

РАСПОЗНАВАНИЕ ДВУЯЗЫЧНОЙ РЕЧИ БЕЗ ПРЕДВАРИТЕЛЬНОЙ ИДЕНТИФИКАЦИИ ЯЗЫКА

Людовик Т. В. (tetyana.lyudovyk@gmail.com),
Пилипенко В. В. (valeriy.pylypenko@gmail.com)

Международный научно-учебный центр
информационных технологий и систем
НАН Украины и МОН Украины, Киев, Украина

Предлагается подход к дикторонезависимому распознаванию слитной украинско-русской речи, не требующий предварительной сегментации речевого сигнала на разноязычные участки и идентификации языков. Также не требуется создание специальных речевого и текстового корпусов для обучения акустической и лингвистической моделей. Подход учитывает особенности фонетических систем русского и украинского языков. Используется разработанная ранее акустическая модель украинской речи. Двужызычная лингвистическая модель обучается на множестве украинских и русских текстов. Лексикон распознавания объединяет словоформы обоих языков, при этом фонемные транскрипции русских словоформ представлены украинскими фонемами.

Предлагаемый подход применим для распознавания двужызычной речи с межфразовым и внутрифразовым языковым переключением.

Ключевые слова: двужызычная речь, языковое переключение, украинский язык, русский язык, автоматическое распознавание речи, автоматическая идентификация языка

BILINGUAL SPEECH RECOGNITION WITHOUT PRELIMINARY LANGUAGE IDENTIFICATION

Lyudovyk T. V. (tetyana.lyudovyk@gmail.com),
Pylypenko V. V. (valeriy.pylypenko@gmail.com)

International Research/Training Center for Information
Technologies and Systems, Kyiv, Ukraine

We presents an approach to speaker-independent recognition of large-vocabulary continuous speech characterized by code-switching between

Ukrainian and Russian. The approach does not require language boundary detection or language identification. Special speech and text corpora are not needed to train acoustic and linguistic models. The approach takes into account peculiarities of phonetic systems of Russian and Ukrainian languages.

A cross-lingual speech recognition system is developed. A previously developed acoustic model of Ukrainian speech serves for both languages. A set of HMM-models representing 54 Ukrainian phonemes and several non-speech units such as breath, fillers and silence are used. Bilingual linguistic model is trained on a set of Ukrainian and Russian texts. Pronunciation lexicon combines word forms in both languages. Phonemic transcription of Russian word forms are generated using Ukrainian phonemes. Recognition post-processing can be applied to smooth recognized word sequences by using a dictionary containing Ukrainian and Russian words which sound equally but are written differently.

The proposed approach can be applied to the recognition of bilingual speech with between-phrase and within-phrase code-switching.

Developed cross-lingual speech recognition system was tested on Ukrainian, Russian, and Ukrainian-Russian speech of one bilingual speaker. Preliminary results show that the proposed approach could achieve a good performance. Accuracy of mixed speech recognition is lower only by 3–7% as compared with monolingual speech recognition accuracy.

Key words: mixed speech, bilingual speech, bilingual code-switching, Ukrainian language, Russian speech, automatic speech recognition, automatic language identification

1. Введение

Большинство разработанных к настоящему времени систем автоматического распознавания речи ориентировано на правильную, нормативную одноязычную речь. Однако, часто приходится иметь дело с двуязычной речью — языковым переключением («code switching»), когда чередуются отрезки речи на разных языках, или смешанной речью («code mixing»), когда в речи появляются слова и обороты, образованные из элементов разных языков.

Проблема распознавания двуязычной речи актуальна в мировом масштабе, однако активно в настоящее время она решается в Гонконге [3, 7], Сингапуре [16, 17], Тайване [11] и Индии [1, 2].

Проблему распознавания речи с языковым переключением можно рассматривать как проблему распознавания многоязычной речи [6, 8]. В такой постановке она решается, например, в Швейцарии на материале пяти европейских языков [8].

В перечисленных работах решаются задачи разной сложности: распознавание изолированных слов узкой предметной области [8], распознавание подготовленной слитной речи [1, 2], распознавание спонтанной слитной речи [16, 17].

В речи с переключением языки могут быть «равноправными» или «примешиваемыми». Наиболее часто примешивается английский язык к китайскому

[16, 17] и хинди [1, 2]. В [11] исследуется переключение между равноправными мандаринским и тайваньским диалектами китайского языка.

В Украине наблюдается как переключение между русским и украинским языками, так и смешение этих языков. Пример межфразового переключения: «Сейчас я задам два вопроса. **Перше питання до вас**». Пример внутрифразового переключения: «В наше время **замовник** становится **загарбником**». Смешение языков может происходить на фразовом и словарном уровнях. Это явление принято называть суржиком.

Данная работа посвящена распознаванию слитной спонтанной украинско-русской речи с языковым переключением. Чередующиеся отрезки речи могут быть произвольной длительности. Дикторами могут быть как билингвы, владеющие украинским и русским языками, так и носители только одного из этих языков. При этом системе распознавания неизвестно, на каком из языков произносятся те или иные отрезки речи. