

УДК 81'243:371.315

Ю.К. ТАРАНЕНКО,

доктор технічних наук, професор кафедри прикладної лінгвістики та методики навчання іноземних мов Дніпропетровського університету імені Альфреда Нобеля

О.Б. ТАРНОПОЛЬСКИЙ,

доктор педагогічних наук, професор, завідувач кафедри прикладної лінгвістики та методики навчання іноземних мов Дніпропетровського університету імені Альфреда Нобеля

М.М. ГАЙДАР,

викладач кафедри прикладної лінгвістики та методики навчання іноземних мов Дніпропетровського університету імені Альфреда Нобеля

Т.В. ЄРМОЛАЄВА,

старший викладач кафедри педагогіки та психології Дніпропетровського університету імені Альфреда Нобеля

ПРОЕКТНА ТЕХНОЛОГІЯ ЯК ОСНОВА ФАХОВОЇ ПІДГОТОВКИ ПРИКЛАДНИХ ЛІНГВІСТІВ

У цій статті на конкретних прикладах використання проектної технології навчання фахівців з «Прикладної лінгвістики» розглянуто систему навчання, основою якої є поетапне виконання фахових завдань практичного спрямування із зростаючою складністю. Проектна технологія навчання орієнтована насамперед на засвоєння систем автоматизованої обробки текстів з елементами самостійного програмування інструментальних засобів аналізу текстів, складність та продуктивність яких постійно зростає.

Ключові слова: проектна технологія прикладна лінгвістика, Natural Language Toolkit (NLTK) – аналіз текстів природною мовою (NLTK), автоматизований морфологічний, семантичний та синтаксичний аналіз, орієнтована на проведення лінгвістичних досліджень мова програмування Python.

Вступ. У сучасних умовах розвитку суспільства якісно змінилися засоби мовної комунікації. Обсяги інформації, яку конче потрібно переробляти у реальному масштабі часу настільки підвищилися, що класичні методи обробки мовної та текстової інформації вже не працюють. Виникла гостра потреба інтегрувати у лінгвістику методи та технічні засоби опрацьовані у інших науках, насамперед інформатики та комп'ютерної техніки з ухилом на розробку спеціального програмного забезпечення. Це спонукало до активного пошуку нових методів професійної підготовки прикладних лінгвістів, серед яких одним з основних є метод проектних технологій. Слід зауважити, що проектний метод нараховує не одне століття. Так проект, у перекладі з латинського «кинутий вперед», завжди передбачає побудову іншої або зміну існуючої реальності. Але при застосуванні проектної методики у підготовці прикладних лінгвістів виникає ряд суттєвих протиріч, які необхідно вирішити.

1. Лінгвіст не може бути програмістом в класичному розумінні, тому що, по-перше, це інший фах, а по-друге, такої потреби у реаліях прикладної лінгвістики немає, але при цьому

він повинен мати мінімально необхідний обсяг знань у галузі мови програмування, орієнтованої на проведення лінгвістичних досліджень, та проводити такі дослідження.

2. При викладанні та використанні проектних методик з дисциплін пов'язаних з обробкою текстів, таких як лінгвостатистика, автоматизований морфологічний, семантичний та синтаксичний аналіз, машинний переклад, інформаційно-пошукові системи, вкрай необхідно мати спільний інструментальних засіб, на якому можна продемонструвати основні принципи їх роботи та навіть розробити нові принципи з врахуванням того, що лінгвіст, на відміну від фахового програміста, добре знає синтаксис та морфологію мови.

3. Існуючі інструментальні засоби автоматизованого аналізу текстів запрограмовані різними штучними мовами програмування здебільшого не орієнтовані на проведення саме лінгвістичних досліджень [1].

Таким чином, виникає певна проблема застосування проектних технологій при фаховій підготовці лінгвістів.

Метою цієї статті є вирішення проблеми застосування проектних технологій, здатних суттєво підвищити якість фахової підготовки прикладних лінгвістів.

