

Актуальність розвідки – у винятковій гетерогенності категорії комічного, що становить постійний інтерес. Дослідження охоплює аналіз трьох основних комічних форм – гумору, іронії, сарказму, пародія ж визначає структурні та сюжетні особливості комічно-сатиричних продовжень першого роману, в якому переважає гумор і м'яка іронія. Комічні акти проаналізовано на рівнях семантики (метафор, метонімії, переважно гумористичних гіпербол, іронічних епітетів, алюзій, парадоксів), на рівні дискурсу (коли умовності, що регулюють використання мови, обігруються один проти одного, створюючи комічний ефект), регістру (зміщення в стилі) і прагматики (коли розмовні максими порушуються, звідки й комічний ефект). Такий підхід до аналізу комічного потрібний, коли комічні акти зумовлені, зокрема, позамовними факторами.

Ключові слова: комічне, гумор, іронія, сарказм, семантика, дискурс, прагматика.

Ладницькая Олеся, Кожемяко Наталия. Комическое в «Университетской трилогии» Дэвида Лоджа: языковые средства и внеязыковые факторы. Стаття посвящена аналізу комического в художественной литературе, аналізу языковых средств и внеязыковых факторов, ответственных за комический эффект в изучаемой трилогии. Актуальность работы – в исключительной неоднородности категории комического, способствующей постоянным ее изучением. Исследование основано на изучении трех основных комических форм – юмора, иронии, сарказма, пародия же определяет структурные и сюжетные особенности комическо-сатирических продолжений первого романа, где преобладает юмор и мягкая ирония. Комические акты подвергаются анализу на уровнях семантики (метафор, метонимии, в основном юмористических гипербол, ироничных эпитетов, аллюзий и парадоксов), на уровне дискурса (когда условности, регулирующие использование языка обыгрываются друг против друга, создавая комический эффект), регистра (сдвиги в стиле) и прагматики (когда разговорные максимы нарушаются, создавая комический эффект). Такой подход к анализу комического необходим в силу действия внеязыковых факторов.

Ключевые слова: комическое, юмор, ирония, сарказм, семантика, дискурс, прагматика.

Стаття надійшла до редколегії
07.04.2015 р.

УДК 81'322

Маргарита Лангенбах

Автоматичний синтаксичний аналіз речення за принципами граматики залежностей

Стаття висвітлює питання організації автоматичного синтаксичного аналізу речень. Актуальність обраної теми визначається її належністю до класу завдань комп'ютерного моделювання мови однієї з ключових проблем у сфері інформаційних технологій. Уперше наведено алгоритм автоматичного моделювання структури речення в термінах граматики залежностей, побудований на основі інформації про його безпосередні складники. Описано хід виконання основних завдань: формалізацію правил установалення зв'язків у реченні та їх автоматизацію; з'ясування основних переваг та недоліків обраної формальної моделі. Схарактеризовано структуру й основні етапи роботи аналізатора. Теоретичне значення дослідження полягає в накопиченні бази синтаксичних моделей речень. Практичне застосування отриманих результатів передбачається в системах комп'ютерного опрацювання мови, у навчально-методичних розробках тощо.

Ключові слова: граматика залежностей, автоматичний синтаксичний аналіз, автоматичне опрацювання тексту, лінгвістичний алгоритм, дерево залежностей, універсальна граматика.

Постановка наукової проблеми та її значення. Моделювання мовної системи, відтворення механізмів її функціонування – одне з пріоритетних завдань і теоретичного, і прикладного мовознавства. Особливої ваги ця проблема набула впродовж останніх десятиліть через активний розвиток комп'ютерних технологій, що спричинив, з одного боку, широку автоматизацію робочих процесів, з іншого – виникнення інформаційного суспільства, базисом якого є постійний та інтенсивний обмін відомостями. Це, своєю чергою, сформувало очевидну потребу створення комп'ютерних моделей мови, що уможливили б автоматичне опрацювання великих масивів інформації та забезпечили

ефективну взаємодію комп'ютера з людиною. Таке завдання передбачає, передовсім, формалізацію опису мовної системи. У цьому аспекті важливе місце посідає питання відтворення структури синтаксичних зв'язків між мовними одиницями, що й зумовлює **актуальність** дослідження.

Аналіз досліджень цієї проблеми. Як відомо, основними синтаксичними одиницями є слово, словосполучення, речення. Усі вони взаємопов'язані, оскільки слова як елементарні синтаксичні елементи, вступаючи у зв'язки, формують словосполучення й речення.

На сьогодні двома найпопулярнішими підходами до формалізації синтаксису та побудови автоматичного синтаксичного аналізу є граматики безпосередніх складників і граматики залежностей. Перша відтворює речення через схему бінарних зв'язків (між словами, словами та словосполученнями тощо), друга передбачає побудову ієрархічної структури із чітко визначеним центром і системою підпорядкувань. Отже, якщо під час першого підходу акцентується увага на існуванні в межах речення менших синтаксичних комплексів, то при другому речення розглядається як єдине ціле з мережею розгалужених міжслівних зв'язків.

Синтаксичний модуль граматики АГАТ, який розробила лабораторія комп'ютерної лінгвістики Інституту філології Київського національного університету імені Тараса Шевченка, організовано у формі двоетапного опрацювання речень, що поєднує обидві методики: перший етап працює на основі граматики безпосередніх складників і передбачає встановлення зв'язків між словами в межах словосполучень, другий завершує аналіз речення через побудову його цілісної схеми у вигляді дерева залежностей. На сьогодні перший етап повністю реалізовано (принципи та результати його роботи описано в роботах Н. П. Дарчук [1; 2]), другий перебуває на стадії розроблення і тестування.

Мета статті – проаналізувати основні засади використання граматики залежностей для побудови автоматичного синтаксичного аналізатора речень. Поставлена мета передбачала виконання таких завдань:

- формалізацію основних правил установаження зв'язків у реченні за допомогою означеної методики;
- автоматизацію побудови цілісної синтаксичної структури речення у вигляді дерева залежностей;
- з'ясування основних переваг і недоліків використання такої формальної моделі, аналіз її ефективності, виявлення типових помилок та окреслення шляхів їх усунення.

Об'єктом дослідження були прості та складні речення; **предметом** – їхня синтаксична структура, описана в термінах граматики залежностей.

Матеріалом слугували тексти Електронного корпусу української мови, опрацьовані системою автоматичного граматичного аналізу АГАТ.

Новизна роботи полягає в тому, що вперше наведено алгоритм реалізації автоматичної побудови синтаксичної схеми речення на основі інформації про його безпосередні складники.

Теоретичне значення дослідження полягає в накопиченні бази синтаксичних моделей речень, а також у виявленні типових і нетипових реченневих конструкцій. **Практичне застосування** отриманих результатів належить до галузі комп'ютерного опрацювання мови та мовлення (системи автоматичного аналізу й синтезу текстів, машинного перекладу, діалогові та інформаційно-пошукові системи тощо). Окрім того, можливе використання накопиченої бази синтаксичних конструкцій як методичних матеріалів до навчання мов.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. Розглянемо принципи моделювання структури речень на базі граматики залежностей. За І. П. Севбо, основні постулати цієї синтаксичної теорії такі:

- у реченні є один головний (вершинний) член – присудок;
- речення будується з одиниць одного таксономічного рівня;
- усі слова в реченні пов'язані між собою зв'язком підпорядкування, що є узагальненням традиційних типів синтаксичних зв'язків;
- ієрархічна структура залежностей між словами речення зображується графічно за допомогою дужок, стрілок або у вигляді дерева залежностей [3, с. 8–9].

Ці правила було покладено в основу аналізу структури речень у системі АГАТ. Для графічного представлення моделі обрано дерево залежностей: різновид математичного графа, схему, що складається з розташованих на різних рівнях вузлів (якими в нашому випадку слугують слова) і ребер (синтаксичних зв'язків між ними). Така модель, на нашу думку, найкраще передає ієрархічність зв'язків у реченні та логіку побудови висловлювання.

Алгоритм моделювання дерева залежностей ґрунтується на двох групах правил:

- 1) правила, які передбачають використання інформації з попереднього етапу синтаксичного аналізу;
- 2) доповнювальні правила, які охоплюють ситуації, що виходять за межі синтаксису словосполучень (визначення головних членів речення у разі відсутності в ньому предикативної пари (підмета з присудком), установлення зв'язку між частинами складного речення тощо).

