL). – 2017. – Рр. 1506–1515. – DOI: 10.18653/v1/P17-1138. – URL: <https://aclanthology.org/D17-1159/> (дата обращения: 13.12.2025).
5. Клеенанканди Дж., Абдул Назир К. А. Усовершенствованная архитектура Tree-LSTM для семантического моделирования предложений с использованием типизированных зависимостей. // arXiv.org: preprint. – 2020. – 23 с. – arXiv:2002.07775 [cs.CL]. – URL: <https://arxiv.org/pdf/2002.07775> (дата обращения: 14.12.2025)
6. Фиораванти С., Зугарини А., Джаннини Ф., Ригутини Л., Маггини М., Дилигенти М. Введение лингвистических функций для эффективной обработки естественного языка // Материалы Международной совместной конференции по нейронным сетям 2023 года (IJCNN) – Голд-Кост, Австралия, 18–23 июня 2023 г. – С. 1–7. – DOI: 10.1109/IJCNN54540.2023.10191680. – URL: <https://doi.org/10.1109/IJCNN54540.2023.10191680> (дата обращения: 20.12.2025).
7. Ли Ч., Ке Ч., Чжао П. Внедрение структурированных знаний в большие языковые модели с помощью графовых нейронных сетей // Труды 1-го совместного семинара по большим языковым моделям и структурному моделированию (XLLM 2025). – Вена, Австрия, 2025. – С. 16–25. – Ассоциация вычислительной лингвистики. – DOI: 10.18653/v1/2025.xllm-1.3. – URL: <https://aclanthology.org/2025.xllm-1.3.pdf> (дата обращения: 17.12.2025).
8. Касас Н., Коста-Хусса М. Р., Фоноллоса Х. А. Р., Алонсо Х. А., Фанло Р. Лингвистические словари, основанные на знаниях, для нейронного машинного перевода // Инженерия естественного языка. – 2020. – Том. 27, № 4, С. 485–506. – DOI: 10.1017/S1351324920000364. – URL: [https://upcommons.upc.edu/bitstream/handle/2117/330835/2020nle\\_linguistic\\_vocab\\_s.pdf](https://upcommons.upc.edu/bitstream/handle/2117/330835/2020nle_linguistic_vocab_s.pdf) (дата обращения: 15.12.2025).
9. Касас Н., Коста-Жусса М. Р., Фоноллоса Ж. А. Р. Разреженно факторизованный нейронный машинный перевод // arXiv.org: препринт. – 2021. – arXiv:2102.08934 [cs.CL]. – URL: <https://arxiv.org/html/2102.08934> (дата обращения: 18.12.2025).
10. Сонг Л. Решение графических задач НЛП с помощью рекуррентных сетей на графах // arXiv.org: препринт. – 2019. – 148 с. – arXiv:1907.06142 [cs.CL]. – URL: <https://arxiv.org/pdf/1907.06142> (дата обращения: 12.12.2025).
11. Дайер К., Канкоро А., Баллестерос М., Смит Н. А. Рекуррентные грамматики нейронных сетей // Труды NAACL-HLT. – 2016. – С. 199–209. – DOI: 10.18653/v1/N16-1024. – URL: <https://arxiv.org/abs/1602.07776> (дата обращения: 22.12.2025).
12. Кабир М. А., Джахин С. М. А., Аль Хасан М. Извлечение причинно-следственных пар из предложения с помощью модели преобразователя с учетом зависимостей // arXiv.org: препринт. – 2025. – arXiv:2507.09925 [cs.LG]. – URL: <https://arxiv.org/html/2507.09925v1> (дата обращения: 11.12.2025).
13. Ся Ц., Ван Р., Ли Ч., Чжан Ю., Чжан М. Семантическая маркировка ролей с использованием гетерогенных синтаксических знаний // Труды 28-й

Международной конференции по вычислительной лингвистике (COLING 2020). – Барселона, Испания (онлайн), 8–13 декабря 2020 г. – С. 2979–2990. – Международный комитет по вычислительной лингвистике. – DOI: 10.18653/v1/2020.coling-main.266. – URL: <https://aclanthology.org/2020.coling-main.266/> (дата обращения: 17.12.2025).

14. Ся Ц., Ли З., Чжан М., Чжан М., Фу Г., Ван Р., Си Л. Синтаксически-ориентированная нейронная семантическая ролевая маркировка // arXiv.org: препринт. – 2019. – arXiv:1907.09312 [cs.CL]. – URL: <https://arxiv.org/abs/1907.09312> (дата обращения: 18.12.2025).

### List of references

1. Kulmizev A., Nivre J. Schrödinger’s tree–On syntax and neural language models // *Frontiers in Artificial Intelligence*. – 2022. – Vol. 5, Article 796788. – DOI: 10.3389/frai.2022.796788. – URL: <https://arxiv.org/abs/2110.08887> (дата обращения: 13.12.2025).