Як це не дивно, але перший крок у вирішенні наведеної проблеми був зроблений ще голландцем Гвідо ван Россумом [2], який розробив достатньо молодю мову програмування Python. Python має досить простий синтаксис. Читати код цією мовою програмування легко, тому що в ньому використовується мінімум допоміжних елементів, а правила мови змушують програмістів робити відступи. З іншого боку, Python – це повноцінна, можна сказати, універсальна мова програмування. Це мова так званого надвисокого рівня: він підтримує об'єктно-орієнтоване програмування (насправді він і розроблявся як об'єктно-орієнтована мова). Також Python поширюється вільно під ліцензією GNU General Public License. Все це робить Python дуже привабливим для використання у проектах, а з урахуванням браку коштів у більшості навчальних закладів і економічно обґрунтованим.

Але по-справжньому Python став орієнтованим на лінгвістичні дослідження після виходу бібліотеки *Natural Language Toolkit (NLTK)*. Наведемо опис даного пакета у Вікіпедії [3] дослівно: «NLTK добре підходить для студентів, що вивчають комп'ютерну лінгвістику або близькі предмети, такі як емпірична лінгвістика, когнітивістика, штучний інтелект, інформаційний пошук і машинне навчання [4]. NLTK з успіхом використовується як навчальний посібник, інструмент індивідуального навчання і платформа для прототипування і створення науково-дослідних систем. NLTK є вільним програмним забезпеченням».

З'ясувавши можливість Python, дослідимо його застосування для використання проектних технологій при викладанні дисципліни «Лінгвостатистика» для студентів другого року навчання, після вивчення дисципліни «Основи програмування на Python» протягом двох триместрів.

Завдання на курсове проектування складається з двох частин. Перша стосується вимог до програмного забезпечення, яке потрібно розробити. Друга – дослідженню, яке за допомогою створеного програмного забезпечення треба здійснити. Перше завдання формулюється на підставі вже вивчених студентом розділів дисципліни «Основи програмування на Python». Друга частина стосується вивченого матеріалу з лінгвостатистики [1,5]. Наприклад:

1. Використовуючи модуль **tkinter** побудувати графічний інтерфейс за вибором відмет на свій розсуд, для здійснення операцій підрахунку у тексті (текст вводиться за допомогою діалогового вікна) англійською мовою іменників, прикметників і дієслів. Для морфологічного розбору використати такі бібліотеки NLTK: саму бібліотеку **import nltk; text = nltk.word_tokenize(«...»); теги обрати за таблицею (додається); цикл пошуку for записати з мінімальною втратою програмного часу**.

2. Обрати художній твір англійської прози одного автора. Зробити десять вибірок по 500 слів кожна (тільки з авторської художньої мови без мови персонажів твору). За допомогою створеної програми підрахувати частотні ряди: іменників, прикметників і дієслів та отримати їх середні частоти. Ту ж саму процедуру здійснити з твором іншого автора. На підставі порівняння середніх частот зробити висновок.

При виконанні подібних курсових робіт, треба застосувати принцип змагальності. Для цього в першу половину завдання треба ввести заміри часу роботи програми з використан-

ням спеціальної функції **timeit**. Це дуже ефективний засіб проти запозичень, тому що час роботи розроблених програм впливатиме на оцінку. Обов'язковим при виконанні проєктів є вимога до єдиної форми пояснювальної записки процедури публічного захисту проєкту. Наявність методичних вказівок з курсового проєктування. Аналіз наведеного прикладу проєкту свідчить про його високу ефективність порівняно з іншими формами навчання тому, що у ньому органічно поєднуються і засвоєння теоретичних знань і надбання практичних навичок.

Розглянемо тепер дисципліну, де курсове проєктування, реалізоване з використанням Python, має не тільки навчальне, а і практичне значення. Дисципліну «Автоматизований семантичний аналіз» студенти спеціальності «Прикладна лінгвістика» вивчають на четвертому курсі, коли вивчення функціонального програмування на Python закінчено. Ця обставина дозволяє найбільш повно реалізувати можливості мови Python.

Перша частина завдання на курсовий проєкт має такий зміст:

1. Встановлення Python 3.4 і підключення необхідних модулів.
2. Розробка інтерфейсу автоматизованого семантичного аналізу з використання бібліотеки модулів tkinter, вбудованої в Python.
3. Розробка процедур для виключення з аналізованих слів стоп-слів, визначення розподілу решти слів за документами, побудова на цій основі матриці слово-документ у середовищі програмування Python.
4. Розробка процедур для перевірки заповнення матриці слова-документи та виключення з неї тих документів, які не пов'язані з відфільтрованими словами в середовищі програмування Python.
5. Розробка процедур для нормалізації отриманої частотної матриці методом TF-IDF з подальшим її сингулярним розкладанням.
6. Зіставлення результатів сингулярного розкладання, виконаного в Python, з результатами математичного пакета Matlab.
7. Визначення семантичних координат заданого ключового слова в середовищі програмування Python.
8. Розробка процедур для графічного аналізу стану ключових слів і документів з подальшим їх виділенням в групі засобами графічного пакета Matplotlib в середовищі програмування Python.
9. Порівняння результатів роботи розробленої програми з результатами аналогів.
10. Тестування програми за заданими критеріями швидкодії та надійності.

Треба звернути увагу на те, що нового з'явилося у першій половині завдання на курсове проєктування для студентів випускного четвертого курсу. Пункт перший свідчить про те, що тепер студент перетворюється у користувача, і всі його дії зі встановлення та налаштування програмного забезпечення він буде виконувати сам на власному робочому місці без порад викладача. Завдання презентуються у більш загальному вигляді, що надає можливості до творчого підходу. Самі вимоги до програмного забезпечення наближаються до вимог реальної практики – це і порівняння з аналогами, і перевірки, і процедура тестування.

Змінюється зміст і другої половини курсового завдання, а саме.

1. Студент використовує розроблене ним програмне забезпечення на місці проведення практики.
2. На вимогу керівника практики за місцем практики вносить необхідні зміни в програмне забезпечення.
3. Можлива з консультуванням керівника практики за місцем навчання розробка нового програмного забезпечення для потреб за місцем практики.
4. Практикант веде журнал практики, в якому відображає всі особливості роботи розробленого ним програмного забезпечення.
5. Оцінювання курсової роботи враховує результати практичного застосування розробки.

Для того, щоб оцінити рівень засвоєння функціонального програмування на Python, розглянемо результати наведеного курсового проєкту для іспанської мови.

```
Python 3.4.3: settings25.py - C:\Users\TJK\Desktop\ACCLISA\settings25.py
File Edit Format Run Options Windows Help

docs=[
" El aumento entra en vigencia el 1° de noviembre, por lo que el incremento deberá pagarse en la quin
" El nuevo espacio televisivo conducido por Jorge Rodríguez, 'La Política en el Divan', se estrenó este
" Representantes del dúo venezolano se encuentran en conversaciones para concretar una nueva fech
" Secretaría de gobierno de la alcaldía afirma que seis cooperativas ofrecen el servicio. Parientes de c
" La familia Marín está a cargo de la Fundación Casa de Ana en Club de Campo, donde cuidan a 37 nif
" El ministro González López informó que Rosales decidió entregarse a la justicia venezolana y se cur
" Representantes de varios consejos comunales trancaron ayer la vía. Aseguraron que la línea sólo tie
" Con los dividendos de la empresa Petro San Félix, se culminó el pago de la deuda de Sidor a los acc
" La versión coreográfica de Claudia Capriles se podrá ver el sábado 24 y el domingo 25 de octubre e
" Cerca de 20.000 personas continúan por tercer día consecutivo sin suministro eléctrico debido a un a
" El cargamento de reses que era trasladado al país y se hundió a principios de este mes, comenzó a
" Desde inicio de 2013, el sueldo mínimo ha aumentado 356,97% tomando en cuenta el nuevo increm
" El presidente Nicolás Maduro decretó por primera vez desde la aprobación de la Constitución del 19
" La directora del ente comicial en la entidad, María Lineros, informó este jueves que se habilitarán dos
" Aseguran que los productos autorizados hasta la fecha con dichas características pueden ser perfec
" La segunda ronda de reuniones entre la organización internacionales y ambos países se retomará en
" El Primer Mandatario Nacional instó a la población a profundizar el socialismo ante los planes de la oli
" Muñoz apuntó que el espacio desmantelado generaba recursos para financiar a dirigentes de la dere
" Con esta simplificación, el contribuyente podrá autogestionar de forma electrónica sus solicitudes y p
" Sin embargo, el jurado del Premio Nobel ha querido dedicar la distinción de este año a la sociedad ci
" El titular de Transporte indicó que para el mes de noviembre iniciará la fabricación de las unidades Y

stopwords=[ '-', 'El', 'en', 'por', 'que', 'de', 'la', 'a', 'donde', 'este', 'para', 'cerca', 'de', 'con', 'desde', 'sin', 'ha
wordd= 'aumento'
stem= 'spanish'
# 'danish', 'dutch', 'english', 'finnish', 'french', 'german', 'hungarian', 'italian',
# 'norwegian', 'porter', 'portuguese', 'romanian', 'russian', 'spanish', 'swedish'

Ln:1 Col:0
```