Перша група правил забезпечує об'єднання всіх побудованих на попередньому етапі словосполучень у ланцюжки підпорядкувань. Словосполучення розглядаються як зв'язані, якщо в них є спільний член. Така ситуація може мати два варіанти:

- 1) спільний член є головним компонентом для одного словосполучення і залежним для іншого (так званий «ланцюжок» підпорядкувань);
- 2) спільний член є головним в обох словосполученнях («гілкування»);

У першому випадку слова формують ланцюжок залежностей, у другому відбувається розгалуження на окремі ланцюжки, об'єднані спільною вершиною. Оскільки граматики залежностей передбачає, що слово може мати кілька залежних, але лише одне керівне, у системі неможливі ситуації, коли словосполучення мають спільний залежний член.

Ці правила лягли в основу алгоритму побудови дерева залежностей (див. рис. 1):

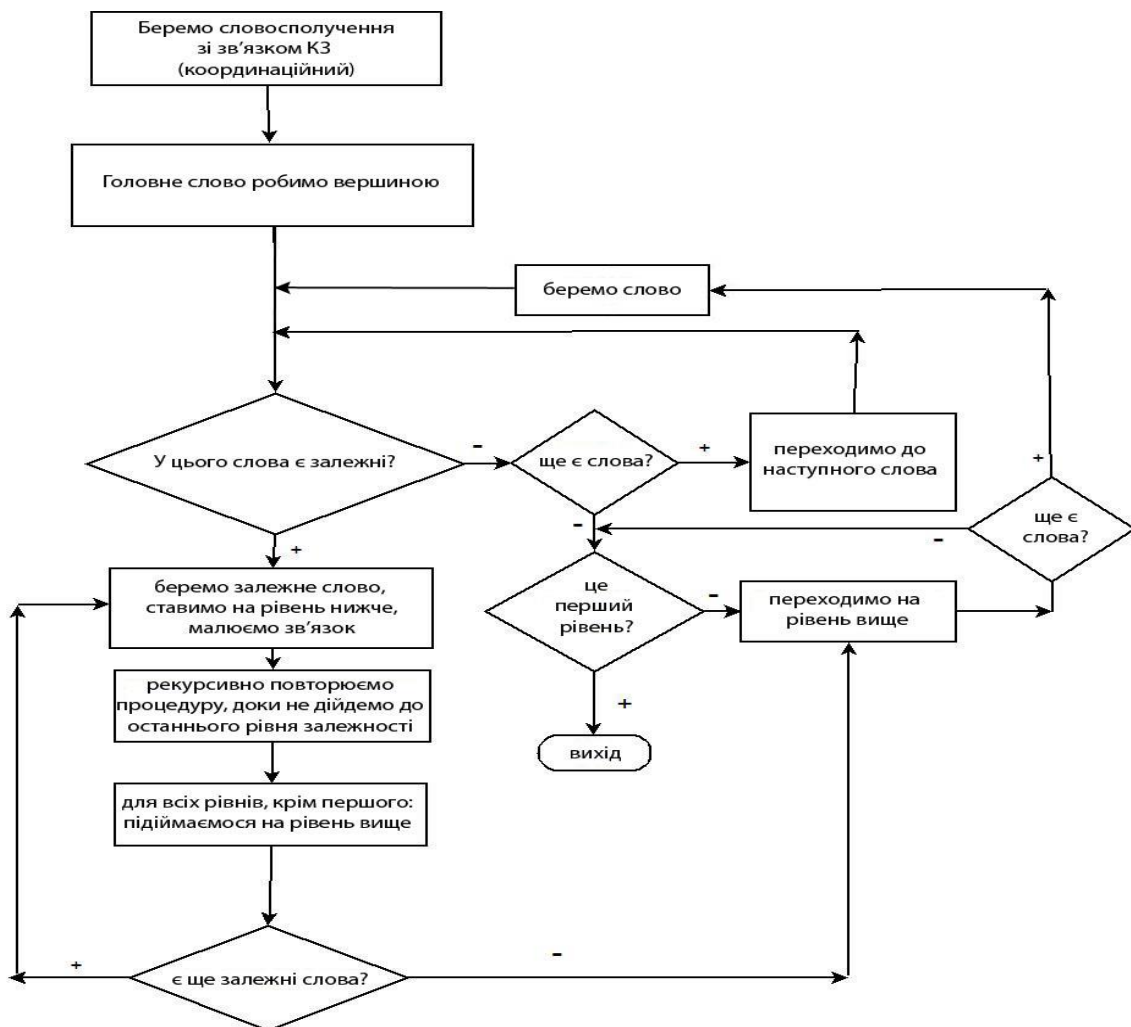


Рис. 1. Алгоритм побудови дерева залежностей

Через рекурсивне проходження за усіма ланцюжками залежностей, які є в реченні, починаючи з граматичного центра, будується його схема (див. рис. 2). Граматичним центром вважається дієслово-член предикативного словосполучення (пари «підмет–присудок»).

