

варіантів реалізації, передбачено більше можливостей для створення слів, відсутніх у базі даних, і більш точного їх інтонування.

Окремим програмним етапом синтезу є конкатенація сегментів, маркування додаткових пауз та безпосереднє звукове відтворення послідовності складів.

Оскільки на даному етапі програма синтезу вміщує невелику базу даних, це унеможливує відтворення власних назв, скорочень, абревіацій, відсутніх у начитаному диктором тексті. Через обмеженість інтонаційних моделей можуть виникати проблеми із синтезом деяких видів питань та емоційно чи стилістично забарвлених речень. Має місце також самостійна проблема омонімії. Так, при синтезі мовлення необхідно розрізняти звучання омонімічних слів. Наприклад, англійське слово *read* матиме різне звучання в залежності від вживання у формах теперішнього чи минулого часу. Слово *breathe* може бути іменником або дієсловом і в залежності від цього вимовлятися по-різному. Частотність таких омонімів є достатньо високою.

У цьому синтезаторі, як і в інших відсутній модуль семантичного аналізу тексту. Варто зазначити, що проблема семантичного аналізу тексту може значно поліпшити синтез мовлення. Наприклад, якщо до звичайного розповідного речення *'My sister goes to school every day.'* поставити різні запитання:

*What does your sister do every day?*

*Who goes to school every day?*

*When does your sister go to school?,*

то відповідь кожного разу матиме різну інтонацію в залежності від зміни позиції логічного наголосу. Без блоку семантичного аналізу в багатьох випадках неможливо автоматично визначити місце логічного наголосу у фразі чи реченні, а отже, генерувати просодію, що відповідатиме семантиці природного мовлення. З цією метою необхідним вбачається етап семантичного аналізу тексту, тобто орієнтація на значення.

#### Література

1. Копелевич Ю. Х., Цвєрава Г. К. Христиан Готлиб Кратценштейн. 1723 – 1795. – Л., 1989.
2. <http://offline.computerra.ru/1999/293/3579/>
3. Dutoit T. An Introduction to Text-To-Speech Synthesis. – Springer, 1996. – 316 p.
4. Mark Tatham, Katherine Morton. Developments in Speech Synthesis. – John Wiley and Sons, 2005. – 328 p.

*І. Біскуб*, к. філол. н.\*

Волинський державний університет імені Лесі Українки (Луцьк)

УДК 81'322.5

#### ЛІНГВІСТИЧНА ПАРАМЕТРИЗАЦІЯ ТЕКСТІВ ДЛЯ ТРЕНУВАННЯ КОРИСТУВАЧІВ У СИСТЕМІ АВТОМАТИЧНОГО РОЗПІЗНАВАННЯ АНГЛІЙСЬКОГО МОВЛЕННЯ DRAGON NATURALLY SPEAKING

*The report deals with the complex analysis of the training session texts taken from Dragon Naturally Speaking program for automatic speech recognition. The new type of software – SEGMIХ® is introduced in order to provide automatic text segmentation into words with statistics frequency calculations attached. The results of the frequency calculations are analyzed from the point of view of artificial text vocabulary simplification which helps to optimize man-machine interaction.*

Завдяки науково-технічному прогресу людство отримало автоматизовані системи управління, керування якими здійснюється за допомогою природної мови. З цією метою використовуються як усний так і письмовий варіанти мови. Керування різноманітними системами управління за допомогою природної мови є однією з актуальних проблем сучасної лінгвістики. Особової ваги набувають дослідження об'єктів на усіх мовних рівнях, оскільки інформація про інвентар та структуру цих об'єктів використовується у багатьох кібернетичних системах, що здійснюють взаємодію між людиною та електронно-обчислювальними машинами (ЕОМ).

Поява швидких ЕОМ значно збільшила потенціал інтелектуальної діяльності людини. ЕОМ довели свою здатність замінити обчислювальну працю сотень і тисяч людей, що спричинило появу нових засобів і методів, які отримали назву кібернетичних.

