

**Предизвикателството „фразеологични единици“ –
изкуственият интелект срещу човешката интуиция**

Владислав Маринов, Анита Тодоранова

**The „Idiomatic Expressions“ Challenge –
AI vs. Human Intuition**

Vladislav Marinov, Anita Todoranova

Abstract:

This report compares, from a pragmatic perspective, the perception of ten idiomatic expressions (with the components honey and butter) by humans and by some modern artificial intelligence (AI) models.

The results of surveys conducted with 35 students showed that a significant portion of the surveyed individuals struggle to determine the precise meaning and contextual usage of the idioms. In contrast, AI demonstrates an exceptionally high level of comprehension, providing accurate definitions and appropriate examples of usage. The analysis reveals that the difficulties experienced by the students result from a lack of knowledge of the cultural and historical context from which the idioms originate.

The report highlights the increasing role of AI in language processing and its potential to assist in the comprehension and acquisition of complex linguistic constructions such as idioms.

In the conclusion, options for optimizing language education are discussed, utilizing the capabilities of artificial intelligence.

Keywords: idioms, phraseological units, artificial intelligence, globalization.

For contacts: Asst. Prof. A. Todoranova, St. Cyril and Methodius University of Veliko Tarnovo,
a.todoranova@ts.uni-vt.bg

Духът на езика най-отчетливо се изразява в непреводимите думи.
Мария фон Ебнер-Ешенбах

ВЪВЕДЕНИЕ

В епохата на глобализация и технологичен напредък, когато светът се смалява до размерите на един екран, е особено важно да съхраним своята идентичност и културно наследство. Фразеологичните единици (ФЕ⁴⁴), като неделима част от езика, са съкровищница на народната мъдрост. Във връзка с това М. Яскот обобщава, че „езикът и културата са взаимосвързани, тъй като езикът, съхранявайки културно знание, на свой ред оказва влияние върху колективната и индивидуалната менталност, като дори ги формира в определена степен“ (Яскот 2019: 94). Фразеологизмите и техните значения, носещи духа и историята на един народ, постепенно избледняват от речника на младото поколение. В ерата на дигиталните технологии, когато информацията е достъпна на един клик разстояние, младите хора все по-рядко се задълбочават в изучаването на родния си език. С оглед на технологичния свят, в който представителите на поколение Z израстват и се обучават, все по-важна роля имат

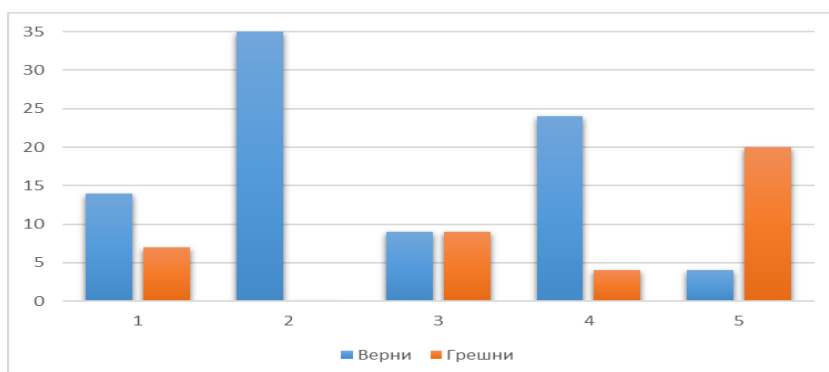
⁴⁴ В българската лексикология редом с използвания тук термин *фразеологична единица* са се наложили и термините *фразеологизъм*, *устойчиво словосъчетание*, *фразеологично словосъчетание*, *фразема* и т.н. (вж. по-подробно у Бояджиев, 2007, с. 293; Вътов, 1995, с. 359 – 360; Радева, 2017, с. 225).

езиковите модели, базирани на изкуствен интелект (Белоев и др. 2024). Както отбелязват Г. Иванова и А. Иванов, чрез изкуствения интелект се улеснява индивидуалният подход в обучението и учебното съдържание може да се адаптира към интересите на обучаемите, което подобрява не само резултатите, но и мотивацията за учене (Ivanova, Ivanov 2024).

ИЗЛОЖЕНИЕ

Целта на доклада е да се съпоставят знанията на представители на поколението Z със „знанията“ на три езикови модела, базирани на изкуствен интелект – ChatGPT (Open AI), Gemini (Google) и Copilot (Microsoft), относно значението на някои фразеологични единици с компонентите *мед* и *масло*. В проведения експеримент 35 първокурсници от специалност „Приложна лингвистика“ във Великотърновския университет „Св. св. Кирил и Методий“ трябваше да посочат значението на пет фразеологизма: **1. Бия водата да пусне масло.** – *Полагам напразни усилия за нещо неизпълнимо; върша безплодна работа.*; **2. Върви по мед и масло.** – *За нещо – ставам, извършвам се без пречки или затруднения, лесно и много добре.*; **3. Светих маслото.** – *Убих, премахнах някого. Обикн. умишлено за наказание или за отмъщение.*; **4. Ни<то> меда, ни<то> жилото не искам (не ща).** – *Абсолютно нищо, нито добро, нито лошо (не искам).*; **5. С мед да го намажеш.** – *Много ще съм доволен, ще се примиря с нещо (за нещо по-малко лошо, отколкото изглежда или по-малко лошо, по-поносимо в сравнение с друго)*⁴⁵. Анкетираните лица имаха възможност да изберат на кои фразеологизми да посочат значението и кои да пропуснат, затова в интерпретацията на данните има различен брой отговори, въпреки че анкетата беше проведена еднократно и обхвана цялата група студенти.

