

## ГЕНЕРИРАНЕ НА ЕЗИКОВИ ФРАГМЕНТИ В КОМПЮТЪРНА СРЕДА В ПОДКРЕПА НА ИЗСЛЕДОВАТЕЛСКИЯ ПОДХОД В ЕЗИКОВОТО ОБРАЗОВАНИЕ<sup>1)</sup>

Евгения Сендова

Институт по математика и информатика – БАН

**Резюме.** В статията са представени идеи за прилагане на изследователски подход при изучаването на естествени езици в специализирана компютърна среда (*микросвят* в езика за програмиране Лого за обработка на думи и списъци). Разгледани са конкретни примери за моделиране на езикови фрагменти като елемент на учебни ресурси, разработени в контекста на *Проблемната група по образованието* (1978 – 1999), и на учебници по информатика в средното училище в подкрепа на интегрирането ѝ с езиковото образование.

Във връзка със засиления интерес към прилагането на изследователския подход в образованието не само по природни науки, но и по математиката, информатиката и естествените езици, се защитава тезата, че в подходяща компютърна среда учениците могат да действат като изследователи в тези области. Като разработват компютърни модели на езикови явления, учениците материализират хипотезите си за тези явления, проверяват поведението им с различни входни данни и виждат доколко универсални са те. Като получат обратна връзка след изпълнението на компютърните си модели, учениците могат да *редактират идеите* си, т.е. да ги коригират, модифицират, прецизират, усъвършенстват и обогатяват. Този типичен за научните изследвания подход се оказва важен фактор за повишаване на мотивацията на учениците да изучават тънкостите на естествения език.

**Keywords:** Inquiry based language education; computer microworlds

*Exceptions prove the rules!*

Популярна пословица

### 1. Увод – как да учим граматика?

Замисляли ли сте се как така „изключенията потвърждават правилото“? Срещат се различни тълкувания, но една правдоподобна хипотеза гласи, че сегашната интерпретация на това твърдение е резултат от погрешен превод. Думата *prove* на английски език е имала освен сегашното си значение (*do-*

казвам) и друго, свързано с латинската дума *probo*, която означава *пробвам, тествам, проверявам истинността на нещо*<sup>2)</sup>. Та друг възможен превод от английски би бил нещо от типа на: *Изключенията тестват правилото*. С други думи, когато намерим изключения на формулирано от нас (или от някой авторитет) правило, бихме могли да се постараем да го коригираме така, че да обхване изключенията.

Дори да е спорен този превод, идеята зад него може чудесно да се използва с образователна цел. Изучаването на граматиката в училище не е достатъчно мотивирано – човек може да говори правилно на родния си език, без да е в състояние да формулира правилата, на които този език се основава. От друга страна, когато изучаваме чужд език, понякога сме в състояние да декламираме граматични правила, без да можем да ги приложим на практика.

След външното оценяване по български език и литература за VII клас (2017 г.) в пресата бяха цитирани тревожни констатации на оценителите от вида:

*Учениците все още масово не могат да разпознават видовете сложно изречение по състав.*

Естествената реакция на голям брой читатели (вкл. езиковеди, автори на книги, редактори на научни списания или просто любители на граматиката) бе, че едва ли компетентността да се разпознават видовете сложни изречения, е основен показател за езикова култура.

Впрочем нека погледнем официалната класификация на сложните изречения<sup>3)</sup>.

- **Сложно съчинено разделително (съотносително) изречение** – изразява действия, които се редуват или се изключват взаимно.

- **Сложно съчинено противоположно изречение** – сказуемите на простите изречения в състава на сложното се противопоставят по смисъл.

Сигурно не само учениците ще се поколебят към коя група да причислят изречението: *Или ще живея свободно, или ще умира*.

Естествено, на всеки етап от познанието хората си изграждат теории за явленията, с които се сблъскват. За да усъвършенстват тези теории, те трябва да ги представят в достатъчно ясен и точен вид. За представянето на лингвистични идеи, отношения и модели е необходима специализирана система на означения. Но ако моделите, които създаваме, са не просто абстрактни системи от правила, а *работещи модели* (например реализирани в компютърна среда), можем да проверяваме поведението им с различни входни данни и да видим доколко универсални са те. Освен това можем да моделираме и подобряваме по-сложните правила с помощта на последователни приближения.

В този смисъл, можем да гледаме на компютъра като на среда, която е *редактор на идеи*, стига тези идеи да се облекат в подходяща форма. Като получим обратна връзка след изпълнението на компютъра, можем да

редактираме идеите, т.е. да ги коригираме, модифицираме, прецизираме, усъвършенстваме и обогатяваме.

И този присъщ на учените дух, който в последно време наричаме *изследователски подход*, търси своето достойно място в образованието не от днес и от вчера...

## **2. Изследователски подход в образованието по естествените езици**

Ако се вгледаме в естествените наклонности на децата, ще си дадем сметка, че още от най-ранна възраст те се *проявяват като учени* – например при играта с езика. Част от тази игра е един вид експеримент. Макар тези експерименти да не са планирани, децата разполагат (в главата си) с невероятен голям запас от данни за езика и непосредствен достъп до тях. Освен това те разработват теории за организиране на хаоса от данни. Възможно е например едно дете да образува думи като *скоци* и *соци* по подобие на множественото число на *вълк* и *поток*. Или да предаде основната идея на *Лебедово езеро* с фразата: *Магьосникът превърна една девойка в лебедка*. Едва след като теорията му се обогати с нови познания или си създаде нова теория, детето може да открие вярната форма.

### **2.1. „Сега знам защо имаме съществителни и глаголи!“**

Идеи за приложение на езикови изследвания като учебна дейност с помощта на компютърни модели се появяват в края на 60-те години на миналия век, когато Сиймър Пепърт и сътрудниците му, включващи експерти по психология, информатика и изкуствен интелект, експериментират първите версии на езика за програмиране Лого с групи от по няколко деца в рамките на амбициозни проекти за училища на бъдещето. Да си припомним едно споделяне на Пепърт, свързано с този експеримент.

Джени, 13-годишна ученичка, била включена със свои връстници в експериментална група за програмиране на Лого. Задачата била да научат компютъра да генерира изречения от вида *Кой какво прави*, например:

- *Лудият Питър обича, защото грозният мъж мрази.*
- *Сладкото куче бяга, защото дебелият вълк подскача.*
- *Лудият вълк мрази, защото мършавият вълк подскача.*

Макар че в училище Джени вече била минала материала за граматически категории, така и не успяла да разбере разликата между съществителните, глаголите и прилагателните. Не че не можела да работи с логически категории, просто не разбирала защо ѝ трябва граматиката и ѝ се струвало безсмислено да я изучава. Когато Джени попитала учителката си за какво служи граматиката, получила отговор, който не я задоволил: „Граматиката ти помага да говориш по-добре!“. Първо, Джени не разбирала как така изучаването на граматически правила ти помага да говориш, още по-малко чувствала нужда от помощ, за да се изразява. Така че граматиката предизвиквала у нея чувство на възмущение. За да научи обаче компютъра да генерира изречения,

й се наложило да раздели думите на категории... Не защото някой й казал, а защото почувствала нуждата да го направи!

Преди започване на експеримента тя се обърнала към Пепърт с думите: *Не виждам защо сте ни избрали... Ние не сме най-големите умници в класа.* Няколко седмици по-късно тя се приближила до учения много развълнувана: *Сега вече знам защо имаме съществителни и глаголи!*

Това, което Джени научила за граматиката от това общуване с компютъра, било съвсем различно от механичното наизустяване на правила. Тя не само разбрала необходимостта от категоризиране на думите, но направила това като собствен инструмент за обучаване на компютъра. В резултат на подобреното й самочувствие цялостното поведение и отношение на Джени към ученето се подобрили съществено.

## **2.2. Интегриране на образованието по математика, информатика и естествени езици в български контекст**

В книгата си *Exploring Language with Logo* авторите Голденберг и Форцайг интегрират философията на научното изследване с математическото моделиране и перспективите за обучението по лингвистика, информатика и когнитивна психология. Много от предложените от тях изследвания са били експериментирани с американски ученици от прогимназията и гимназията. Някои от тези идеи бяха адаптирани за обучението по български език и за някои аспекти на обучението по чужди езици, а също и в рамките на интегрирания предмет *Език и математика* за училищата на *Проблемната група по образованието* (ПГО). Работата с ученици и с учители от тези училища показва, че езикът може да се превърне в привлекателен обект за изследвания на различни равнища – нещо, което може, от една страна, да мотивира изучаването на структурата на езика, а от друга – да спомогне за опознаването на някои научни методи, присъщи на природните науки. При компютърното моделиране на фрагменти от естествения език дори най-абстрактни понятия, свързани със структурата на езика, могат да станат *посезаеми и конкретни* за учениците. За да накарат компютъра да генерира фрагменти от езика, те трябва да формализират наблюденията си във вид на програма и да я тестват с достатъчно много примери. Когато генерираният от компютъра текст не отговаря на очакванията им, се налага да ревизират хипотезата си.

## **2.3. ФилоЛОГО – един компютърен микросвят за езикови експерименти**

Учебниците и компютърните *микросветове* за работа с езика в изследователски стил бяха разработени в рамките на ПГО на основата на Лого – като компютърен език и като образователна философия, известна като *конструкционизъм*. Като език, Лого е представител на фамилията езици от областта *Изкуствен интелект* и основното му предназначение е за обработка на думи и списъци. В първите версии на Лого вградените операции са

около стотина и включват само основните операции за обработка на думи и списъци – *извличане на първи и последен елемент, отрязване на първи или последен елемент и обединение на думи и списъци* (съответно ПЪРВИ, ПОСЛЕДЕН, БЕЗПЪРВИ, БЕЗПОСЛЕДЕН, ДУМА, ОБЕДИНЕНИЕ). В един по-специализиран микросвят, ФилоЛОГО, са вградени допълнителни операции, включващи *извличане на произволен елемент от дума или списък (ИЗБРАНОТ), обръщане на елементите на обекта (ОБЪРНАТ), случайно разбъркване на елементите на обекта (РАЗБЪРКАН), намиране на дължината на обекта (БРОЙ), проверка дали дадена буква е гласна (ГЛАСНА?), проверка дали дадена дума има окончание, принадлежащо към даден списък (ЗАВЪРШВА\_НА)* и т.н.

Да разгледаме няколко представителни примера на моделиране в този езиков микросвят, които биха могли да дадат представа как учебният процес в часовете по език и литература може да бъде допълнително мотивиран, като изграждаме у учениците стил на научно изследване.