Рис. 1. Файл setting.py для налаштування роботи програми

З рис. 1 виходить, що програма може працювати на 15 мовах, що пов'язано з можливостями модуля nltk.stem, який проводить стеммінг, тобто виділяє основу слова. У файл setting.py вводяться також документи, що аналізуються у списку docs, стоп-слова, які не впливають на зміст документів у списку stopwords, змінна іспанської мови stem='spanish' та змінна ключового слова word, від якого вимірюється семантична відстань до слів та документів.

Основна форма програми має меню з вибором пункту Start, з якого починається аналіз. Первинні графічні результати наведено на рис. 2.

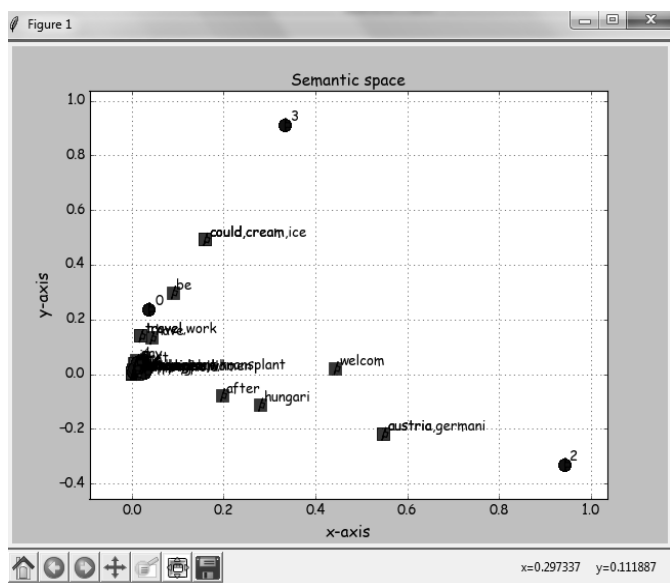


Рис. 2. Семантичний простір для розподілення документів та ключових слів за групами – це три групи, серед яких нижня розріджена, середня зі щільним накопиченням слів та документів і верхня також розріджена

Для детального розгляду щільної групи скористаємося інструментом Zoom to rectangle, отримаємо її зображення рис. 3.

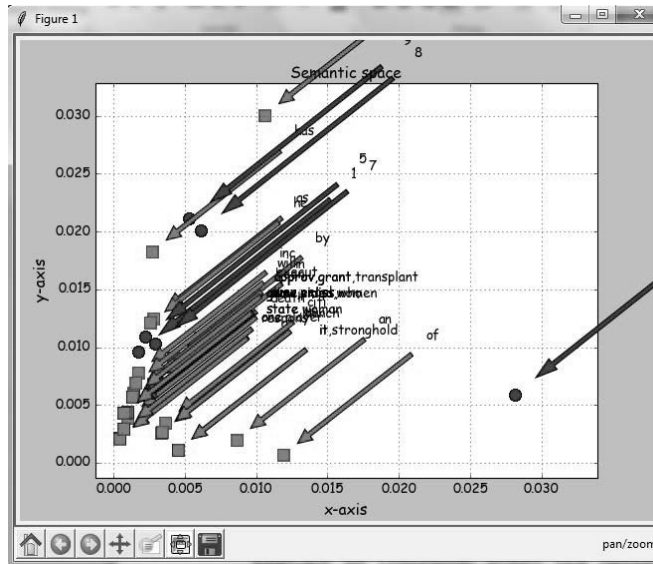


Рис. 3. Найбільш щільна група документів і слів, аналіз якої дозволяє визначити ключові слова, за якими можна здійснювати ефективний пошук певних документів у мережі Internet

При необхідності можна отримати роздруковку усіх етапів автоматизованого латентно-семантичного аналізу, яку через її великий обсяг, тут не наводимо.

