

УДК 004.934

А. Н. Продеус, д-р техн. наук

РЕЧЕВЫЕ КОРПУСА: СОЗДАНИЕ И ПРОБЛЕМЫ

Проведен обзор наиболее востребованных речевых корпусов, используемых для разработки систем распознавания и преобразования речи. Выработаны рекомендации по созданию речевых корпусов зашумленной украинской и русской речи, предназначенных для тестирования и оптимизации систем обработки речи.

Ключевые слова: речевой корпус, синтез речи, распознавание речи

A. N. Prodeus, ScD.

SPEECH CASES: CREATION AND PROBLEMS

Review of the most claimed speech corpora, used at designing of speech recognition and transformation systems, is made. Offers on creation of noisy Ukrainian and Russian speech corpora, destined for testing and optimization of speech processing systems, are made.

Keywords: speech corpus, speech synthesis, speech recognition

A. M. Продеус, д-р техн. наук

МОВЛЕННЄВІ КОРПУСИ: СТВОРЕННЯ Й ПРОБЛЕМИ

Наведено огляд характеристик мовленнєвих корпусів, що користуються найбільшим попитом, застосовуваних для розробки систем розпізнавання й перетворення мови. Запропоновано рекомендації із створення мовленнєвих корпусів зашумленої української й російської мов, застосовуваних для тестування й оптимізації систем обробки мови.

Ключові слова: мовленнєвий корпус, синтез мови, розпізнавання мови

Введение

Речевые корпуса представляют собой базы данных в виде звуковых и соответствующих им текстовых файлов [1]. Звуковые файлы содержат образцы элементов речи (звуки, слоги, слова, фразы), в текстовых файлах размещены соответствующие транскрипции. Обычно в состав речевых корпусов включают вспомогательный программный инструментарий, обеспечивающий редактирование файлов, автоматизацию транскрибирования звуковых файлов и т.п. В речевые корпуса также включают файлы с разнообразной служебной информацией (инструкции по эксплуатации, исходные тексты программ и т.п.).

Актуальность проблемы создания речевых корпусов обусловлена тем, что сегодня никакие работы по созданию математического обеспечения систем коррекции речи, низкоскоростного кодирования речи, автоматического распознавания речи (АРР) не могут считаться завершёнными, если при разработке таких систем не использованы те или иные речевые корпуса, ставшие стандартами де факто [1].

Целью данной работы является обзор наиболее востребованных речевых корпусов,

предназначенных для разработки и испытаний математического обеспечения систем цифровой обработки речевых сигналов. Наличие такого обзора должно помочь построению перспективных речевых корпусов украинской и русской речи.

Организация и финансирование создания речевых корпусов

Первые речевые корпуса появились в середине 80-х годов прошлого столетия в США, их разработка финансировалась главным образом Министерством обороны. В результате этого были созданы корпуса: TIDIGITS – для тестирования систем распознавания изолированных цифр и цифровых последовательностей; Road Rally – для анализа и распознавания ключевых слов (word spotting); King Corpus – для систем идентификации говорящего (speaker recognition). Это же министерство финансировало создание корпуса TIMIT, который послужил прототипом для многих других речевых баз данных. Наконец, при той же финансовой поддержке были разработаны специализированные речевые корпуса Resource Management (RM) и Wall Street Journal (WSJ) для исследований в области распознавания слитной речи, а также Air Travel Information Service (ATIS) для исследования спонтанной

© Продеус А.Н., 2013

речи и понимания естественного языка в диалоговых системах [1].

Поскольку создание качественных речевых корпусов сопряжено со значительными организационно-финансовыми затратами (коммерческая цена речевых корпусов составляет сотни и тысячи долларов [2,3]), для решения этих задач в 90-е годы прошлого века были созданы специальные координационные центры. Среди наиболее мощных центров:

Linguistic Data Consortium (LDC) [2];

European Language Resources Association (ELRA) [3].