#### **2.4. Моделиране на изречения от тип *клюка***

Да отбележим, че един проект за изследване на структурата на изречението, който се радва на еднакъв успех и при ученици на различна възраст, и при студенти, и при учители, е свързан с моделиране на *клюки*. Вероятната причината е, че *клюката* има напълно стереотипна форма, може да бъде забавна (защото е свързана с конкретни личности) и когато бъде генерирана от компютър, се възприема от създателите на програмата като безпристрастна.

Досегашният ми преподавателски опит показва, че дори и да *мразят* граматиката, учениците с удоволствие я впрягат в модела си за *клюки*. Приятно се изненадват, че могат да накарат компютъра да се държи *по човешки*, и се ентусиазират, че могат да управляват поведението му (дори цената да е използването на някое граматическо правило). И те като цитираната по-горе Джени често възкликват: *Чак сега да разбера защо трябва някои думи да се определят като съществителни имена, а други – като глаголи.*

Самият процес на моделиране протича по следния начин – за да стигнем до най-простия възможен модел на *клюката*, предлагаме на учениците дадат примери на *клюки*. След това се опитваме заедно да формализираме структурата на *клюката*.

При първото приближение на модела могат да се вземат само най-основните структурни елементи – можем да ги наречем условно: КОЙ и КАКВО\_ПРАВИ. Нека например научим компютъра да генерира *клюки* от типа на следните изречения:

*Георги хърка.*

*Иван флиртува.*

*Таня бърбори като картеница.*

*Пенка философства.*

Съответните процедури на Лого са:

ЗА КОЙ

ИЗХОД ИЗБРАН ОТ [ГЕОРГИ ИВАН ТАНЯ ПЕНКА]

КРАЙ

ЗА КАКВО ПРАВИ

ИЗХОД ИЗБРАН ОТ [ХЪРКА ФЛИРТУВА ФИЛОСОФСТВА [БЪРБОРИ  
КАТО КАРТЕЧНИЦА]]

КРАЙ

Разбира се, получените клюки няма непременно да съвпадат с примерните (поради това, че резултатът от операцията ИЗБРАН ОТ е случайно избран елемент от конкретния списък):

– ИВАН ХЪРКА

– ТАНЯ ФЛИРТУВА

– ТАНЯ ФИЛОСОФСТВА

– ПЕНКА БЪРБОРИ КАТО КАРТЕЧНИЦА

Ако допълнително искаме компютърът да *клюкарства* в стила на даден автор, например Чудомир, можем да опишем съответните процедури въз основа на най-характерните клюки на неговите героини:

ЗА КЛЮКА

ИЗХОД (ОБЕДИНЕНИЕ КОЙ КАКВО\_ПРАВИ)

КРАЙ

ЗА НЕ\_СЪМ\_ОТ\_ТЯХ

ИЗВЕДИ УВОД

ИЗВЕДИ КЛЮКА

ИЗВЕДИ ЗАКЛЮЧЕНИЕ

КРАЙ

ЗА УВОД

ИЗХОД ИЗБРАН ОТ [[НЕ МИ Е РАБОТА И НЕ СЪМ ОТ ТИЯ, ДЕТО  
СЕ БЪРКАТ В~

ХОРСКИТЕ РАБОТИ, АМА][ДА ЗАПОЧНА ДА ТИ РАЗПРАВЯМ,~  
АМА НЕ МИ Е ПО ХАРАКТЕРА.]]

КРАЙ

ЗА КОЙ

ИЗХОД ИЗБРАН ОТ [[ТАНА ПАПУЧКИНА][БОНА КЕВИНА] [НАШ  
ПРОЙКО]]

КРАЙ

ЗА КАКВО\_ПРАВИ

ИЗХОД ИЗБРАН ОТ [[ВСЕ ПО МРЪКНАЛО ИЗЛИЗА.][НЕ ЗНАЕ НИ  
ДА СГОТВИ,~

НИ ДА ОПЕРЕ, НИ НЕЩО ДА УШИЕ.]~

[ДА ИЗЛЪЖЕ НЯКОГО – ПЕЙ ДАВА.]]

КРАЙ

ЗА ЗАКЛЮЧЕНИЕ

ИЗХОД ИЗБРАНОТ [[НЕ МИ ТРЯБВА ДА ГИ РАЗПРАВЯМ.]

[АМА ДА СИ МЪЛЧИ ЧОВЕК.] [ЧУНКИМ У ДОМА ~

ВСИЧКО Е ПОТЕКЛО ПО ВОДА...]]

КРАЙ

Възможни резултати от програмата КЛЮКА са изречения от вида:

– ТАНА ПАПУЧКИНА СТОИ, НЕ ЗНАЕ НИ ДА СГОТВИ, НИ ДА ОПЕРЕ, НИ НЕЩО ДА УШИЕ.

– НАШ ПРОЙКО ДА ИЗЛЪЖЕ НЯКОГО – ПЕЙ ДАВА.

– НАШ ПРОЙКО НЕ ЗНАЕ НИ ДА СГОТВИ, НИ ДА ОПЕРЕ, НИ ДА УШИЕ.

Последната клюка не звучи много правдоподобно за Чудомировата как ‘Сийка. Затова в едно следващо приближение на процедурата КЛЮКА трябва да се избегне обединяването на този герой с това действие.

По-строго описание на простите изречение е направено в учебника *Информатика в стил Лого*.

## 2.5. Естественият език в един учебник по информатика за X клас

*Информатика в стил Лого* съдържа ядро от задължителни по програмата информатични елементи, последвано от няколко направления, в които авторите предоставят избор на учителя и учениците да правят изследвания в подходяща информатична среда. Едно от тези направления бе естественият език. Ето как изглежда например фрагмент от урока за генериране на прости изречения.

### Естествени и формални езици

**Естествени езици** са тези, на които хората говорят и пишат, на които общуват помежду си – български, английски, испански, японски и т.н. Много учени са имали амбицията да осъществят използването на естествените езици като средство за общуване с компютъра, но засега това все още не е постигнато. Главните направления, в които се работи, са свързани с обработката на текст, с компютърния превод, с отговорите на въпроси от дадена област. Те се основават на теорията на т. нар. **формални езици** – съвкупности от думи, построени по дадени правила. Тази теория се занимава с това как се задават, разпознават и преобразуват формалните езици. Основна нейна задача е да се моделират и изследват свойствата на естествените езици и на някои изкуствени езици – например езиците за програмиране. Моделирането на естествен език е свързано с въпроса как да се опише някаква негова (достатъчно голяма) подсвкупност. Един възможен подход е да се състави система от правила, с които може да се определи кои изречения попадат в интересувашата ни подсвкупност от естествения език и кои са извън нея.

## 2. Най-прост модел на изречение

Да разгледаме изреченията:

Иван рисува.

Манрико пее.

Георги хърка.

Таня се смее.

Пенка философства.

(1)

Всички те са от вида *еди-кой си* (съществително собствено име) *прави еди-какво си* (глагол). Да опишем система от правила, с помощта на които ще свързваме дадени думи (речник) така, че да се получават изречения от типа (1). За да генерираме изречения с такава структура, ще обособим в отделни списъци съществителните имена (N) и глаголите (V):

$N = \{\text{Иван, Манрико, Георги, Таня, Пенка}\}$  (2)

$V = \{\text{рисува, пее, хърка, смее се, философства}\}$

Правилото, по което се получават всички изречения (1) и други от същия тип, може да се запише символично така:

$\sum_i N + V$  (3)

Знакът  $\Sigma$  (гръцката буква сигма) означава изречение. Долният индекс означава, че правилото (3) описва само един вид изречения. Това граматическо правило се чете по следния начин: *Всяко свързване на думи, което се получава при заместването на N и V с елемент от съответния списък, е изречение от типа (1).*

## 3. Програмна реализация на модела

Да опишем процедура на Лого, с която се генерират изречения от този тип:

ЗА ИЗРЕЧЕНИЕ1

ИЗХОД ОБЕДИНЕНИЕ СЪЩЕСТВИТЕЛНО ГЛАГОЛ

КРАЙ

ЗА СЪЩЕСТВИТЕЛНО

ИЗХОД ИЗБРАН ОТ [ИВАН МАНРИКО ГЕОРГИ ТАНЯ ПЕНКА]

КРАЙ

ЗА ГЛАГОЛ

ИЗХОД ИЗБРАН ОТ [ПЕЕ РИСУВА ХЪРКА [СЕ СМЕЕ] ФИЛОСОФСВА]

КРАЙ

С операцията ИЗБРАН ОТ, която очаква един вход – списък, се избира по случаен начин елемент от входния списък. Следователно изреченията,

генерирани от процедурата ИЗРЕЧЕНИЕ1, са **обединение** на **съществително**, което е елемент, **избран** по случаен начин **от** списъка N, и **глагол**, **избран** по случаен начин **от** списъка V, т.е. са от вида (1). Да се убедим в това:

ПОВТОРИ 5 [ИЗВЕДИ ИЗРЕЧЕНИЕ1]

**МАНРИКО ХЪРКА**

**ГЕОРГИ ФИЛОСОФСТВА**

**ГЕОРГИ СЕ СМЕЕ**

**ТАНЯ РИСУВА**

**ИВАН ПЕЕ**

Всъщност нито едно от получените изречения не съвпада напълно с примерите от (1), но при достатъчен брой повторения и това ще стане.

За да актуализираме по-лесно списъците N и V, по-удобно е предварително да ги направим стойности на променливи, а в процедурите **СЪЩЕСТВИТЕЛНО** и **ГЛАГОЛ** да използваме тези стойности:

НАПРАВИ "N [ИВАН МАНРИКО ГЕОРГИ ТАНЯ ПЕНКА КУЧЕТО КОТКАТА ЖЕНАТА МЪЖЪТ ТРУБАДУРЪТ КОНЯТ ЩАСТИЕТО СЛАВАТА УСПЕХЪТ]

НАПРАВИ "V [ПЕЕ РИСУВА ХЪРКА [СЕ СМЕЕ] ФИЛОСОФСТВА СКАЧА РЪМЖИ БЯГА ПРЕПУСКА [СЕ ИЗПЛЪЗВА] ОМАЙВА ЗА-СЛЕПЯВА]

ЗА **СЪЩЕСТВИТЕЛНО**

ИЗХОД ИЗБРАН ОТ: N

**КРАЙ**

ЗА **ГЛАГОЛ**

ИЗХОД ИЗБРАН ОТ: V

**КРАЙ**

Обогатили сме списъка N със съществителни нарицателни (в членуваната им форма) и списъка V – с няколко нови глагола. Сега вече с процедурата ИЗРЕЧЕНИЕ1 бихме могли да получим 168 (14 ' 12) различни изречения, между които например:

**КУЧЕТО СКАЧА**

**КОНЯТ ПРЕПУСКА**

**ЖЕНАТА ОМАЙВА**

**ЩАСТИЕТО СЕ ИЗПЛЪЗВА**

#### **4. Пораждаща граматика**

Да означим с G граматиката, която се състои от речника (2) и правилото (3). Всички изречения, които могат да се получат, като свържем думите от (2) според правилото (3), наричаме *генерирани (породени) от граматиката G*, а самата граматика – *пораждаща*. Това са различните изречения, които могат да се генерират с процедурата ИЗРЕЧЕНИЕ1.