\* © І. Біскуб, 2006

Кібернетика за своєю генезою є суміжною наукою. До її теоретичного арсеналу належать методи і концепції математики, статистичної фізики, логіки, фізіології вищої нервової діяльності, теорії автоматів, теорії зв'язку, теорії інформатики, соціології, гносеології. Кібернетика покликана здійснювати взаємозв'язок між точними і суспільними науками. Вона є однією із наймолодших наук, існуючи близько 40 років.

Кібернетика – це наука про процеси керування і передачі інформації у машинах і живих організмах із застосуванням математичних методів.

Орієнтація на звукове мовлення при розробленні систем, які забезпечують спілкування людини й комп'ютера, спричиняє диференційований підхід до вирішення задач автоматичного розпізнавання ізольованих команд та розпізнавання зв'язного мовлення. Розв'язання цих завдань ускладнюється через існування низки факторів, а саме:

- 1) варіативність мовленнєвого сигналу, який подається на автоматичний розпізнавач;
- 2) залежність продуктивності системи від індивідуальних властивостей вимови мовця, його психологічного стану, гучності, темпу та ступеня чіткості вимови;
- 3) характеру акустичних шумів;
- 4) обсягу лексичної, синтаксичної та семантичної інформації, що потребує оброблення.

Кінцевою метою процесу розпізнавання мовлення є ведення діалогу між людиною і комп'ютером із використанням природної мови. Для використовуються спеціальні діалогові системи, наприклад, система KEAL. Ця система передбачає розв'язання цілої низки завдань, а саме запити інформації, віддавання наказів або автоматичного введення інформації. Режим діалогу реалізується шляхом постановки питання користувачем. Система працює у циклічному режимі. Кожен цикл роботи включає:

- 1) надсилання усного повідомлення;
- 2) отримання усного повідомлення;
- 3) розроблення відповіді;
- 4) виконання дії, що передбачена відповіддю;
- 5) перехід до наступного циклу.

Система KEAL складається із наступних частин: фонетичного аналізатора, детектора фонетичних слів, детектора акустичних слів, синтаксичного аналізатора, інтерпретатора аналізу, програми просодичної сегментації, програми фонетико-орфографічної транскрипції, генератора мовлення, діалогового автомата.

Фонетичний аналізатор дозволяє перейти від реального мовленнєвого сигналу до його фонетичної інтерпретації. Кожному сегменту, який виділений фонетичним аналізатором, відповідає певна множина впорядкованих висловлювань, які розташовані згідно із зменшенням вірогідності їх використання. Ця множина висловлювань має назву фонетичного спектру фрази.

Фонетичний аналіз проводиться поетапно: на першому етапі відбувається кодування акустичного сигналу в цифровий формат, яке здійснюється за допомогою вокодера; на другому етапі здійснюється сегментація звукового сигналу; на третьому етапі відбувається ідентифікація фонем.

Центральною частиною усієї системи є діалоговий автомат, який виконує дві функції:

- 1) визначає, які типи фраз мають найбільшу вірогідність бути вимовленими оператором у відповідності до контексту діалогу;
- 2) вирішує, яке спрямування буде мати діалог (запит інформації, відповіді на питання та ін.).

Досвід використання цієї діалогової системи підтверджує гіпотезу про те, що спілкування між людиною і машиною допоможе виправити помилки у розпізнаванні.

Важливою проблемою, яка спричиняє значні складності при розпізнаванні мовлення, є проблема адаптованості систем до конкретних дикторів. Нормалізація параметрів або ознак для усунення відмінностей між індивідуальним дикторським мовленням та адаптацією систем до роботи без орієнтації на конкретного диктора відбувається шляхом запровадження універсальної шкали формантних частот. Дана шкала успішно застосовується для ідентифікації стаціонарних голосних звуків. Однак її ефективність значно зменшується при різноманітних ефектах коартикуляції. Можливі шляхи розв'язання цієї проблеми передбачають вивчення формантних дистрибуцій для

ідентифікації голосних звуків із урахуванням довготи голосного, контекстно-зумовлених часових характеристик і спектральних характеристик консонантного оточення.

При створенні систем автоматичного розпізнавання звукового мовлення особливого значення набувають експерименти в галузі сприйняття мовлення. Результати таких експериментів часто покладені в основу функціонування тієї чи іншої системи. ЕОМ, що розпізнає мовлення, часто копіює не тільки функції людського вуха, що покликані аналізувати звуки, а й деякі функції людського мозку, які відповідають за запам'ятовування та генерування логічних зв'язків між елементами мови.

Однак, як відзначають дослідники, відсоток помилок при голосовому введенні інформації все ще складає 10%, у той час, як при введенні тексту він є у 10 разів меншим і складає 0,1%. Таким чином, постає питання про оптимізацію систем розпізнавання.

Висуваючи гіпотезу про відмінність функціонування природної мови як засобу спілкування між людьми та засобу ведення діалогу із комп'ютером, ми проаналізували лінгвістичні параметри текстів, які використовуються для тренування користувачів у системах автоматичного розпізнавання мовлення. Матеріалом для дослідження стала низка текстів із відомої системи автоматичного розпізнавання англійського мовлення Dragon Naturally Speaking.

За основу в нашому дослідженні ми беремо гіпотезу про те, що сучасні системи розпізнавання розглядають мовлення як знакову систему, а комп'ютер, який розпізнає мовлення, виконуватиме функції перетворювача пропускну здатності. Визначення шляхів оптимізації роботи систем базуватиметься на аналізі надлишковості мовленнєвих знаків. Надлишковість у цьому контексті ми розуміємо як наявність для однієї мовної одиниці декількох формальних моделей відповідних джерел інформації, які можна виміряти і описати за допомогою акустичних та лінгвістичних параметрів.

Аналіз наявних теорій автоматичного розпізнавання доводить, що обидва компоненти мовлення – фонетика і просодика – є однаково важливими для розпізнавання фонетичних знаків. Збільшення словника призводить до збільшення надмірності у природному мовленні та до зменшення надмірності при закодованому мовленні. Крім того, збільшення словника веде до експонентного збільшення помилок у природному мовленні та до зменшення помилок у закодованому мовленні. Таким чином, ми вважаємо, що аналіз надлишковості знакових систем допомагає оптимізувати процес розпізнавання мовлення, встановлює кращий розподіл у процесі кодування адитивних та мисленнєвих функцій.

Тексти для тренування користувачів передбачають генерування шаблонів (templates) для їх подальшого співставлення із потоком мовлення на вході до системи, індивідуалізацією та розпізнаванням на основі ймовірнісного моделювання.

Вивчення надлишковості лінгвістичних параметрів текстів у автоматичних розпізнавачах проводилося у декілька етапів. На початку дослідження була створена програма автоматичної сегментації англійських друкованих текстів SEG MIX<sup>®</sup>, яка дозволяє генерувати списки слів із зв'язного тексту із автоматичним визначенням кількості вживань слів та встановленням абсолютної та відносної частотності усіх одиниць тексту. З метою запобігання двозначності при розпізнаванні слів (наявність тире, крапок з комами, інших знаків пунктуації), програма передбачає визначення довжини слів у літерах для оптимізації побудови списків. Програма SEG MIX<sup>®</sup> також передбачає інтегрування отриманих даних у середовище Microsoft Excel та Microsoft Access для подальшої статистичної обробки.