Основанный в 1992 г., центр LDC финансировался в соответствии с грантом от Агентства передовых оборонных исследовательских проектов (Defense Advanced Research Projects Agency – DARPA). В настоящее время центр LDC финансируется по гранту отдела информационных и интеллектуальных систем Национального научного фонда США, а также американской компанией Sun Microsystems. Коллекция центра LDC насчитывает несколько сотен речевых корпусов для 87 языков мира, в их число входят корпуса русского, эстонского, литовского, узбекского и грузинского языков, однако корпуса украинской речи в коллекции центра LDC нет.

Европейский центр ELRA, основанный в Люксембурге в 1995 г., в рамках проекта ONOMASTICA подготовил речевые корпуса для 11 западноевропейских языков. Несмотря на то, что в 1995 г. проект ONOMASTICA завершился, работы в данном направлении были продолжены благодаря финансовой поддержке программы Еврокомиссии Copernicus. В результате этого были созданы еще 8 речевых корпусов для восточноевропейских языков, в том числе и для украинского языка. Полученный набор речевых корпусов, содержащий 1783390 транскрипций для 1705653 речевых объектов, был назван базой данных ONOMASTICA-Copernicus. Корпус украинского языка, входящий в состав этого речевого ресурса и содержащий 258430 транскрипций для 251579 речевых объектов, был разработан сотрудником института кибернетики НАН Украины Е. Людовицом [3]. Корпус русского языка разработан с

участием Санкт-Петербургской компании «Одитек» [1].

Проекты по созданию речевых корпусов ведутся также во множестве зарубежных университетов, причем, как правило, это делается не в одиночку, а объединенными усилиями. Одним из примеров такого рода является Центр речевых технологий (Centre for Speech Technology Research – CSTR) Эдинбургского государственного научно-исследовательского университета (Шотландия) [4], где, совместно со специалистами американского университета Карнеги–Меллон (Carnegie Mellon University – CMU) [5], разработано несколько речевых корпусов. В частности, это система конкатенативного синтеза речи Festival, о которой подробнее будет сказано ниже, и речевой корпус АМІ. Другим примером является центр речевых технологий Маквайрского университета (Сидней, Австралия), где разработана система EMU, к развитию которой привлекались специалисты Эдинбургского университета, а также университетов Мюнхена и Киля (Германия) [6].

Корпуса для систем речевого синтеза

Корпуса английской речи. В течение последних 15 лет система конкатенативного синтеза речи Festival, разработанная в 1996 г. в Эдинбургском университете, стала стандартом де-факто и послужила отправной точкой для создания по меньшей мере трех коммерческих систем синтеза речи (фирм Nuance, AT&T и Cepstral) [4]. Система Festival предназначалась для работы с английским (Великобритания, США), испанским и валлийским языками. В проекте Festival приняли участие сотрудники американского университета Карнеги-Меллон, разработавшие систему FestVox автоматизации построения дифонных баз, что позволило существенно сократить – с нескольких месяцев до нескольких недель – время создания речевых баз данных на основе речевых корпусов. Еще один важный вклад сотрудников CMU состоял в развитии и внедрении технологии Unit Selection [7,8], отличительной особенностью которой является то, что в базах данных хранятся не уникальные дифоны, а несколько вариантов каждого из дифонов. Это позволило существенно повысить естествен-

ность звучания синтезированной речи, не прибегая к сложным методам обработки сигналов.

Корпуса русской речи. В России, в рамках работы по созданию системы синтеза русской речи «Живой голос» (VitalVoice), осуществляемой ООО «Центр речевых технологий» (ЦРТ), г. Санкт-Петербург, в основу системы было решено положить технологию Unit Selection, совместив ее с аллофонным синтезом [9, 10]. В результате этого на речевых корпусах разного объема – от полутора до восьми часов звучащей речи – в системе синтеза VitalVoice были созданы десять основных голосов (четыре мужских и шесть женских). Опыт создания голосов показал, что сбалансированного корпуса размером в полтора-два часа достаточно для довольно качественного синтеза, но для максимальной естественности звучания речи желательно увеличить объём корпуса до 6–7 часов, в зависимости от характеристик диктора.

Корпуса украинской речи. В Украине сотрудниками Международного научно-учебного центра информационных технологий и систем (МНУЦИТС) Т. Людовик и Н. Сажок созданы речевые корпуса, использовавшиеся при разработке синтезатора украинского языка для компьютерной программы «Вимова плюс», а также для телекоммуникационного проекта SMS2Voice компании Global Message Services [11–14]. Для ускорения процедуры фонемного транскрибирования использовались программа автоматической расстановки ударений и программа автоматического преобразования буква-фонема. Соответствие полученной транскрипции реальной речи проверялось вручную с помощью стандартных программ анализа речевых файлов [14], позволявших прослушивать и визуализировать речевые фрагменты.

Корпуса для систем автоматического распознавания речи

Корпуса английской речи. В дальнейшем зарубежье одним из наиболее востребованных является речевой корпус TIMIT, предназначенный как для фонетических исследований, так и для разработки и тестирования автоматических систем распознавания слитной речи в рамках американского варианта англий-

ского языка [2]. В его разработке (1988-1990 гг.) принимало участие несколько широко известных организаций и исследовательских центров: компания Texas Instruments, Inc., Массачусетский технологический институт (Massachusetts Institute of Technology - MIT), Стэнфордский исследовательский институт (Stanford Research Institute International - SRI), американский национальный институт стандартов (National Institute of Standards and Technology – NIST).

Текстовый материал TIMIT включает 6300 предложений, по 10 предложений для каждого из 630 дикторов из 8 диалектных зон США. Каждый записанный диалект представлен около 70 % дикторов-мужчин и 30 % женщин. При подборе и записи дикторов учитывались возраст, пол, рост, расовая принадлежность, уровень образования, время записи речи. Записанные предложения делятся на два вида: 1) специально сконструированные фразы, в которых наилучшим образом проявляется диалектная принадлежность диктора; 2) специальные фонетически сбалансированные предложения, обеспечивающие полное покрытие фонемного материала. В корпусе TIMIT звуковые файлы, полученные от разных дикторов, разделены на тренировочную и тестовую части.

Корпуса русской речи. Первый представительный речевой корпус ISABASE для русского языка, в значительной степени создававшийся по образцу корпуса TIMIT, был разработан в период с 1996 по 1998 гг. в Институте системного анализа (ИСА) Российской академии наук (РАН) при поддержке Российского фонда фундаментальных исследований (РФФИ) [1]. В составе корпуса ISABASE 4653 речевых фрагмента (фонетически сбалансированные предложения), зачитанных 20 мужчинами и 18 женщинами. При транскрибировании использовалось 110 монофонов. Вместе с тем, в разработке корпуса ISABASE свойственны существенные недостатки: малое (36) количество дикторов и весьма низкий технологический уровень разработки.

В 2000–2001 гг. в ИСА РАН по заказу корпорации Intel был создан самый представительный на сегодняшний день речевой корпус русского языка RuSpeech, который

может быть использован для разработки систем автоматического распознавания русской речи [1]. Общие характеристики корпуса: 237 дикторов (127 мужчин и 110 женщин), зачитывавших непрерывную речь (более 50000 предложений); 203 диктора (111 мужчин и 92 женщины), зачитывавших дискретную речь (3130 предложений). Типы аннотации: полный текст, фонетическая транскрипция, вручную размеченные слова и фонемы. При транскрибировании использовалось 114 монофонов. Помимо самой речевой базы, важным результатом проекта RuSpeech явилась отлаженная технология создания речевых корпусов, а также средства автоматизации в виде комплекса соответствующих компьютерных программ.

Санкт-Петербургской компанией ЦРТ [15] создана речевая телефонная база RUSTEN (250 фонограмм телефонных переговоров между двумя персонами), использованная, в частности, при разработке автоматизированного комплекса распознавания дикторов в телефонном канале «Трал». Для работы над еще одним проектом «Руссограф» (технология распознавания слитной устной русской речи) в ЦРТ использовался набор речевых баз данных, созданных по результатам обработки записей более чем 3000 дикторов общей длительностью около 300 часов и собранных с учетом 5 диалектных групп русского языка.