Много вероятно е при изпълнението на тази процедура да получим и изречения, които са граматически правилни, но смислово неприемливи, например:

ЩАСТИЕТО РЪМЖИ

КОНЯТ СЕ СМЕЕ

КУЧЕТО ОМАЙВА

За да намалим вероятността да се получават смислово неприемливи изречения, ще усъвършенстваме нашата граматика, като направим нова класификация на думите от  $N$  и  $V$  и добавим нови правила:

$N = \{N_1, N_2, N_3\}$

$N_1 = \{\text{Иван, Манрико, Георги, Таня, Пенка, жената, мъжът, трубадурът}\}$

$N_2 = \{\text{кучето, котката, конят}\}$

$N_3 = \{\text{щастие, славата, успехът, жената}\}$

$V = \{V_1, V_2, V_3\}$

$V_1 = \{\text{пее, рисува, хърка, се смее, философства, скача, бяга}\}$

$V_2 = \{\text{скача, ръмжи, бяга, препуска}\}$

$V_3 = \{\text{се изплъзва, омайва, заслепява}\}$

Тази класификация ни позволява да формулираме правила, според които ще се генерират по-малък брой смислово неприемливи изречения. Сега ще заместим правилото (2) за образуване на изречения с няколко нови правила, всяко от които определя някакъв тип изречения:

$\sum_1: N1 + V1$

(напр. Мъжът философства.)

$\sum_2: N1 + V_2$

(напр. Жената омайва.)

$\sum_3: N2 + V2$

(напр. Конят препуска.)

$\sum_4: N3 + V3$

(напр. Щастие, омайва.)

Разбира се, можем да усложним граматиката – за да изключим например изречения от типа на *Конят ръмжи*, но усложняването трябва да става в разумни граници – да се елиминират всички смислово неприемливи изречения, едва ли е възможно, а и самото понятие за тях не е точно определено.

Според новите правила можем да свързваме елементите на  $N_2$  само с елементите на  $V_2$ , елементите на  $N_3$  – само с тези на  $V_3$ , а елементите на  $N_1$  можем да свързваме както с елементите на  $V_1$ , така и с тези на  $V_3$ . Новата граматика ще означим с  $G_2$ . Редактираната в съответствие с  $G_2$  програма ще генерира изречения от вида:

СЛАВАТА ЗАСЛЕПЯВА  
КУЧЕТО РЪМЖИ  
ЖЕНАТА СЕ ИЗПЛЪЗВА  
КОНЯТ РЪМЖИ  
МЪЖЪТ ОМАЙВА

Повечето изречения са приемливи, но някои са спорни в смислово отношение.

И така, дефинирахме граматиката G2, която се състои от речник (с определена класификация на думите в него) и правила, по които тези думи могат да се свързват в изречения. Всички изречения, които могат да бъдат породени от G2, са *правилни в рамките на тази граматика*. Те образуват едно скромно подмножество на множеството от изреченията, които могат да се изречат на български език. Между тях все още има такива, които не са съвсем приемливи от смислова гледна точка. С обогатяване на речника ще се увеличава броят на изреченията, които могат да се генерират в съответствие с тези правила. Ако обаче не изменим по подходящ начин правилата, броят на неприемливите изречения ще нарасне.

Задачите към този урок имаха за цел да насърчат учениците да правят вариации върху предложените модели.

### **ЗАДАЧИ**

1. Изменете съдържанието на списъците N и V и дефинирайте нова класификация на думите в тях и правила за генериране на изречения в рамките на новата граматика.

2. Дефинирайте процедура, с която да се генерират изречения от вида:  
*Някои мъже са умни. Всички жени са приказливи. Повечето момчета са благородни.*

3. Изберете няколко пословици от вида:

*Който много бърза, далеч не стига. Който не работи, не трябва да яде.*

*Който гроб копае другиму, сам пада в него. Който нож вади, от нож умира.*

Оформете два списъка съответно от първите и вторите части на тези пословици. Дефинирайте процедура, която да групира по случаен начин първата част от една пословица с втората част от друга, например: *Който гроб копае другиму, от нож умира*. Изследвайте колко често се получават смислени пословици.

Може би си струва да отбележим, че при задача за формализиране на правилото за съгласуване на прилагателно със съществително от среден род първата предложена версия на програмата бе формулирана като процедура на Лого по следния начин:

ЗА СЪГЛАСУВАНИ :СЪЩ :ПРИЛ  
АКО НЯКОЕОТ(ПОСЛЕДЕН :СЪЩ)= „О (ПОСЛЕДЕН :СЪЩ)= „Е ~  
ИЗХОД ОБЕДИНЕНИЕ ДУМА :ПРИЛ „О :СЪЩ  
КРАЙ

Или казано на малко „по-естествен“ език: *Ако съществителното завършва на някоя от гласните о или е, към формата на прилагателното от мъжки род да се добави о.* Действително процедурата работеше много добре с първите зададени думи:

ИЗВЕДИ СЪГЛАСУВАНИ "СЕЛО "ХУБАВ

**ХУБАВО СЕЛО**

ИЗВЕДИ СЪГЛАСУВАНИ "ПОЛЕ "ГОЛЯМ

**ГОЛЯМО ПОЛЕ**

Но при следния пример се видя, че не е универсална:

ИЗВЕДИ СЪГЛАСУВАНИ "ДЯДО "СТАР

**СТАРО ДЯДО**

И това се установи от ученик, който имаше за задача да тества процедурата с произволни входни думи. Това наложи добавяне на информация в програмата за рода на съществителното.

Видяхме как могат да се моделират прости изречения. Сега ще се спрем на въпроса за генериране на някои форми, които традиционно трудно се подават на формализиране без някои специфични условия за структурата им.

## **2.6. Как да научим компютъра да спори**

Един бегъл поглед върху учебни материали, свързани с изучаването на темата *Синоними и антоними*<sup>4)</sup>, ни води до следното определение:

*За антоними говорим тогава, когато противопоставянето на думите е пълно, т.е. когато те заемат полюсно положение в границите на семантичното поле.*

**Пример:** думите *добър* и *лош* са антоними, но *добър* и *недобър* и *лош* и *нелош* не са антоними.

Вярно е, но какво могат да заключат учениците от горните примери за двойките думи: *прекъснат/непрекъснат, промокаем/непромокаем, зависим/независим*?

Друго е, когато впрегнат познанията си за антоними, за да научат компютъра да спори. Да разгледаме примерен диалог между ученик (играч) и компютър, в който играчът спори с компютъра и в процеса на спора получава противоположен по смисъл отговор. Когато компютърът не знае

антонима на някоя дума, играчът му помага с предложение, като по този начин обогатява *знанието* му.

Общуването започва, след като играчът въведе командата СПОР:

Компютърът: Каж ми нещо, което обичаш или мразиш.

Играчът: *Аз обичам сладко.*

Компютърът: **Аз мразя сладко, аз обичам горчиво.**

Компютърът: Каж ми нещо, което харесваш или недолюбваш.

Играчът: *Аз харесвам математиката.*

Компютърът: **Аз недолюбвам математиката, аз харесвам физкултурата.**

Компютърът: Каж ми нещо, което обожаваш или мразиш.

Играчът: *Аз обожавам лятото.*

Компютърът: **Аз мразя лятото. Каж нещо противоположно на лято.**

Играчът: Зима.

Компютърът: *Аз обожавам зимата.*

Естествено, играчът трябва да играе *честно*, защото от него зависи какви антоними компютърът ще научи. Структурата на фразите в спора е фиксирана и компютърът реагира, като замества думата *обичам* (*обожавам*, *харесвам*) със съответни техни антоними. Ако компютърът *знае* антонима на обекта на чувствата на играча, той го включва в отговора си. В противен случай пита играча за дума с „противоположно“ значение. Ако в следващата игра играчът използва същата дума, компютърът *знае* как да отговори благодарение на обогатения си от играча речник. Такъв подход е важен и от психологическа гледна точка, защото учениците са тези, които учат компютъра и обогатяват знанията му!

Реализация на този сценарий в компютърния микросвят ФИЛОЛОГО изглежда така:

За СПОР :ФРАЗА

ИЗВЕДИ (ОБЕДИНЕНИЕ "АЗ АНТОНИМ\_НА ВТОРИ :ФРАЗА ПОСЛЕДЕН :ФРАЗА)

ИЗВЕДИ (ОБЕДИНЕНИЕ "АЗ ВТОРИ :ФРАЗА АНТОНИМ\_НА ПОСЛЕДЕН :ФРАЗА

КРАЙ

Операцията АНТОНИМ\_НА ПОСЛЕДЕН търси в списък от антоними, предварително създаден в програмата, думата, която е антоним на дадената.

Например резултатът от

АНТОНИМ\_НА "БАХ [[ЛЯТО ЗИМА] [СЛАДКО ГОРЧИВО]  
[МАТЕМАТИКА ФИЗКУЛТУРА][БАХ КУИЙН]]

ще бъде КУИЙН,

а резултатът от  
АНТОНИМ НА "МАТЕМАТИКА [[ЛЯТО ЗИМА] [СЛАДКО ГОРЧИВО]  
[МАТЕМАТИКА ФИЗКУЛТУРА][БАХ КУЙН]]  
ще бъде ФИЗКУЛТУРА.

Разбира се, за да звучи спорът на правилен български език, резултатът във втория случай трябва да се членува. Това става, като се добави членът ТА (чрез операцията ДУМА за конкатенация на думи). В общия случай програмата СПОР трябва да включва проверка дали резултатът е съществително нарицателно име и от кой род е.

Именно редактирането на програмата с цел да се генерират правилни от езикова и смислова гледна точка изречения, е процес, в който учениците прилагат знанията си в дейност, която е в основата на изкуствения интелект – компютърът да води диалог с потребителя и да дава адекватни в конкретен контекст отговори.