Слід зауважити, що саме завдяки тривалому використанню цієї програми, була виявлена ситуація, коли внаслідок накладання передбачених методом обмежень частотна матриця методу мала пусті стовпчики. Цей випадок не був передбачений теоретично. Тому у програму була введена певна перевірка, що при появі згаданої ситуації, спонукала спеціальну циклічність обробки інформації, це, у свою чергу, дозволило коректно виключити з аналізу неінформативні документи.

Розглянемо також результати курсового проектування з дисципліни «Автоматизований синтаксичний аналіз». При вивченні дисциплін, пов'язаних з автоматизованою обробкою тексту, основну увагу було зосереджено на словах. Вирішувалися такі завдання:

1. Ідентифікація послідовності слів, n -грами, аналіз структури слова.
2. Морфологічний аналіз слів з метою встановити їх значення.

Ці методи не стосуються дослідження пропозицій як граматичної конструкції, побудованої з одного або декількох слів певної мови. Граматична конструкція містить окрему, відносно незалежну думку. Це змістова, граматична та інтонаційне ціле, що виражає якусь думку щодо її до дійсності (предикативність, створена категоріями модальності, часу і особи) одним словом чи сполученням слів.

Автоматизований аналіз вирішує завдання обробки необмеженої безлічі довільних пропозицій. Для вирішення такого складного і трудомісткого завдання не обійтися без розробки спеціального програмного забезпечення для аналізу структури пропозиції та дослідження значення цих пропозицій.

Визначимося з роллю граматики в автоматизованому синтаксичному аналізі. Почнемо зі властивості пропозицій. Менша за кількістю слів пропозиція може складатися у більшу пропозицію. Розглянемо приклади:

(1)

- a. Usain Bolt broke the record of 100m.
- b. The Jamaica Observer had reported that Usain Bolt broke the record of 100m.
- c. Andre said the Jamaica Observer had reported that Usain Bolt broke the record of 100m.

d. I think Andre said the Jamaica Observer had reported that Usain Bolt broke the record of 100m.

Замінімо перше речення на символ S, тоді наступні пропозиції будуються за шаблонами, такими як Andre said S та I think S. Ці шаблони і подібні шаблони (S but S і S when S) дозволяють на основі одного речення побудувати великі пропозиції. За подібним шаблоном побудовано величезну пропозицію в казці «Вінні-Пух».

«Winnie the Pooh» story by A. A. Milne, in which Piglet is entirely surrounded by water: [You can imagine Piglet's joy when at last the ship came in sight of him.] In after years he liked to think that he had been in Very Great Danger during the Terrible Flood, but the only danger he had really been in was the last half-hour of his imprisonment, when Owl, who had just flown up, sat on a branch of his tree to comfort him, and told him a very long story about an aunt who had once laid a seagull's egg by mistake, and the story went on and on, rather like this sentence, until Piglet who was listening out of his window without much hope, went to sleep quietly and naturally, slipping slowly out of the window towards the water until he was only hanging on by his toes, at which moment, luckily, a sudden loud squawk from Owl, which was really part of the story, being what his aunt said, woke the Piglet up and just gave him time to jerk himself back into safety and say, «How interesting, and she did?» when – well, you can imagine his joy when at last he saw the good ship, Brain of Pooh (Captain, C. Robin; 1st Mate, P. Bear) coming over the sea to rescue him...

Цій пропозиції відповідає проста структура, починаючи з S S but when S. З цього прикладу можна зробити висновок, що мові властиві конструкції, які дозволяють умовно нескінченно розширювати пропозицію. Крім того, ми можемо зрозуміти пропозицію довільної довжини, про яку раніше ніколи не чули. Досить легко придумати зовсім нову пропозицію, яка ніколи раніше в даній природній мові не зустрічалася, але всі носії мови її зрозуміють.

Мета граматики – дати опис природної мови. Щоб описати мову потрібно визначити, що вважати природною мовою, і вивчити основні підходи до її подання.

Розглянемо формальне подання відтворюючої граматики, згідно з якою мова представляється як множина всіх граматично правильних речень, а граMATика – це формальна система, яка може бути використана для генерації елементів цієї множини.

Практичні роботи найбільш ефективні саме при розгляді простих граматик, тому що дозволяють покращити мовну підготовку студентів-лінгвістів.

Для кращого розуміння завдань проектування розглянемо таблицю основних лексичних категорій (табл. 1).

Таблиця 1

Синтаксичні категорії

Символ	Значення	Приклад
S	sentence	<i>the man walked</i>
NP	noun phrase	<i>a dog</i>
VP	verb phrase	<i>saw a park</i>
PP	prepositional phrase	<i>with a telescope</i>
Det	determiner	<i>the</i>
N	noun	<i>dog</i>
V	verb	<i>walked</i>
P	preposition	<i>in</i>

Проаналізуємо неоднозначність виразу: ***I shot an elephant in my pajamas***. Спочатку потрібно визначити (розробити) просту граматику на підставі табл. 1:

S -> NP VP

PP -> P NP

NP -> Det N | Det N PP | 'I'

VP -> V NP | VP PP

Det -> 'an' | 'my'

N -> 'elephant' | 'pajamas'
V -> 'shot'
P -> 'in'

Для аналізу та побудови синтаксичного дерева розглянемо роботу програми, яка розробляється самими студентами у цьому курсовому проектуванні з використанням nltk Python 2.7.

Як і у попередньому проектуванні (див. рис. 1), у програмі використовується файл setting.py для налаштування роботи програми (рис. 4).

```

settings.py - C:\Users\TJK\Desktop\ACC\SA\settings.py (2.7.10)
File Edit Format Run Options Window Help
#!/usr/bin/env python
# -*- coding: utf-8 -*-
grammar="""
S -> NP VP
PP -> P NP
NP -> Det N | Det N PP | 'I'
VP -> V NP | VP PP
Det -> 'an' | 'my'
N -> 'elephant' | 'pajamas'
V -> 'shot'
P -> 'in'
"""

sen='I shot an elephant in my pajamas'.split()
hel="""
Символ    Значення          Пример
S          sentence          the man walked
NP         noun phrase       a dog
VP         verb phrase       saw a park
PP         prepositional
           phrase         with a telescope
Det        determiner        the
N          noun              dog
V          verb              walked
P          preposition       in
"""

```

Рис. 4. Файл setting.py для налаштування роботи програми

Слід зауважити, що використання файлів setting.py в усіх програмних продуктах, що розробляють студенти напряму «Прикладна лінгвістика», роблять ці програми гнучкими і такими, що легко перелаштовуються. Наприклад, у наведеному на рис. 5 файлі є можливість не тільки утримувати значну кількість граматик (qramar1..n) і відповідних їм речень (sen1..n), але і довідкові дані (hel 1..n), що робить програмні продукти зручними для використання.

SA	Symbol	Meaning	Example
	S	sentence	the man walked
	NP	noun phrase	a dog
	VP	verb phrase	saw a park
	PP	prepositional phrase	with a telescope
	Det	determiner	the
	N	noun	dog
	V	verb	walked
	P	preposition	in

```

S -> NP VP
PP -> P NP
NP -> Det N | Det N PP | 'I'
VP -> V NP | VP PP
Det -> 'an' | 'my'
N -> 'elephant' | 'pajamas'
V -> 'shot'
P -> 'in'

```

['I', 'shot', 'an', 'elephant', 'in', 'my', 'pajamas']
[(('T', 'PRP'), ('shot', 'VBD'), ('an', 'DT'), ('elephant', 'NN'), ('in', 'IN'), ('my', 'PRPS'), ('pajamas', 'N'))]

```

(S
  (NP I)
  (VP
    (V shot)
    (NP (Det an) (N elephant) (PP (P in) (NP (Det my) (N pajamas))))))

```

Рис. 5. Результати автоматизованого синтаксичного аналізу речення

Наведена вище програма дозволяє здійснити аналіз речення двома способами залежно від того, чи прийменниковий вираз *in tu rajamas* описує слона чи подію «стріляння». Результати графічного аналізу надано у вигляді дерева з відповідними гілками.

