

ХТО ГОВОРИТЬ: ЛЮДИНА АБО МАШИНА?

Морозова І. Б.

*доктор філологічних наук,
професор кафедри граматики англійської мови
Одеській національній університет імені І. І. Мечникова
вул. Дворянська, 2, Одеса, Україна
orcid.org/0000-0002-1905-7563
torpo@ukr.net*

Боцуляк Г. В.

*студентка факультету романо-германської філології
Одеській національній університет імені І. І. Мечникова
вул. Дворянська, 2, Одеса, Україна
orcid.org/0000-0001-5993-8317
annaboculak@gmail.com*

Ключові слова: *штучний інтелект, природне мовлення, англomовний діалог, граMATика залежностей, синтаксична структура, речення.*

Стаття присвячена розгляду особливостей синтезу мовлення штучним інтелектом (ШІ) в аспекті лінгвальних відмінностей від природної генерації мовлення людиною. Вибір теми дослідження мотивовано розвитком комп'ютерних технологій і сучасної прикладної лінгвістики, які пов'язані з генерацією мовлення людини, а саме: обробці текстів природною мовою, ПМ (Natural Language Processing, NLP) і галузі робототехніки, яка є втіленим ШІ (embodied AI). Основна мета дослідження полягає у спробі встановлення формальних комунікативних відмінностей між генерацією машинного і природного мовлення на прикладі англomовного діалогу моделі людина-робот (Л-Р). Попри величезний теоретичний і практичний внесок лінгвістів і фахівців галузі комп'ютерних технологій щодо генерації машинної мови (Джозеф Вайценбаум, Роджер Шенк, Уїльям Аарон Вудс, Террі Виноград, Н. Хомський), багато аспектів залишаються не вирішеними. До них відносяться питання морфологічної омонімії (Morphological Disambiguation) англійської мови, обмеження граматики залежностей у розпізнаванні й синтезі складних синтаксичних структур (складносурядні та складнопідрядні речення), обмеження когнітивної системи сучасних роботів розпізнавати абстрактні поняття й категоризувати так звані «для цього випадку» (ad-hoc) поняття, що створює людина залежно від життєвої ситуації, розуміти метафоричні висловлювання як одне з найважливіших лінгвістичних явищ, що відображає наше мислення. У більш широкому сенсі головною проблемою залишається створення роботів, які здатні передбачувати соціальну поведінку людини, розуміти її фізичний та емоційний стан завдяки синхронній системі «людина-робот» (на противагу лише фізичній симуляції поведінки машинами). Наукова новизна статті полягає в з'ясуванні лінгво-гносеологічних відмінностей між природним і машинним мовленням. У результаті проведеної роботи конкретизовано поняття «штучний інтелект», розглянуто основні етапи синтезу машинного мовлення, відокремлені базові відмінності генерування штучного та природного англomовного діалогу.

WHO IS SPEAKING: MAN OR MACHINE?

Morozova I. B.

*Doctor of Philological Sciences,
Professor at the Department of English Grammar
Odesa I. I. Mechnikov National University
Dvoryanska str., 2, Odesa, Ukraine
orcid.org/0000-0002-1905-7563
morpo@ukr.net*

Botsuliak G. V.

*Student at the Faculty of Romance and Germanic Philology
Odesa I. I. Mechnikov National University
Dvoryanska str., 2, Odesa, Ukraine
orcid.org/0000-0001-5993-8317
annaboculak@gmail.com*

Key words: *artificial intelligence, natural speech, English dialogue, dependency grammar, syntactic structure, sentence.*