### 2.5. Моделиране на афоризми с фиксирана структура

Идеята за генериране на езиков фрагмент във вид на *спор* може да се обогати с моделиране на афоризми с определена структура. Да разгледаме например следната мисъл на Карл Сандбърг:

*Поезията е постигане на синтез между зюмбюл и бисквита.*

При моделиране на този афоризъм трябва да направим списък от думи, които две по две нямат противоположно значение, но нямат и нищо общо наглед, например:

[[зюмбюл бисквита][брилянт локва][свиня ябълка][печка лале][куче шкаф]  
[ухо балон]~

[вентилатор кокиче][компютър сонет]]

Ето как изглежда една програма, написана на Лого:

ЗА СИНТЕЗ

НАПРАВИ "ДВОЙКА ИЗБРАН ОТ [[зюмбюл бисквита][брилянт локва]~

[свиня ябълка][печка лале] [куче шкаф]~

[вентилатор кокиче][компютър сонет]]

(ИЗВЕДИ [поезията е като синтез на] ПЪРВИ :ДВОЙКА "с ~

ДУМА ПОСЛЕДЕН :ДВОЙКА ".)

КРАЙ

Ето няколко резултата от изпълнението на програмата СИНТЕЗ:

– *Поезията е като синтез на брилянт с локва.*

– *Поезията е като синтез на компютър със сонет.*

– *Поезията е като синтез на куче и шкаф.*

В традиционното учебно съдържание учениците се сблъскват с понятия като *синоними*, *антоними*, *пароними*, *омоними*. Но едва ли ще срещнат официално название за двойки от типа на [брилянт, локва] и [куче, шкаф]. Затова пък биха могли да предложат свое название на такива двойки думи.

Впрочем в известната *Граматика на фантазията* Джани Родари<sup>5)</sup> нарича такива двойки *фантазиен бином* и {*куче, шкаф*} е една от използваните от него в клас двойки думи за разпалване на фантазията на учениците:

*Нужно е двете думи да са разделени от известна дистанция, едната да е достатъчно чужда за другата, та съседството им да е малко нещо необичайно — само така въображението ще бъде принудено да се задвижи и да установи между въпросните думи сходство, да изгради единна, в дадения случай — фантастична цялост, в която двата разнородни елемента биха могли да съжителстват.*

А дали има клас от двойки думи, за които не е известен опит да бъдат категоризирани? Това също може да е интересен въпрос за изследване, който възниква, ако решим да моделираме друг афоризъм, този път на Георг Кристоф Лихтенберг:

*Конят заприлича на магаре — като превод от немски на холандски език.*

Ако анализираме този афоризъм, ще видим, че идеята за това колко се губи при превода, може да се предаде чрез сравнението на двойка обекти, първият от които е подобен на втория, но значително го превъзхожда според някакви критерии. Следователно, ако искаме да направим вариации на горния афоризъм, трябва да създадем асоциативен списък от подобни двойки, например:

[[кон магаре] [ария чалга] [сметана айрян] [вино оцет] [диня тиква] [заседание пазар]~ [процес сеир][конференция панаир]]

Тогава една възможна процедура за генериране на вариации на афоризма на Лихтенберг ще изглежда така:

ЗА АФОРИЗЪМ

НАПРАВИ "А ИЗБРАНОТ [[Конят магаре] [Арията чалга]~

[Сметаната айрян][Виното оцет]~

[Динята тиква][Заседанието пазар]~

[Процесът сеир] [Конференцията панаир]]

НАПРАВИ "Е ИЗБРАНОТ[[немски холандски][френски италиански]~

[български македонски][руски български]]

(ИЗВЕДИ ПЪРВИ :А [заприлича на] ПОСЛЕДЕН :а "-)

(ИЗВЕДИ [като превод от] ПЪРВИ :Е "НА ДУМА ПОСЛЕДЕН :Е ".)

КРАЙ

Да чуем сега няколко компютърни афоризма:

– *Динята заприлича на тиква — като превод от руски на български.*

– *Арията заприлича на чалга — като превод от български на македонски.*

– *Конференцията заприлича на панаир — като превод от френски на италиански.*

На последния афоризъм можем да се засмеем от сърце, защото италианския език свързваме несъзнателно с говорене на висок глас, оживление, жестикулиране. Там, където участва *македонски*, ще почувстваме необходимост от

автоцензура, а в случая на *руски* и *български* може дори да се почувстваме засегнати. Това показва колко пристрастен е всъщност слушателят и от колко много фактори зависи успехът на шегата...

Създаването на асоциативни списъци от *близки* или *далечни* по смисъл думи би могло да бъде интересен елемент от учебна дейност, свързана с естествените езици и информатиката. При това дали изборът на двойките думи е подходящ в конкретен контекст, ще се оцени допълнително при разглеждане на генерираните афоризми от горните видове. От друга страна, как да дефинираме „семантичната близост“ в контекста на обработката на естествени езици и на изкуствените невронни мрежи, се оказва основен елемент от моделирането в тези научни области.

## 7. Заключение

Изследователският подход в учебния процес има за цел да се изградят у учениците такива ценни компетентности като *планирането на експеримента, изграждането на хипотези, построяването на модели, тестването* им и *изграждането на следващи, усъвършенствани модели* на изследваното явление. При това моделирането на системи от реалния свят (каквито са езиковите явления) и изследването на вътрешното поведение на абстрактния модел (тестването и евентуалното подобряване на модела) са приложими и в друг контекст. Със същите програмни средства, които са използвани при моделиране на *клюката* и *афоризмите*, учениците могат да моделират други миниатюрни езикови форми – петстишия, малки приказки, *тъпи* вицове.