2: Одномовний тлумачний словник за самим своїм типом становить єдність протилежностей .



Рис. 2. Схема речення (дерево залежностей) у системі АГАТ

Проте в мовленні досить поширені речення, які не мають повної граматичної основи, отож, у них відсутнє й предикативне словосполучення, за яким система знаходить дієслово-присудок. Оскільки ж за алгоритмом необхідною є наявність вершинного компонента, створена група доповнювальних правил, які охоплюють ситуації з неповними реченнями. У разі відсутності в реченні предикативної пари запускається додатковий блок алгоритму, спрямований на визначення слова, що може виконувати роль головного члена речення (див. рис. 3).

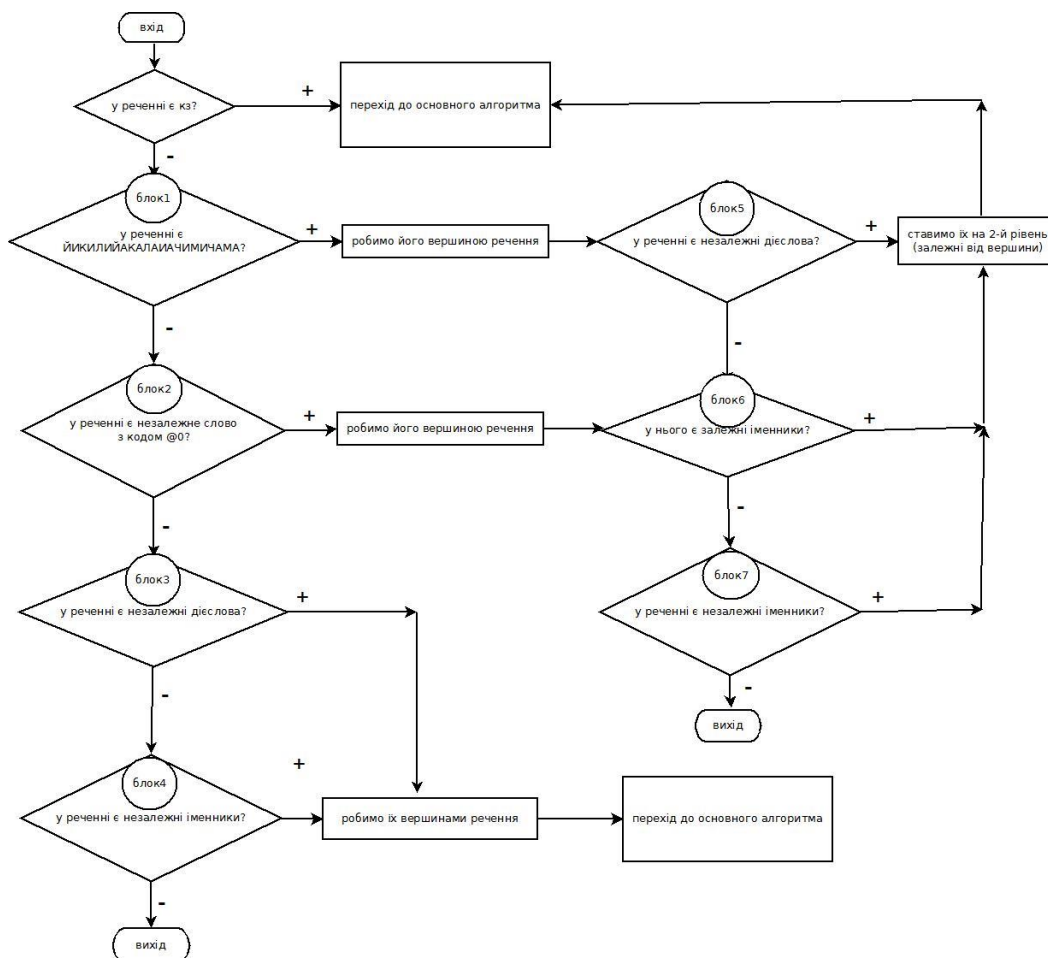


Рис. 3. Алгоритм визначення головного члена неповного речення

Побудова структури складного речення відбувається в такий спосіб:

- виявляються межі та граматичні центри кожної із частин речення, будуються їх схеми;
- з'ясовується, за допомогою яких засобів поєднані ці речення: сполучників (сполучних слів) чи розділових знаків;
- для сполучникового зв'язку виявляється тип – сурядний чи підрядний;
- у разі безсполучникового зв'язку вершиною визначається або сам розділовий знак (якщо це кома чи тире), або головний член речення, що закінчується розділовим знаком (якщо це двокрапка);
- сурядні сполучники, що з'єднують речення, устанавлюються їх спільною вершиною;
- якщо речення складнопідрядне, то вершиною визначається головний член основного речення, залежні ж приєднуються через поєднання цієї вершини або зі сполучником, або з присудком залежної частини (для речень, у яких замість сполучника виступає сполучне слово) [3, с. 61–63].

Коректність роботи алгоритму і ступінь «покриття» ним аналізованого матеріалу залежить від двох основних чинників: якості результатів роботи попереднього рівня аналізу (виявлення словосполучень) і складності структури речення та нормативності/типовості його конструкції. Перший чинник представлено такими проблемами:

1. Пропуск зв'язків. Причинами можуть бути:

– *непрямий порядок слів (більшість моделей керування передбачає розгортання конструкції в реченні зліва направо, тому в ситуаціях типу **Орієнтиром має бути** передусім відповідність нової лексики новим поняттям система може пропустити зв'язок присудка з додатком);*

– відсутність слова або моделі керування в словнику;

– неправильні підпорядкування, спричинені порядком слів або граматичною омонімією, зокрема омонімією граматичних форм і синтаксичних конструкцій. Наприклад, у реченні *Попри розуміння сутності проблеми, розв'язання її досі немає* іменники *проблеми* та *розв'язання* мають однаковий відмінок, тому система може кваліфікувати їх як однорідні члени; у реченні *Офіційний курс гривні цього року перебуватиме в коридорі, заявленому Нацбанком* іменник *року* система визначить як залежне від слова *гривні* за моделлю іменникового керування «гривня + іменник у Р. в.» (пор. *За результатами ревізії документів було нараховано 21 223 гривні несплаченого боргу*).

*Другий чинник переважно призводить до нездатності системи виявити в реченні головне слово. Найчастіше така ситуація виникає в неповних реченнях (Ми схильні виправдовувати свою некоректну поведінку в певних ситуаціях. А от таку саму ситуацію, але з іншими людьми – часто ні) або реченнях, які містять частини мови в синтаксичній транспозиції (у реченні Антоніми «державний-недержавний» знайомі українській мові вже давно, але **показове** переосмислення останнього з них у сучасних суспільно-політичних реаліях України **прикметник** показове **перебирає** на себе присудкову функцію пропущеного дієслова-зв'язки (є показовим)).*

Усунення цих проблем і вдосконалення якості автоматичного аналізу можливе за умови застосування комплексу різноманітних підходів, зокрема розширення словника моделей керувань, доповнення інформації про позиційні варіанти реалізації цих моделей, ускладнення правил контекстного аналізу та покращення розпізнавання омонімії, використання методик імовірнісного аналізу тощо.

Висновки та перспективи подальшого дослідження. Перспективність розроблення модуля синтаксичного аналізу речення за деревом залежностей зумовлена, по-перше, тим, що цей рівень – базис для моделювання смислової структури речення (а отже слугуватиме основою для подальшого переходу до семантичного аналізу речень і текстів); по-друге, його універсальністю. Оскільки алгоритм безпосередньо не спирається на словникову інформацію, можливе його застосування до будь-якого мовного матеріалу, у тому числі й не українськомовного. Цьому також сприяє виведення результатів у вигляді дерева залежностей, яке також є універсальною схемою, незалежною від специфіки граматичних конструкцій певних мов. Ця властивість аналізатора має велике теоретичне та практичне значення: з одного боку, у такий спосіб викристалізовується набір граматичних правил, спільних для різних мов (фактично, ідеться про універсальну граматику, на важливості створення якої наголошував Н. Хомський [4, с. 11]); з іншого – подібні аналізатори можуть бути застосовані в системах інформаційного пошуку та машинного перекладу.