За допомогою розробленої програми ми проаналізували 5 текстів для тренування користувачів у програмі автоматичного розпізнавання мовлення Dragon Naturally Speaking. Тематика текстів належить до тем „Інформаційні технології”, „Автоматичний синтез та розпізнавання мовлення”, „Кіберпростір”, „Комп'ютерний гумор”. Особливо показовим з точки зору статистики повторення слів є текст „Talking to Your Computer” (Instructional). Цей текст є по суті інструкцією до користування автоматичним розпізнавачем, і одночасно міні базою даних, яка створюється для покращення процесу розпізнавання шляхом утворення індивідуальних кальок для кожного окремого користувача.

Текст „Talking to Your Computer” складається із 404 слів. Фрагмент статистичних результатів застосування програми SEG MIX<sup>®</sup> для сегментації цього тексту наведений у таблиці 1.

Кількісний аналіз засвідчив, що більшість слів (156 = 39%) мають частоту вживання 2, що свідчить про намагання системи підтвердити звукову модель кожного з цих слів шляхом вживання в позиції всередині речення, а також у фінальній або ж ініціальній позиції. Відносна частота таких слів складає 0,5%.

Абсолютну частоту 3 (0,75%) мають тематично нейтральні слова, наприклад *first, another, using, just, have clearly* та ін. Слова із частотою 4 (1%) умовно поділяються на дві групи: теоретично нейтральні (наприклад *this, more, each* та ін), а також слова, що є відносно тематично забарвленими (наприклад *noise, speaker, naturally* та ін.).

Таблиця 1. Результати автоматичної сегментації тексту „Talking to Your Computer” (file: 'Source4.doc').

СЛОВО	АБСОЛЮТНА КІЛЬКІСТЬ	ВІДНОСНА КІЛЬКІСТЬ
We'd	3	0,75%
like	5	1,24%
read	3	0,75%
aloud	2	0,50%
minutes	3	0,75%
while	2	0,50%
computer	14	3,48%
listens	3	0,75%
learns	2	0,50%
speak	6	1,49%
When	9	2,24%
you've	2	0,50%
finished	2	0,50%
reading	2	0,50%
we'll	2	0,50%
make	2	0,50%
some	2	0,50%
adjustments	2	0,50%
then	4	1,00%
you'll	2	0,50%
able	3	0,75%
talk	4	1,00%
your	4	1,00%
words	8	1,99%
appear	2	0,50%
screen	2	0,50%

Слова із частотністю 5, 6, 7 (1,25 – 1,75%) також поділяються на дві вищезгадані групи (перша: *like, they, other, that* та ін; друга – *person, recognize, recognition* та ін.).

Найбільшу частотність 8 і більше мають тематично забарвлені слова, які вказують на топікальні характеристики цілого тексту, наприклад *computer* (14 – 3,48%), *speech* (12 – 2,99%), *people* (11 – 2,74%), *words* (8 – 1,99%), за винятком *what* (11 – 2,74%) та *when* (9 – 2,24%), які вживаються для ініціації інтерогативно-інформаційних речень.

Загальна статистика свідчить про те, що переважна більшість слів (79%) мають надзвичайно низьку частотність (0,5 – 1%), що сприяє оптимальному створенню великої кількості кальок на основі відносно невеликого за об'ємом зв'язного тексту, який несе, окрім усього іншого, і велике інформаційне навантаження.

Аналіз словникового складу текстів продемонстрував, що переважна більшість слів (близько 80%) належать до теоретично нейтрального лексикону сучасної англійської мови, а також до лексикону словникових визначень (за інформацією Longman Dictionary of Contemporary English (2005)). Як відомо, у словникових дефініціях електронних словників видавництва LONGMAN використовується спеціальна добірка слів *The Longman Defining Vocabulary*, яка включає близько 2000 слів для оформлення пояснення слів у словниках. Слова для цього вокабул'яру добираються на основі розрахунків

частотності, проведених на матеріалі електронних текстових корпусів, які входять до складу мережі The Longman Corpus Network. І хоча колекція з п'яти текстів є досить різноманітною за тематикою, словниковий склад текстів є відносно однорідним, що свідчить про штучну симпліфікацію добору слів при спілкування людини і машини.