Корпуса украинской речи. Лидерами в Украине, в области распознавания и синтеза речи, являются отдел распознавания и синтеза звуковых образов Киевского Международного научно-учебного центра информационных технологий и систем, а также Украинская ассоциация по обработке информации и распознаванию образов [12]. Этими организациями для исследований в области автоматического распознавания украинской речи разработано три речевых корпуса: многодикторный речевой корпус UkReso, речевой корпус выступлений депутатов Верховной Рады Украины, а в настоящее время продолжается разработка акустического корпуса украинской эфирной речи (АКУЭР).

Речевой корпус UkReso (автор – Н.Н. Сажок) содержит свыше 30000 реализаций

слов и тысячи предложений около 100 дикторов, живущих в разных областях Украины. Реализации слов сохраняют частотные пропорции фонем и являются фонетически сбалансированными, при подборе слов также учитывались их частотные характеристики.

Речевой корпус выступлений депутатов Верховной Рады Украины содержит 67 часов речевых сигналов (400 дикторов), записанных с выхода телевизионного канала. Для корпуса характерны такие особенности: спонтанная речь, быстрый темп, эмоциональная окраска, высокое качество записи.

В настоящее время продолжается разработка корпуса АКУЭР [16], содержащего читаемую, подготовленную и спонтанную речь, в общей сложности более 260 часов аннотированной речи. Словарь корпуса содержит почти 45000 слов украинского языка и почти 50000 слов русского языка, произнесенных 1500 дикторов разного возраста, пола, социального положения и профессий. Кроме общеупотребительных слов, был создан словарь суржика (более 1700 слов), словарь территориальных и социальных диалектов (более 800 слов).

Корпуса речи, записанной в естественных условиях

Поскольку эффективность функционирования систем АРР зависит от систем передачи речевых сигналов (радио и телевидение, телефонные сети, IP-телефония) и от акустического окружения (шум различной природы, реверберация) [17], в последнее десятилетие особое внимание исследователей привлекают корпуса речи, записанной в «естественных условиях».

Корпуса английской речи. Речевой корпус NTIMIT (Network TIMIT) разработан в лаборатории Science and Technology Speech Communication Group американской телефонной компании NYNEX. Корпус NTIMIT представляет собой надстройку к корпусу TIMIT в виде речевых сигналов в полосе частот, используемой в телефонии [18], и предназначен для проведения исследований влияния телефонных линий на качество и разборчивость речи, на эффективность систем АРР. Речевой корпус NTIMIT сформирован посредством записи и оцифровки всех 6300 оригинальных сигналов корпуса TIMIT,

которые были подвержены специальным искажениям. Так, речевые сигналы корпуса TIMIT передавались по телефонным сетям десяти зон, половина из которых были зонами дальней связи. Кроме того, передаваемые сигналы подвергались искажениям с помощью системы «искусственный рот» и системы, моделирующей телефонную трубку.

Перечень двадцати речевых корпусов, разработанных в центре изучения разговорной речи (Center for Spoken Language Understanding) аспирантуры при Орегонском медицинском университете (Oregon Health & Science University) и предназначенных для исследований алгоритмов систем шумоподавления и АРР, приведен в [19]. Большинство корпусов представляют собой записи «телефонной» речи. В частности, разработанный в 2000 г. корпус «22 Language» интересен тем, что в его состав включены образцы «телефонной» речи 22 языков мира, в том числе русской речи. Записи представляют собой фиксированные выражения (например, дни недели), а также образцы слитной речи (в виде ответов, в свободной форме, респондента на вопросы, задаваемые по фиксированному сценарию). Речевые сигналы записывались для цифровых телефонных линий со следующими параметрами кодирования: 8 кГц, 8 бит, нелинейное компаундирование по μ -закону.

В упомянутый выше перечень двадцати корпусов входит и речевой корпус ISOLET, разработанный в 1990–1996 гг. и модифицированный в 2000–2002 гг. Данный корпус содержит записи отдельных звуков (26 звуков) английской речи, произнесенных 150 дикторами (75 женщин и 75 мужчин). В общей сложности корпус содержит 7800 записей, протяженностью около 1,25 часа. Параметры оцифровки сигналов: 16 кГц, 16 бит, моно.