Висновок. Запропонована і досліджена на практиці проектна технологія, як основа фахової за спеціальністю «Прикладна лінгвістика» допомагає створити конкурентоспроможні умови їх працевлаштування на сучасному ринку праці України, яка інтегрується у європейський інформаційний простір. Однією з основних технологій такого проекту є автоматизований аналіз мовної та текстової інформації.

Список використаних джерел

1. Лутц М. Программирование на Python: пер. с англ. 4-е изд. / М. Лутц. – СПб.: Символ-Плюс, 2011. – Т. I. – 992 с.
2. Тараненко Ю.К. Лингвистическая статистика / Ю.К. Тараненко, О.Б. Тарнопольский. – Днепропетровск: Университет имени Альфреда Нобеля, 2014. – 165 с.
3. Тараненко Ю.К. Сборник задач по лингвистической статистике / Ю.К. Тараненко, О.Б. Тарнопольский, М.А. Снятовская. – Днепропетровск: Университет имени Альфреда Нобеля, 2014. – 80 с.
4. Bird S. Multidisciplinary Instruction with the Natural Language Toolkit / S. Bird, E. Klein, E. Loper, J. Balbridge. // Proceedings of the Third Workshop on Issues in Teaching Computational Linguistics. – ACL, 2008.
5. https://ru.wikipedia.org/wiki/Natural_Language_Toolkit.

References

1. Lutc, M. (2011). *Programmirovanie na Python* [Programming in Python]. Translated from English. 4th ed. SPb., Simvol-Pljus Publ., Vol. 1. 992 p. (In Russian).
2. Taranenko, Ju.K. Tarnopol'skij, O.B. (2014). *Lingvisticheskaja statistika* [Linguistic Statistics]. Dnipropetrovsk, Alfred Nobel University university, 165 p. (In Russian).
3. Taranenko, Ju.K., Tarnopol'skij, O.B., Snjatovskaja, M.A. (2014). *Sbornik zadach po lingvisticheskoi statistike* [Casebook on Linguistic Statistics]. Dnipropetrovsk, Alfred Nobel University university, 80 p. (In Russian).
4. Bird, S. Multidisciplinary Instruction with the Natural Language Toolkit / S. Bird, E. Klein, E. Loper, J. Balbridge. // Proceedings of the Third Workshop on Issues in Teaching Computational Linguistics. – ACL, 2008.
5. https://ru.wikipedia.org/wiki/Natural_Language_Toolkit.

В статье на конкретных примерах использования проектной технологии обучения специалистов «Прикладной лингвистики» рассмотрена система обучения, основой которой является поэтапное выполнение практических заданий по специальности с возрастающей сложностью. Проектная технология обучения ориентирована, в первую очередь, на усвоение систем автоматизированной обработки текстов с элементами самостоятельного программирования инструментальных методов анализа текстов, сложность и продуктивность которых постоянно возрастает.

Ключевые слова: проектная технология, прикладная лингвистика, Natural Language Toolkit (NLTK) – анализ текстов на природном языке (NLTK), автоматизированный морфологический, семантический и синтаксический анализ, ориентированный на проведение лингвистических исследований язык программирования Python.

In this article the system of learning is considered on the basis of particular examples of project technology used for training applied linguists. The basis of this system is the completion of practical tasks with gradually rising complexity. Project technology is aimed at mastering the systems of computerized processing of texts with the elements of independent programming of the instruments to process the texts whose complexity and productivity rise gradually.

Key words: project technology, applied linguistics, Natural Language Toolkit (NLTK) – text analysis in the natural language (NLTK), computerized morphological, semantic and syntactic analysis, Python – the language of programming aimed at linguistic research.

Одержано 12.01.2016.