The presented paper gives a detailed survey of peculiarities of the AI-powered speech synthesis in the aspect of their lingual differences from the natural speech generation. The chosen subject matter for investigation is motivated by the rapid development of Information Technologies and modern applied linguistics in the field of human speech generation, and namely – Natural Language Processing (NLP) versus robotics (embodied AI) speech production. The ultimate objective of our investigation consists in making an attempt to establish formal and communicative differences between machine and natural speech synthesis on the material of the English dialogue carried out after the “human-robot (H-R)” model. Despite the enormous theoretical and practical contribution of linguists and computer scientists to machine language generation (Joseph Weizenbaum, Roger Schenk, William Aaron Woods, Terry Vinograd, N. Chomsky), many aspects of this problem require further consideration. The latter include morphological disambiguation of the English language, dependency grammar limitations to recognize and synthesize lengthy syntactic structures (compound and complex sentences), restrictions in the cognitive system of modern robots to identify abstract concepts and determine the so-called ad-hoc categories that people can form on the spot, grounding on the situations they face, understanding metaphorical phrases which belong to one of the most important linguistic phenomena reflecting human natural way of thinking. In a broader sense, the remaining problem is building robots that will be capable to predict the social behaviour of people, understand their physical and emotional states due to synchronic system “human-robot” (not just the physical simulation). The scientific novelty of the article consists in figuring out lingual and epistemological differences between human and machine speech. Our investigation results in concretising the notion “artificial intelligence”, establishing the main stages of AI-powered speech synthesis and foregrounding the basic differences in generating the artificial and natural English dialogue.

Постановка проблеми. Мовлення – динамічне, системне, когнітивне та соціальне явище, тому, щоб зрозуміти й запрограмувати його з наукової точки зору, нам необхідна модель із сенсорно-моторною системою, яка може бути залучена в комунікативний процес з людиною, і це, безперечно, є робот [1, с. 722].

Відомо, що обробка текстів природною мовою, ПМ (Natural Language Processing, NLP) – це галузь інформатики, яка зосереджена на лінгвістичній взаємодії людини та машини. Більшість істориків відстежує першоджерела цієї галузі з початком комп’ютерної ери, коли Алан Тюрінг запропонував ідею, що машини здатні проявляти

інтелектуально обумовлену поведінку, що тотожна поведінці людини [2, с. 14–15]. У 1965 р. Джозеф Вайценбаум написав ELIZA – перший чат-бот, який працював за принципом простого синтаксичного аналізу речень і підбору ключових слів у фіксовані фрази [3, с. 104–105]. Надалі такі вчені, як Роджер Шенк, Уільям Аарон Вудс, Террі Виноград, установили, що мова машин прив'язана до поверхневого синтаксису англійської мови, і пояснили значення семантики в обчислюваних системах. Контекстно-вільна граматики (парадигматичний синтаксис, заснований на залученні певних структурних моделей), яка складається з відрізків тексту, зокрема слів, словосполучень, речень загалом, за Н. Хомським, лежить в основі сучасної обробки текстів системами штучного інтелекту (далі – ШІ) [2, с. 15–17].

Робот нового покоління – це втілений ШІ (embodied AI), розумний робот із можливостями машинного навчання, який не лише використовує велетенський об'єм даних, а й отримує їх із власного досвіду та попередньо засвоєної інформації [4, с. 2]. Такими є соціальні роботи або людиноподібні роботи-гуманоїди (андроїди), які зовнішньо схожі на людину та визначаються як «sociable», що означає «комунікативний, дружній, контактний» [5, с. 16–17]. Але, з іншого боку, поява таких роботів викликає неспокій у людства. Такі психологічні стани відомі як ефект «моторошної долини», «ефект ELIZA», стан когнітивного дисонансу, теорія психопатів тощо. Усі вони є наслідком несвідомості й непередбачуваності поведінки, відсутності емоційної реакції з боку людиноподібного роботу [6, с. 10–11].

Попри величезний теоретичний і практичний внесок лінгвістів і фахівців галузі комп'ютерних технологій щодо генерації машинної мови та робототехніки, оволодіння мовою та її розуміння програмами до сих пір пов'язані з широким колом проблем, що виникають при дослідженні ШІ. Сучасні програми дають змогу роботам розуміти значення речень, але міжособистісна комунікація пов'язана з непередбачуваністю й динамічністю мови [7, с. 759]. У більш широкому сенсі – це підтримання природного діалогу з людиною, проблеми візуальної та невербальної комунікації, уміння сприймати й виражати емоції, упізнавати моделі поведінки людей, установлювати та підтримувати соціальні відносини [8, с. 305].