Речевой корпус Auroga разработан в период с 1997 по 2002 гг. одноименной рабочей группой Европейского института телекоммуникационных стандартов (European Telecommunications Standard Institute – ETSI) [20]. Различные версии корпуса Auroga отличаются речевыми прототипами и способами добавления шумового фона. Параметры

оцифровки сигналов: 8 или 16 кГц, 16 бит, моно.

Англоязычный речевой корпус NOIZEUS, разработанный в 2007 г. в лаборатории обработки речи университета Далласа (Speech Processing Lab of the University at Texas at Dallas), содержит 30 предложений, произнесенных 6 дикторами (3 женщины и 3 мужчины), искаженных 8 видами реальных шумов (заимствованных из корпуса Auroga) для различных отношений сигнал-шум [21]. Параметры оцифровки сигналов: 8 кГц, 16 бит, моно.

Корпуса русской речи. Санкт-Петербургской компьютерной фирмой «Одитек» разработаны два специализированных корпуса русской «телефонной» речи для задач верификации диктора и автоматического распознавания речи [1,22]. Эта же фирма разработала речевой корпус для стационарных телефонных сетей Russian SpeechDat(E) Database, содержащий записи речи 2500 дикторов [23,24]. Параметры оцифровки сигналов: 8 бит, 8 кГц, компрессия по А-закону. Фирмой «Одитек» также создан речевой корпус русского языка в рамках европейского проекта SPEECON [25]. Записи выполнены в различных акустических средах: офис, автомобиль, общественное место (в том числе улица), дом. Всего записано 550 взрослых дикторов, 50 детей. При организации записей учтены возраст, пол, особенности диалекта. Запись производилась по четырем каналам одновременно (для четырех различных микрофонов), параметры оцифровки: 16 кГц, 16 бит.

Корпуса украинской речи. Киевским МНУЦИТС, а также Украинской ассоциацией по обработке информации и распознаванию образов разрабатывается телефонный речевой корпус [12]. Его особенности: русский и украинский языки, реальные записи для разных моделей мобильных телефонов, частота дискретизации 8 кГц, спонтанная речь. Объем корпуса близок 5 Гб, корпус пока не аннотирован.

Сравнение характеристик методов и подходов разработки речевых корпусов

Для украинских разработчиков в первую очередь интересны методы и подходы к разработке речевых корпусов, используемые

российскими специалистами. Причин тому несколько. Во-первых, Украина является двуязычной страной, поэтому при испытаниях телекоммуникационных систем целесообразно располагать корпусами как украинской, так и русской речи. Во-вторых, российские специалисты пока опережают украинских в плане разработки национальных речевых корпусов. Наконец, российские специалисты внимательно следят за появлением в дальнем зарубежье новых идей в данной области, оперативно внедряя удачные технические решения.

Примером такой оперативности является использование компанией ЦРТ, в рамках работы по созданию системы синтеза русской речи «Живой голос», технологии Unit Selection [9, 10], предложенной сотрудниками американского университета Карнеги-Меллон [7, 8]. Примерами значительного продвижения Санкт-Петербургских компаний ЦРТ и фирмы «Одитек» в области создания корпусов зашумленной речи являются упомянутые выше речевая телефонная база RUSTEN [15], специализированные корпуса русской «телефонной» речи для задач верификации диктора и автоматического распознавания речи [1,22], речевой корпус русского языка в рамках европейского проекта SPEECON [25].

Рекомендации по разработке перспективных корпусов зашумленной речи

При разработке речевых корпусов, предназначенных для отладки систем АРР, следует учитывать различие видов речи: дискретная (команды), подготовленная (чтение), спонтанная [17]. Следствием такого различия является неизбежное разнообразие алгоритмов добавления шума к исходным «чистым» речевым сигналам, вызванное различной ролью пауз в каждом из указанных видов речи.

При создании речевых корпусов, помимо требований фонетического характера, безусловно важными являются требования технического характера, а именно: полоса частот, отношение сигнал–шум, характер окрашенности шума (при использовании «искусственных» шумов), источник происхождения «естественного» шума (шумы природных явлений, шумы урбанистической среды,

промышленные шумы и т.п.), характер и степень нестационарности шумовой помехи, наличие дополнительных видов помех и искажений (реверберация, эхо, особенности частотной характеристики тракта, наличие и способ кодирования сигнала). В частности, при создании корпусов узкополосной (300–3400 Гц) «телефонной» речи можно использовать опыт разработчиков корпуса NOIZEUS, заимствовавших записи шумов речевых корпусов Aurora [21].