Комплексний аналіз лінгвістичних параметрів текстів, що вживаються для тренування користувачів у програмах автоматичного розпізнавання мовлення засвідчив, що переважна більшість слів (близько 80%) мають надзвичайно низьку частотність вживання, що можна легко пояснити оптимізованим підбором слів для тренінгу, який досягається шляхом укладання зв'язного тексту із мінімальним повторенням слів, і, відповідно, збільшенням кількості слів, які будуть вимовлені користувачем у процесі підготовки до роботи із системою розпізнавання мовлення. Вокабуляр згаданих текстів на 95 % складається із спрощеного лексикону сучасної англійської мови, що іще раз доводить гіпотезу про спрощення стратегій спілкування при веденні діалогу між людиною і комп'ютером.

*Н. Філіппова*, к. філол. н.\*

Національний університет кораблебудування імені адмірала Макарова (Миколаїв)  
УДК 81'33

### **КОГНІТИВНІ ЗАСАДИ АНАЛІЗУ КОНЦЕПТУ „ПРИКЛАДНА ЛІНГВІСТИКА” ЯК СПЕЦІАЛЬНОСТІ**

*Тема статті пов'язана з концептуальним аналізом поняття „Прикладна лінгвістика” як основи навчального плану спеціальності “Прикладна лінгвістика” (6.030505 – „бакалавр”, 7.030505 – „спеціаліст”, 8.030505 – „магістр”).*

Актуальність зазначеної теми базується на необхідності створення науковообґрунтованої і прагматичної моделі навчання із названої спеціальності.

Як відомо, прикладна лінгвістика з'явилась наприкінці 20-х р.р. ХХ ст., а як самостійна лінгвістична дисципліна вона виокремилася у другій половині ХХ ст., хоч проблеми прикладного характеру (наприклад, створення алфавітів, систем письма, навчання іноземній мові, тлумачення й коментування текстів тощо) мовознавство розв'язувало ще з давніх-давен. Прикладна лінгвістика (ПЛ) – це така сфера мовознавства, яка займається лінгвістичним конструюванням і вирішує конкретні завдання машинної обробки мови та її використання. У широкому розумінні – це моделювання мови, мовлення, породження мови і сприйняття мови, мовного впливу.

ПЛ – це діяльність з використання наукових знань про устрій і функціонування мови в нелінгвістичних наукових дисциплінах і в різних сферах практичної діяльності людини, а також теоретичне осмислення такої діяльності.

ПЛ – це комплексна дисципліна, пов'язана з психологією, філософією, фізіологією, математикою, логікою, соціологією, інформатикою, програмуванням та ін.

Основний корпус функціональних досліджень, які проводилися у надрах ПЛ, врешті, зводилися до описання й моделювання фонетичної, граматичної, семантичної структури наукових, технічних, організаційних, усних та письмових документів, створення словників, тому одне з завдань ПЛ – це побудова формальних моделей, тип моделі вказує на шляхи алгоритмізації. Тільки за допомогою формальних моделей можна автоматизувати обробку даних.

Основний предмет ПЛ – розробка й удосконалення структурних та формальних методів аналізу й опису мови, опис (модель) проблемної області для вирішення конкретної задачі, вивчення й розробка засобів оптимізації функцій мови (якісна характеристика методів ПЛ): комунікативної функції (фатична (контактно-встановлювальна), інформаційна, впливу); соціальної функції (соціолінгвістика, мовна політика, орфографія, орфоєпія, політична лінгвістика, теорії перекладу, машинний переклад, викладання мов, інформаційно-пошукові системи, створення штучних мов, теорія кодування); епістемічної функції (мовна система як засіб збереження та передачі інформації; відображення національної самосвідомості; лексиграфія, комп'ютерна лексикографія, термінологія, термінографія, корпусна лінгвістика); когнітивної функції

\* © Н. Філіппова, 2006