Ключовата роля на технологиите е именно да се помогне на учениците да формулират идеите си, да ги подложат на проверка и да ги усъвършенстват. За учениците става все по-важно умението да формулират оригинални идеи, за да могат по-късно да отговорят на въпроси, които днес нито преподавателите им, нито те предполагат, че ще възникнат.

Опитът от работата с ученици от различни възрасти и с учители по информатика и български език ни прави оптимисти, че изследването и моделирането на езикови явления може да играе решаваща роля при мотивацията на учениците да усъвършенстват езиковата си култура, наред с инициативи като *Написаното остава. Пиши правилно!*<sup>6)</sup>.

## NOTES/БЕЛЕЖКИ

1. Текстът е представен по време на форума „Изследователски подходи в обучението по български език“ (София, 2016 г.). Организатор на събитието е Институтът за български език – Българска академия на науките.

2. Merriam-Webster, <https://www.merriam-webster.com/dictionary/prove> (1/07/2017).
3. Български език и литература за VII клас – външно оценяване, <http://u4ili6teto.bg/mod/page/view.php?id=3235> (1/07/2017).
3. Антоними, [http://balgarskiezik.blogspot.com/2010/10/blog-post\\_1916.html](http://balgarskiezik.blogspot.com/2010/10/blog-post_1916.html).
4. <https://chitanka.info/book/2759-gramatika-na-fantazijata> (10/7/2017)
5. <http://ibl.bas.bg/napisanoto-ostava-pishi-pravilno/>.

## REFERENCES/ЛИТЕРАТУРА

- Goldenberg, E. P. & Feirzeig, W. (1987) *Exploring Language with Logo*, The MIT Press, Cambridge, MA.
- Kenderov, P., Sendova, E. & Chehlarova, T. (2012) IBME and ICT – the experience in Bulgaria. In: Baptist, P., Raab, D. (Eds) *Implementing Inquiry in Mathematics Education*, Bayreuth, pp 47 – 54.
- Goldenberg, E. P. (1991) Linguistics, science and mathematics: modeling by computer. *Mathematics and Informatics Education*, 6. (translated in Bulgarian)
- Sendova, E. (2017) Constructionism as an educational philosophy and a culture – a tribute to Seymour Papert, in *Mathematics and Education in Mathematics*, Proceedings of the Forty-sixth Spring Conference of the Union of Bulgarian Mathematicians Borovets, April 9 – 13, pp 29 – 51, ISSN 1313-3330 (in Bulgarian).
- Papert, S. (1996) *The Connected Family: Bridging the Digital Generation Gap*, Longstreet Press.
- Nikolov, R. & Sendova, E. (1984). Language and Mathematics. Logo. Experimental textbook for 6<sup>th</sup> grade, Research Group on Education, Sofia.
- Sendova, E. (2013) Assisting the art of discovery at school age – a Bulgarian experience. Third chapter of: *Talent Development around the World* (coordinator Pedro S´anchez-Escobedo, Merida Yucatan), pp 39 – 98.
- Sendova, E. (1988) *Logo for philologists - handbook for teachers*. Research Group on Education, Sofia (In Bulgarian).
- Sendova, E. (2000) How to teach the computer to generate gossips and dull jokes, *Mathematics and Informatics*, 5, 6 (In Bulgarian).
- Dicheva, D., Nikolov, R. & Sendova, E. (1997), *Informatics in Logo Style*, Prosveta, Sofia (in Bulgarian).
- Dicheva, D., Nikolov, R. & Sendova, E. (1997) School informatics in Logo style: a textbook facing the new challenges of the Bulgarian informatics, In: Turcsanyi-Szabo, M. (Ed.) *Learning and Exploring with Logo*, Proceedings of the Sixth European Logo Conference Eurologo'97, Budapest, Hungary, 20 – 23 August, pp. 234 – 239.

- Sendova E. (2000) *Modelling of creative processes in abstract art, poetry and music*, IJ ITA vol. 8 No. 3, , pp. 122 – 133.
- Sendova, E. (2008) Materializing model-like hypotheses in language, poetry and humor by means of a computer microworld, *Etudes cognitive*, 8, SOW, Warszawa 2008.

## GENERATING LANGUAGE FRAGMENTS IN A COMPUTER MICROWORLD IN SUPPORT OF THE INQUIRY BASED LANGUAGE EDUCATION

**Abstract.** The paper deals with ideas for exploring language in a specialised Logo microworld in an inquiry-based style. Specific examples for modeling language fragments, developed in the frames of the Research Group of Education (1978 – 1999) as elements of the *Language and Mathematics* integrated subject for 5 – 7 graders, are considered as the roots of an innovative approach to integrating informatics with mathematics and natural languages in secondary school.

In concordance with the growing interest to the inquiry-based approach not only to science education but also to mathematics, informatics and language education, we argue that in an appropriately designed microworld for exploring language, the students can act like researchers.

By developing computer models of language phenomena the students materialise hypotheses about these phenomena, check the models' behavior with various input data and verify how general these models are. The feedback the students get after running their computer models helps them to *edit their ideas*, i.e. to correct, modify, refine and enrich them. This approach, typical for scientific research, turns out to be an important factor for the motivation of students to study natural languages in depth.

✉ **Dr. Evgeniya Sendova, Assos. Prof.**

Institute of Mathematics and Informatics

Bulgarian Academy of Sciences

Akad. G. Bonchev St., bl. 8

Sofia, Bulgaria

E-mail: jenny.sendova@gmail.com