Наконец, при создании речевых корпусов целесообразно учитывать значительное разнообразие критериев и методов экспертизы систем автоматического распознавания и преобразования речи. В частности, при использовании объективных (инструментальных) методов экспертизы необходимо учитывать ограничения, свойственные выбранному методу экспертизы [26].

Выводы

Обзор наиболее востребованных речевых корпусов, используемых при разработке систем синтеза речи и создании систем АРР, свидетельствует, что украинскими специалистами сделан существенный вклад в решение проблемы создания речевых корпусов.

Вместе с тем, украинские специалисты пока отстают в создании корпусов речи, записанной в естественных условиях. В этой связи представляется разумным использование опыта зарубежных исследователей, кооперирующих свои усилия и обменивающихся элементами технологии создания речевых корпусов.

При разработке перспективных корпусов украинской и русской зашумленной речи следует учитывать не только разнообразие систем АРР, но и разнообразие критериев и методов экспертизы указанных систем АРР.

Список использованной литературы

1. Кривнова, О. Ф. Речевые корпуса на новом технологическом витке / О. Ф. Кривнова // Речевые технологии. – 2008. – № 2. – С. 13 – 23.
2. LDC Top Ten Corpora: [Электронный ресурс] – Режим доступа к ресурсу: <http://www ldc.upenn.edu/Catalog/topten.jsp> – 16.01.2013.

3. European Language Resources Association. ONOMASTICA-Copernicus Database: [Электронный ресурс] – Режим доступа к ресурсу: <http://catalog.elra.info/> – 16.01.2013.
4. The Centre for Speech Technology Research. The University of Edinburgh: [Электронный ресурс] – Режим доступа к ресурсу: <http://www.cstr.ed.ac.uk> – 16.01.2013.
5. Carnegie Mellon University. Open Source Speech Software from Carnegie Mellon University: [Электронный ресурс] – Режим доступа к ресурсу: <http://www.speech.cs.cmu.edu> – 16.01.2013.
6. The Emu Speech Database System. Macquarie University. Sidney, Australia: [Электронный ресурс] – Режим доступа к ресурсу: <http://emu.sourceforge.net/manual/> – 16.01.2013.
7. Black, A. W. Optimising selection of units from speech databases for concatenative synthesis / A.W. Black, N. Campbell // In: Proceedings of the Eurospeech'95. – Madrid, Spain, 1995. – P. 581 – 584.
8. Clark, R. A. G. Multisyn: Open-domain unit selection for the Festival speech synthesis system / R. A. G. Clark, K. Richmond, S. King // Speech Communication. – 2007. – Vol. 49. – P. 317 – 330.
9. Корольков, Е. А. Синтез естественной русской речи при помощи метода Unit Selection / Е.А. Корольков, И.А. Главатских, А. О. Таланов, В. В. Киселев // Санкт-Петербург: ООО «Центр речевых технологий: [Электронный ресурс] – Режим доступа к ресурсу: <http://mytts.forum2x2.ru/t261-topic> – 16.01.2013.
10. Продан, А. И. Система подготовки нового голоса для системы синтеза VitalVoice / А. И. Продан, А. О. Таланов, П. Г. Чистиков // Труды международной конференции Диалог 2010: [Электронный ресурс] – Режим доступа к ресурсу: <http://www.dialog21.ru/digests/dialog2010/materials/pdf/60.pdf> – 16.01.2013.
11. Людовик, Т. В. Розроблення мовленнєвих баз даних для синтезу мовлення за текстом / Т.В. Людовик, М.М. Сажок, Р. А. Селюх // Автоматиз. системи управління і нові інформаційні технології. – 2005. – № 3. – С. 1 – 16.
12. Розпізнавання та синтез мовлення в Україні: [Электронный ресурс] – Режим доступа к ресурсу: <http://www.speech.com.ua/developers.html> – 16.01.2013.
13. Центр Глобальних Повідомлень Україна – Global Message Services [Электронный ресурс]: Режим доступа к ресурсу: <http://www.gmsu.ua> – 16.01.2013.
14. Сажок, М. Автоматизовані засоби формування баз даних і знань для озвучення українських текстів / М. Сажок – К: АКД, 2004. – 20 с.
15. Центр Речевих Технологій [Электронный ресурс]: – Режим доступа к ресурсу: <http://speechpro.ru> – 16.01.2013.
16. Васильева, Н. Б. Корпус украинской эфирной речи / Н. Б. Васильева, В. В. Пилипенко, А. М. Радужкий, В. В. Робейко // Речевые технологии. – 2012. – № 2. – С. 12 – 21.
17. Huang, X. Spoken Language Processing / X. Huang – Prentice-Hall, 2001. – 965 p.
18. Jankowski, C. NTIMIT: A phonetically balanced, continuous speech, telephone bandwidth speech database / C. Jankowski, A. Kalyanswamy, S. Basson, J. Spitz // Proc. ICASSP-90. – 1990. – P. 109 – 112.
19. Corpora Group at CSLU [Электронный ресурс]: – Режим доступа к ресурсу: <http://www.cslu.ogi.edu/corpora/> – 16.01.2013.
20. Hirsch, H.-G. The Aurora-5 Experimental Framework for the Performance Evaluation of Speech Recognition in Case of a Hands-free Speech Input in Noisy Environments [Электронный ресурс]: – Режим доступа к ресурсу: <http://aurora.hsnr.de/background.html> – 16.01.2013.
21. Speech Processing Lab. Noisy Speech Corpus. [Электронный ресурс]: Режим доступа к ресурсу: <http://www.utdallas.edu/~loizou/> – 16.01.2013.
22. Викторов, А. Б. Речевые базы данных для задач автоматического распознавания речи и верификации говорящего / А. Б. Викторов, К. О. Викторова, А. В. Воронцова // Современные речевые технологии. Сб. трудов IX сессии Российского акустического общества. – 1999. – С. 5 – 15.
23. SpeechDat-Car data base [Электронный ресурс]: – Режим доступа к ресурсу:

<http://www.fee.vutbr.cz/SPEECHDAT-E/> – 16.01.2013.

24. Auditech.Ltd [Электронный ресурс]: Режим доступа к ресурсу: <http://www.auditech.ru/> – 16.01.2013.

25. Wideband Speech Database for Russian [Электронный ресурс]: Режим доступа к ресурсу: <http://www.auditech.ru/> – 16.01.2013.

26. Продеус, А. Н. Акустическая экспертиза каналов речевой коммуникации. Монография / А. Н. Продеус, В. С. Дидковский, М. В. Дидковская. – К. : Имэкс-ЛТД, 2008. – 420 с.

Получено 16.01.2013

References

1. Kryvnova, O. F. Speech Corpora on New Technologic Level / O. F. Kryvnova // *Speech Technology*. – 2008. – № 2. – P. 13 – 23 [in Russian].

2. LDC Top Ten Corpora [Online]. Available: <http://www ldc.upenn.edu/Catalog/> – 16.01.2013 [in English].

3. European Language Resources Association. ONOMASTICA-Copernicus Database [Online]. Available: <http://catalog.elra.info/> – 16.01.2013 [in English].

4. The Centre for Speech Technology Research. The University of Edinburgh. [Online]. Available: <http://www.cstr.ed.ac.uk> – 16.01.2013 [in English].

5. Open Source Speech Software from Carnegie Mellon University. [Online]. Available: <http://www.speech.cs.cmu.edu> – 16.01.2013 [in English].

6. The Emu Speech Database System. Macquarie University. [Online]. Available: <http://emu.sourceforge.net/manual/> – 16.01.2013 [in English].

7. Black, A. W. Optimising selection of units from speech databases for concatenative synthesis / A. W Black, N. Campbell // In: *Proceedings of the Eurospeech'95*. – Madrid, Spain, 1995. – P. 581 – 584 [in English].

8. Clark, R.A.G. Multisyn: Open-domain unit selection for the Festival speech synthesis system / R. A. G Clark, K.Richmond, S. King // *Speech Communication*. – 2007.–Vol. 49. –P. 317 – 330 [in English].