Отже, постановка проблеми дослідження зумовлена лінгво-гносеологічною необхідністю заповнити цю прогалину, а також практичною потребою з'ясувати відмінності між природним та машинним мовленням.

Мета й завдання статті. Метою дослідження є спроба встановлення формальних комунікативних відмінностей між генерацією машинного і

природного мовлення. Поставлена мета мотивує низку завдань, головними з яких є такі: конкретизація поняття «штучний інтелект», огляд становлення штучного синтезування мовлення, поверхнево-структурний аналіз реплік у діалогах моделі людина-робот (далі – Л-Р), відокремлення базових відмінностей штучних і природних реплік в англомовному діалозі на матеріалі аналізу мовленнєвих зразків робота – андроїда Софії. Сьогодні існує близько двох десятків людиноподібних роботів, але ми обрали саме Софію, так як остання вважається найбільш досконалою моделлю. Вона є представником роботів-андроїдів інтелектуальної лінійки лабораторії Девіда Хедсона, поєднання технології розпізнання голосу Google Chrome Alphabet та іншого апаратного авторського інструментарію дає їй змогу розуміти мову [9].

Предмет та об'єкт досліджень. Об'єктом досліджень є англійське природне мовлення й мовлення, що генерується ШІ. Предметом досліджень є репліки англомовного діалогу в моделі людина-робот (Л-Р) на прикладі аналізу діалогових зразків робота-андроїда Софії.

Виклад основного матеріалу дослідження. Сьогодні існує величезна кількість визначень поняття «штучний інтелект», проте, відповідно до теми роботи, ми визначаємо це поняття як когнітивну здатність машин моделювати психічні процеси, які притаманні людині.

Складність формального опису природної мови та її обробки призводить до необхідності поділення цього процесу на окремі етапи, які відповідають рівням мови. Більшість сучасних лінгвістичних процесорів відносяться до модульного типу, у якому кожному етапу аналізу чи синтезу тексту відноситься окремий модуль процесора. На початку процесу відбувається графематичний аналіз, тобто виділення в тексті речень і словоформ – перехід від символів до слів. Морфологічний аналіз – це перехід від словоформ до їх лем (словарних форм лексем) або основ (ядерних частин слова). Для англійської мови цей процес включає вирішення морфологічної омонімії (Morphological Disambiguation) – усунення морфологічного багатозначення, так як у складі англійської мови 20 із 50 найбільш уживаних слів – омофони. Синтаксичний аналіз – виявлення синтаксичних зв'язків слів і граматичної структури речень. Граматичними моделями є дерева залежностей/підрядності й дерева складників. Надалі відбувається поєднання синтаксичної структури з динамічною інформацією («grounding»). Остання представлена лінгвістичними ресурсами (комп'ютерні словники, тезауруси, онтології, корпуси текстів, ресурси змішаного типу – FrameNet), що містять слова та речення з лінгвальними даними [3, с. 156–158; 10, с. 410–411].

Отже, для процесу обробки й синтезу природної мови характерна формальність (чіткий алгоритм) і функціональність, тобто відсутність моделі аналізу й синтезу саме мовлення людини, залежно від ситуації навколишнього середовища, що характерно для природного комунікативного процесу.

Наступним етапом цієї роботи був поверхнево-структурний аналіз реплік у діалогах робота Софії та відокремлення базових відмінностей штучних і природних реплік в англомовному діалозі. Загальна кількість зразків реплік становила 520 одиниць.