Отже, можна констатувати, що використання граматики залежностей для створення формалізованого опису синтаксичних структур є стратегією, перспективною і в аспекті теоретичних досліджень, і практичного застосування.

Джерела та література

1. Дарчук Н. П. Автоматичний синтаксичний аналіз текстів корпусу української мови / Н. П. Дарчук // Укр. мовознавство. – 2013. – Вип. 43. – С. 11–19.
2. Дарчук Н. П. Електронний словник мови Тараса Шевченка: методика і технології укладання / Н. П. Дарчук, М. О. Лангенбах // Укр. мовознавство. – 2014. – Вип. 1. – С. 106–115.
3. Севбо И. П. Графическое представление синтаксических структур и стилистическая диагностика / И. П. Севбо. – Київ : Наук. думка, 1981. – 192 с.
4. Хомский Н. Аспекты теории синтаксиса / Н. Хомский. – М. : Изд-во Моск. ун-та, 1972. – 129 с.

Лангенбах Маргарита. Автоматический синтаксический анализ предложения на основе грамматики зависимостей. Стаття описує принципи автоматизації синтаксического аналізу пропозицій. Вперше представлений алгоритм автоматического моделирования структуры предложения в терминах грамматики зависимостей, построенный на основе информации о его непосредственных составляющих. Описаны основные задания: формализация правил установления связей в предложении и автоматизации построения его схемы; определение основных преимуществ и недостатков использованной модели. Теоретическое значение исследования состоит в составлении базы синтаксических моделей предложений. Практическое использование полученных результатов предполагается в области компьютерной обработки языка, в учебно-методических разработках.

Ключевые слова: грамматика зависимостей, автоматический синтаксический анализ, автоматическая обработка текста, лингвистический алгоритм, дерево зависимостей, универсальная грамматика.

Langenbakh Margaryta. The Automatic Syntactic Sentence Analysis Based on the Dependency Grammar. The article highlights the principles of natural language sentences parsing. The importance of the chosen topic is determined by its belonging to a class of problems of automatic language processing, that appears to be the one of the urgent needs at information technologies. The aim of the research was to analyze the principles and prospects of the dependency grammar for the automatic syntactic text processing. The sentences were chosen as the object of the study, their syntactic structure was the subject of research. The textual materials were taken from the Ukrainian Language Corpus. The paper contains the algorithm of automatic dependency tree building using the information about sentence constituents, that makes the novelty of the research. Such tasks were performed: the formalization of the basic linkage rules in the sentence; the automation of the sentences parsing; the analysis of the main advantages and disadvantages of chosen model. The theoretical value of the study is determined as the accumulation of syntactic model sentence structures. The practical use of the results is supposed to lie in the field of the computer language and speech processing, the methodic of the grammar study etc.

Key words: dependency grammar, automatic text processing, linguistic algorithm, dependency tree, universal grammar.

Стаття надійшла до редколегії
20.04.2015 р.

УДК 811.112.2 '37:373.43

Наталія Лисецька

До питання актуальних фразеологічних інновацій сучасної німецької мови

У статті досліджують фразеологічні інновації сучасної німецької мови, що виникли або набули додаткового значення у 2013–2014 рр., функціонують у медійному дискурсі та виступають засобами поповнення німецькомовного дискурсу, німецької мовної та концептуальної картин світу. Витлумачено поняття *неологізм*, *фразеологічна інновація*, *концепт*. Фразеологічні інновації розглянуто як неологізми в галузі фразеології, оказіональні утворення, продукти вторинної опосередкованої номінації, які охоплюють власне фразеологічні інновації (новизна форми і значення); новосемантизми (фразеологічні одиниці з новонабутою семантикою); модифіковані мовні форми; запозичення. Увага приділена таким фразеологічним інноваціям: *Lichtgrenze 2014 / Lichtgrenze zum 25. Jubiläum*, *Schwarze Null*, *Generation Kopf unten*, *Social Freezing*, *Freund hört mit*, *Big Data*. Проаналізовано специфіку фразеологічних інновацій, показано їх ситуативне функціонування в дискурсі, названо концепти, які вони активують. До таких ключових концептів належать: *Hoffnung*, *Gerechtigkeit* bzw. *Ungerechtigkeit*,

© Лисецька Н., 2015