9. Korolkov, E. A. Synthesis of Natural Russian Speech by Means of Unit Selection Method / E. A. Korolkov, I. A. Glavatskih, A. O. Talanov // St-Petersburg: Ltd “Center of Speech technology”. [Online]. Available: <http://mytts.forum2x2.ru/> – 16.01.2013 [in Russian].

10. Prodan, A.I. System of New Voice Preparation for VitalVoice Synthesis System /A. I. Prodan, A.O. Talanov, P. G. Chistikov // *Proc. Intern. Conf. Dialog'2010*. [Online]. Available: <http://www.dialog-21.ru/digests/dialog2010/> – 16.01.2013 [in Russian].

11. Liudovik, T.V. Design of Speech Data Bases for Text-to-Speech Synthesis / T. V. Liudovik, M. M. Sazhok, R. A Seliuh // *Automat. Control Systems and New Information Technology*. – 2005. – № 3. – P. 1–16 [in Russian].

12. Speech Recognition and Synthesis in Ukraine [Online]. Available: <http://www.speech.com.ua/> – 16.01.2013 [in Russian].

13. Global Message Services Ukraine [Online]. Available: <http://www.gmsu.ua> – 16.01.2013 [in Russian].

14. Sazhok, M. Automatic Means of Data Base and Data Knowledge for Sounding of Ukrainian Texts / M. Sazhok. – Kyiv : AKD, 2004. – 20 p. [in Russian].

15. Speech Technology Center [Online]. Available: <http://speechpro.ru> – 16.01.2013 [in Russian].

16. Vasilieva, N. B. Corpus of Ukrainian Broadcast Speech / N. B. Vasilieva, V. V. Piliipenko, A. M. Raduckiy, V. V. Robeiko // *Speech Technology*. – 2012. – № 2. – P. 12 – 21 [in Russian].

17. Huang, X. Spoken Language Processing / X. Huang.– Prentice-Hall, 2001. – 965 p. [in English].

18. Jankowski, C. NTIMIT: A Phonetically Balanced, Cont. Speech, Phone BW Speech DB / C. Jankowski, A. Kalyanswamy, S. Basson, J. Spitz // *Proc. ICASSP-90*. – 1990. – P. 109 – 112 [in English].

19. Corpora Group at CSLU [Online]. Available: <http://www.cslu.ogi.edu/> – 16.01.2013 [in English].

20. Hirsch, H.-G. The Aurora-5 Experimental Framework for the Performance Evaluation of Speech Recognition in Case of a Hands-free

Speech Input in Noisy Environments [Online] / Hirsch H.-G. Available: <http://aurora.hsnr.de/> – 16.01.2013 [in English].

21. Speech Processing Lab. Noisy Speech Corpus. [Online]. Available: <http://www.utdallas.edu/~loizou/> [in English] (16.01.2013).

22. Victorov, A. B. Speech Data Bases for ASR and Dictor Verification Tasks / A. B Victorov, K. O Victorova, A. V. Voroncova / Modern Speech Technology. Proc. IX Session of Russian Acoustical Society. – 1999. – P. 5 – 15 [in Russian].

23. SpeechDat-Car data base [Online]. Available: <http://www.fee.vutbr.cz/SPEECHDAT-E/sample/russian.html> – 16.01.2013 [in Russian].

24. Auditech.Ltd [Online]. Available: <http://www.auditech.ru> – 16.01.2013 [in Russian].

25. Wideband Speech DB for Russian [Online]. Available: <http://www.auditech.ru> – 16.01.2013 [in Russian].

26. Prodeus, A. N. Acoustic Expertise of Speech Communication Channels. Monograph / A. N Prodeus, V. S. Didkovskiy, M. V. Didkovskaia. – Kyiv : Imex-Ltd, 2008. – 420 p. [in Russian].



Продеус
Аркадий Николаевич,
д-р техн. наук, проф. каф.
акустики и акустоэлек-
троники
Нац. технич. ун-та Украи-
ны «КПИ»,
тел.: +38067-232-2896,
e-mail:
aprodeus@gmail.com