У процесі детального вивчення практичного матеріалу дослідження встановлено, що питома вага різних типів речень у діалогових репліках Софії становить: прості речення – 54%, ускладнені – 20,3%, складні – 25,7%, еліптичні – 0%. Аналіз базових відмінностей штучних і природних реплік проводився шляхом порівняння питомої ваги різних типів речень, що генеруються Софією, з кількісним розподілом різних типів речень, які характерні для діалогового мовлення персонажів як аналогу усного розмовного мовлення [11, с. 86].

У результаті порівняння мовлення робота-андроїда Софії з природним мовленням людей визначено, що в діалогових репліках Софії спостерігається порівняно з натуральним мовленням людини менша кількість простих речень і відсутність еліптичних речень. Питома вага простих речень у діалогах Софії – 54%; тоді як у мовленнєвому потоці людей вони становлять 56,5% відповідно. Питома вага еліптичних конструкцій у репліках Софії – 0%; зазначимо, що в природному англійському мовленні вони налічують близько 13,6%. Софія демонструє тяжіння до ускладнених англомовних речень, використовуючи їх у 20,3% порівняно з природним мовленням, де останні становлять у середньому не більше ніж 15,1%. Установлена закономірність указує на нездатність ШІ «домислювати» інформацію від співрозмовника, тобто з контексту комунікативної ситуації, що характерно для діалогічного мовлення людини. Складні речення в діалогах Софії менш чисельні порівняно з природним мовленням і становлять 25,7% і 28,4% відповідно. Засвідчені мовленнєві відмінності зумовлені обмеженнями здатності генерації складних синтаксичних структур на сучасному етапі розвитку ШІ. Репліки, які є складними граматичними структурами в діалогах Софії, являють собою визначення понять, доступних у загальновідомій електронній енциклопедії «Wikipedia». Інакше кажучи, з лінгвістичної точки зору система за допомогою граматики залежностей обирає ключові слова та поєднує її з інформацією, що вже існує в інтернет-ресурсах, або зі стійкими запрограмованими комунікативними шаблонами. Когнітивний потенціал робота спостерігається

лише в його здатності відносно нескладно трансформувати попереднє висловлювання за завантаженими формальними синтаксичними моделями, спираючись на тематично-ремагічні зв'язки словоформ, згідно з граматику залежностей (наприклад, можливість робота переробити попереднє твердження в загальне питання).

Висновки й перспективи подальших розробок. У результаті проведених досліджень зроблено такі висновки:

1. Обробка та синтез природного мовлення сучасними системами ШІ є формальним алгоритмом, в основі якого лежить визначення ключових слів речень за допомогою граматики залежностей і подальше поєднання їх з інформацією інтернет-ресурсів або заздалегідь запрограмованими комунікативними шаблонами.

2. Когнітивна здатність робота Софії є обмеженою; вона полягає в можливості відносно нескладно трансформувати попереднє висловлювання (наприклад, можливість переробити попереднє твердження в загальне питання). Формування реплік, пов'язаних із мисленням, тобто вміння актуально аналізувати синтаксис речення на обов'язкові для існування речення частину (рему) та тему – фон реми (необов'язкова частина) – відсутнє.

Отже, когнітивні здібності робота Софії не відповідають визначенню поняття «штучний інтелект» на сучасному рівні, тобто спостерігається симуляція наявності інтелекту.

Перспективи подальших розробок бачимо у вивченні комунікативних синтаксичних моделей робота, що імітують мисленнєві процеси свідомості людини, а також вплив навколишнього середовища або комунікативної ситуації на мовленнєву реакцію андроїда Софії.

ЛІТЕРАТУРА

1. Tangiuchi T., Mochihashi D., Nagai T. et al. Survey on frontiers of language and robotics. *Advanced Robotics*. 2019. Vol. 33. P. 700–730.
2. Близнюк Б.О. Современные методы обработки естественного языка. *Вісник Харківського національного університету імені В.Н. Каразіна*. 2017. № 36. С. 14–26.
3. Большакова Е.И. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика : учебное пособие. Москва : МИЭМ, 2011. 269 с.
4. Eidenmueller H. The Rise of Robots and the Law of Humans. *Oxford Legal Studies Research Paper*. 2017. Vol. 27. 15 p. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2941001.
5. Зильберман Н.Н. Функциональная классификация социальных роботов. *Гуманитарная информатика*. 2014. № 8. 31 с.