Резултатите от анкетата показаха, че значителна част от студентите не познават точното значение на изследваните фразеологизми, а се опитват да възстановят значенията им според семантиката на компонентите. Фокусът в анализа е върху някои от по-интересните отговори, тъй като те представят разнообразието на асоциативните връзки, които прави човешкият мозък въз основа на индивидуалния опит и усвоените знания, а обобщените резултати са представени на фиг. 1.



Фигура 1. Резултати от анкетните карти

⁴⁵ При представянето на значенията на ФЕ е използван *Фразеологичният речник на българския език, т. 1 и 2*, с автори К. Ничева, С. Спасова-Михайлова и Кр. Чолакова.

По отношение на първия фразеологизъм **Бия водата да пусне масло.** отговор са дали 21 студенти, като верните отговори са 14. Сред останалите отговори на преден план е тълкуването на маслото като вещество с определена консистенция и плътност, различаващо се от водата, срв. **Да е меко самото нещо.; От нищо правя нещо.** (тук вероятно има асоциации *прозрачност = нищо, плътност = нещо*). Усилията при извършването на действието са в основата на други две тълкувания – **Правя нещо грешно.; Карам човек да учи и чакам добър резултат.** В първото е доловена липсата на смисъл в действието, която е интерпретирана като неправилен начин за извършването му. При второто тълкуване вероятно контекстът при попълване на анкетата (учебна зала) е повлиял върху обвързването му с процеса на обучение, но и тук за студента наличието на масло като резултат от дейността е нещо хубаво, т.е. отново имаме противопоставяне между веществата *вода и масло*, като маслото е маркирано като *добър резултат = положително*.

За отговорите на **Върви по мед и масло.** може да се отбележи, че всички анкетиращи лица са посочили вярно значението му, като двама са го обвързали само с работата, срв. **Лека работа.; Работата ми върви страхотно.**, а за един анкетиран значението е общо и се свързва не толкова с липсата на пречки при извършването на нещо, а по-скоро с цялостното усещане – **Всичко ни е хубаво.** Може да се допусне, че високият процент на верни отговори се дължи на факта, че този фразеологизъм е познат и анкетираните лица често са го срещали в ежедневието си.

Третият фразеологизъм **Светих маслото.** е затруднил студентите и половината от тях са предпочели да не отговарят. Само 9 са посочили вярно значението му (26% от общия брой анкетиращи лица). Сред останалите отговори прави впечатление представянето на различни етапи от справяне с някаква дейност, с нейното разбиране, което вероятно е резултат от асоциирането на изследвания фразеологизъм с разговорната конструкция **светнах се** <за нещо> със значение *изясни ми се*, срв. **Разбрах нещата.; Осъзнах се.; Справих се.; Открих нещо.** Последният отговор може да се обвърже и с осветяването на някакво място, в резултат от което нещо вече е видимо.

Висок процент на посочили точното значение се наблюдава и при фразеологизма **Ни<то> меда, ни<то> жилото не искам (не ща).** – 86% (24 от 28 анкетиращи лица). От останалите четирима студенти трима обвързват фразеологизма с неопределеност, при която преобладава негативната конотация, срв. **За нищо не става – ни едното, ни другото направил.; Разочарован съм.; Хем така, хем иначе.**, докато четвъртият е посочил някакво значение, без да го обвърже с нещо конкретно. Основание за това предположение е въпросителният знак в края на отговора, срв. **Стана ми хубаво от нещо?**

За последната ФЕ **С мед да го намажеш.** може да се отбележи, че това е фразеологизмът, за който най-малко анкетиращи лица са посочили вярно значението му – само 17 % (4 от 24 отговорили). Сред останалите отговори могат да се обособят следните групи: 1) Анкетираните лица свързват меда с нещо приятно като усещане, срв. **Много съм доволен от нещо.; Ще съм доволен.; Изразява щастие.; Да го направиш хубаво, приятно.; Да го глезиш.**; 2) За

анкетираните лица са важни лечебните свойства на меда, срв. **Да го излекуваш с някакъв крем примерно.; Ранен си, грижи се добре.; Да му помагаш или да му даваш препоръки.;** 3) Асоцииране на меда с нещо добро и оттам със загриженост и добро отношение към другия, срв. **Да му галиш егото.; Да го глезиш.** 4) Асоцииране на меда с нещо сладко, което води до добро/положително в характера на човека, срв. **Добър човек.; Човек, който е благ.; Добър, оправен във всички сфери.; Да се измъкне лесно от проблем.; Добре се е справил.; Да се съглася с нещо.; Ще се задоволя с нещо.** Интересно е, че отговорите **Да го глезиш.** и **Добър човек.** се срещат по два пъти.