2. Casas Manzanares N. Injection of linguistic knowledge into neural text generation models – Thesis 20/21 (PhD) / Noé Casas Manzanares; Universitat Politècnica de Catalunya, Departament de Teoria del Senyal i Comunicacions. – Barcelona, Spain, 14 Dec. 2020. – 133 с. – URL: <https://mt.cs.upc.edu/2021/10/11/injection-of-linguistic-knowledge-into-neural-text-generation-models-thesis-20-21-phd/> (дата обращения: 10.12.2025).

3. Shi T., Malioutov I., Irsoy O. Semantic role labeling as syntactic dependency parsing // *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. – 2020. – Pp. 7551–7571. – DOI: 10.18653/v1/2020.emnlp-main.610. – URL: <https://aclanthology.org/2020.emnlp-main.610.pdf> (дата обращения: 16.12.2025).

4. Marcheggiani D., Titov I. Encoding sentences with graph convolutional networks for semantic role labeling // *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL)*. – 2017. – Pp. 1506–1515. – DOI: 10.18653/v1/P17-1138. – URL: <https://aclanthology.org/D17-1159/> (дата обращения: 13.12.2025).

5. Kleenankandy J., Abdul Nazeer K. A. An enhanced Tree-LSTM architecture for sentence semantic modeling using typed dependencies // arXiv.org: препринт. – 2020. – 23 p. – arXiv:2002.07775 [cs.CL]. – URL: <https://arxiv.org/pdf/2002.07775> (дата обращения: 14.12.2025)

6. Fioravanti S., Zugarini A., Giannini F., Rigutini L., Maggini M., Diligenti M. Linguistic Feature Injection for Efficient Natural Language Processing // *Proceedings of the 2023 International Joint Conference on Neural Networks (IJCNN) – Gold Coast, Australia, June 18–23, 2023*. – P. 1–7. – DOI: 10.1109/IJCNN54540.2023.10191680. – URL: <https://doi.org/10.1109/IJCNN54540.2023.10191680> (дата обращения: 20.12.2025).

7. Li Z., Ke Z., Zhao P. Injecting Structured Knowledge into LLMs via Graph Neural Networks // *Proceedings of the 1st Joint Workshop on Large Language Models and Structure Modeling (XLLM 2025)*. – Vienna, Austria, 2025. – Pp. 16–25. –

Association for Computational Linguistics. – DOI: 10.18653/v1/2025.xllm-1.3. – URL: <https://aclanthology.org/2025.xllm-1.3.pdf> (дата обращения: 17.12.2025).

8. Casas N., Costa-jussà M. R., Fonollosa J. A. R., Alonso J. A., Fanlo R. Linguistic knowledge-based vocabularies for neural machine translation // *Natural Language Engineering*. – 2020. – Vol. 27, № 4, Pp. 485–506. – DOI: 10.1017/S1351324920000364. – URL: [https://upcommons.upc.edu/bitstream/handle/2117/330835/2020nle\\_linguistic\\_vocab\\_s.pdf](https://upcommons.upc.edu/bitstream/handle/2117/330835/2020nle_linguistic_vocab_s.pdf) (дата обращения: 15.12.2025).

9. Casas N., Costa-jussà M. R., Fonollosa J. A. R. Sparsely factored neural machine translation // *arXiv.org*: препринт. – 2021. – arXiv:2102.08934 [cs.CL]. – URL: <https://arxiv.labs.arxiv.org/html/2102.08934> (дата обращения: 18.12.2025).

10. Song L. Tackling graphical NLP problems with graph recurrent networks // *arXiv.org*: препринт. – 2019. – 148 с. – arXiv:1907.06142 [cs.CL]. – URL: <https://arxiv.org/pdf/1907.06142> (дата обращения: 12.12.2025).

11. Dyer C., Kuncoro A., Ballesteros M., Smith N. A. Recurrent neural network grammars // *Proceedings of NAACL-HLT*. – 2016. – Pp. 199–209. – DOI: 10.18653/v1/N16-1024. – URL: <https://arxiv.org/abs/1602.07776> (дата обращения: 22.12.2025).

12. Kabir M. A., Jahin S. M. A., Al Hasan M. Extracting cause-effect pairs from a sentence with a dependency-aware transformer model // *arXiv.org*: препринт. – 2025. – arXiv:2507.09925 [cs.LG]. – URL: <https://arxiv.org/html/2507.09925v1> (дата обращения: 11.12.2025).

13. Xia Q., Wang R., Li Z., Zhang Y., Zhang M. Semantic role labeling with heterogeneous syntactic knowledge // *Proceedings of the 28th International Conference on Computational Linguistics (COLING 2020)*. – Barcelona, Spain (Online), 8–13 Dec. 2020. – P. 2979–2990. – International Committee on Computational Linguistics. – DOI: 10.18653/v1/2020.coling-main.266. – URL: <https://aclanthology.org/2020.coling-main.266/> (дата обращения: 17.12.2025).

14. Xia Q., Li Z., Zhang M., Zhang M., Fu G., Wang R., Si L. Syntax-aware neural semantic role labeling // *arXiv.org*: препринт. – 2019. – arXiv:1907.09312 [cs.CL]. – URL: <https://arxiv.org/abs/1907.09312> (дата обращения: 18.12.2025).