6. Зильберман Н.Н., Слободская А.В. Восприятие различных типов культурного интерфейса социального робота. *Universum*. 2014. № 10. С. 10–13.
7. Nakadai K., Takahashi T., Okuno HG. et al. Design and implementation of robot audition system 'hark' open source software for listening to three simultaneous speakers. *Advanced Robotics*. 2010. Vol. 24. P. 739–761. URL: <https://www.tandfonline.com/doi/abs/10.1163/016918610X493561>.
8. Зильберман Н.Н., Стефанцова М.А. Социальный робот: подходы к определению понятия. *Современные исследования социальных проблем*. 2016. № 11 (67). С. 305.
9. Sophia: веб-сайт. URL: <https://www.hansonrobotics.com/sophia/>.
10. Cynthia Matuszek, Evan Herbst, Luke Zettlemoyer and Dieter Fox. Learning to Parse Natural Language Commands to a Robot Control System. *Experimental Robotics*. 2013. Vol. 88. P. 403–415.
11. Морозова І.Б. Парадигматичний аналіз структури і семантики елементарних комунікативних одиниць у світлі гештальт-теорії в сучасній англійській мові : монографія. Одеса : Друкарський дім, 2009. 384 с.
4. Eidenmueller H. The Rise of Robots and the Law of Humans. (2017). *Oxford Legal Studies Research Paper*, 27, 15. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2941001
5. Zilberman N. N. (2014). Functional classification of social robots [Funktionalnaia klassifikatsiia sotsialnih robotov]. *Humanitarian Informatics*, 8, 31.
6. Zilberman N. N., Slobodskaya A. V. (2014). Perception of different types of social robot cultural interface [Vospriatie razlichnyh tipov kulturnogo interfeisa sotsialnogo robota]. *Universum*, 10, 10-13.
7. Nakadai K., Takahashi T., Okuno HG. et al. (2010). Design and implementation of robot audition system 'hark' open source software for listening to three simultaneous speakers. *Advanced Robotics*, 24, 739-761. Retrieved from <https://www.tandfonline.com/doi/abs/10.1163/016918610X493561>
8. Zilberman N. N., Stefantsova M.A. (2016). A social robot: approaches to defining the concept [Sotsialnii robot: podhody k opredele-niiu poniatii]. *Modern research on social problems*, 11(67), 305.
9. Sophia (2021). Retrieved from <https://www.hansonrobotics.com/sophia/>
10. Cynthia Matuszek, Evan Herbst, Luke Zettlemoyer and Dieter Fox. (2013). Learning to Parse Natural Language Commands to a Robot Control System. *Experimental Robotics*, 88, 403-415.
11. Morozova I. B. (2009). Paradigmatic analysis of the elementary communicative unit's structure and semantics in the light of the Gestalt-theory in modern English [Paradygmatychnyi analiz struktury i semantyky elementarnykh komunikatyvnykh ody-nyts u svitli geshtalt-teorii v suchasniy angliiskii movi: monografia]. Odesa: Drukarskyi dim Publ.

REFERENCES

1. Tangiuchi T., Mochihashi D., Nagai T. et al. (2019). Survey on frontiers of language and robotics. *Advanced Robotics*, 33, 700-730.
2. Blizniuk B. O. (2017). Modern methods of natural language processing [Sovremennye metody obrabotki estestvennogo iazika]. *The Journal of V. N. Karazin Kharkiv National University*, 36, 14-26.
3. Bolshakova E. I. (2011). *Automatic natural language processing and computer linguistics: tutorial guide book* [Avtomaticheskaia obrabotka tekstov na estestvennom iazike i compiuternaia lingvistika: uchebnoe posobie]. M: MIEM.