От друга страна, може да се отбележи, че трите езикови модела (ChatGPT, Gemini и Copilot) като цяло успешно предоставят точни дефиниции на фразеологичните единици, но се различават по начина на представяне на отговорите. ChatGPT дава най-подробни обяснения, често с контекст или етимология на ФЕ, Copilot се отличава с краткост и директност, а Gemini предлага ясни интерпретации, като понякога дава и допълнителни детайли. По отношение на **Бия водата да пусне масло.** всички са точни, но ChatGPT добавя най-много детайли, а Copilot е най-лаконичен. За **Светих маслото.** и трите модела правилно предават значението. При **Ни<то> меда, ни<то> жилото не искам (не ща).** ChatGPT отново се отличава с обяснение на произхода, докато другите са по-кратки. Четвъртият фразеологизъм **С мед да го намажеш.** затруднява езиковите модели и нито един от тях не дава истинския му смисъл (*примирение с нещо по-малко лошо*), като Copilot и Gemini наблягат на приятния характер на човека, а според ChatGPT фразеологизмът означава, че нещо е много хубаво, приятно, ценно или желано. Във връзка с **Вървя по мед и масло.** всички дават идентични и точни дефиниции. Резултатите показват, че моделите са напреднали в разбирането на фразеологизми, но могат да пропускат нюанси, особено при по-сложни или иронични изрази. Тези изводи обаче не бива да се абсолютизират, тъй като „знанията“ на всеки от езиковите модели зависят от обема на текстовете, с които са обучавани. По отношение на фразеологизмите езиковите модели достигат до значенията им чрез анализ на значенията на компонентите на ФЕ, както и анализ на семантиката на думите в близък или далечен контекст, и статистическа обработка на честотата на прояви на съответния фразеологизъм в определен контекст. С други думи, налице е статистическо разпознаване на семантиката на ФЕ чрез контекста. Затова тук обемът на текстовете (брой и дължина) е важен за степента на определяне на точното значение на фразеологизмите, особено при тези, които имат и допълнителни конотации.

ЗАКЛЮЧЕНИЕ

Настоящото изследване показва как различните „интелекти“ (човешки и изкуствен) се справят с един аспект на езика, който е едновременно често срещан, но и доста коварен. От анализа на резултатите може да се обобщи, че студентите при дефинирането на значението на непознати ФЕ разчитат предимно на интуиция и на лични асоциации или извеждат значението чрез тълкуване на семантиката на отделни думи от ФЕ, което води до разнообразни, но често неточни интерпретации. В рамките на проведения експеримент точността на

техните тълкувания варира значително – от 100% за познати изрази до 17% за непознати за тях фразеологизми.

Езиковите модели, от друга страна, демонстрират по-висока обща точност благодарение на статистическата обработка на данни, особено при често срещани фразеологизми. Използваните модели пропускат нюанси в семантиката и предлагат стандартизирани интерпретации. Както вече беше отбелязано, човешкото тълкуване е динамично и силно зависи от непосредствения контекст и субективния опит, докато изкуственият интелект разчита на голям обем от текстове и данните от статистическата им обработка.

Съпоставката между двата „интелекта“ очертава фундаментални разлики – човекът проявява креативност и контекстуална гъвкавост, но често неточност поради липса на предишен опит, свързан с натрупване на познание, докато езиковите модели предлагат по-голяма надеждност при добре документиран изрази, но засега им липсва човешката изобретателност. Това показва, че въпреки напредъка на изкуствения интелект в езиковия анализ, езиковите модели все още не могат напълно да заместят човешкото разбиране на културно обусловените фразеологизми, но потенциалът на ChatGPT, Gemini и Copilot може да се използва успешно при оптимизацията на базовото езиково обучение в областта на фразеологията.

ЛИТЕРАТУРА

1. Белоев и др. 2024: Белоев, Хр., В. Войноховска, А. Смрикаров. Концептуална рамка за използване на изкуствения интелект във висшето образование. – В: Стратегии на образователната и научната политика, 32 (5s), 11 – 22.
2. Бояджиев 2007: Бояджиев, Т. Българска лексикология. София: УИ „Св. Кл. Охридски“.
3. Вътов 1995: Вътов, В. Фонетика и лексикология на българския език. Велико Търново: Абагар.
4. Радева 2017: Радева, В. Българска лексикология и лексикография. София: Изток-Запад.
5. Яскот 2019: Яскот, М. Отражение на културни компоненти във фразеологичните единици. – В: Български език, 66 (2), 93 – 106.
6. Ivanova, Ivanov 2024: Ivanova, G., A. Ivanov. Comparative Analysis of Utilizing Popular Intelligent Computer Systems in Education. – В: Стратегии на образователната и научната политика, 32 (5s), 53 – 63.

ИЗТОЧНИЦИ

Фразеологичен речник на българския език. Авт. К. Ничева, С. Спасова-Михайлова, Кр. Чолакова. Т. 1 (1974); Т. 2 (1975), София: Издателство на БАН.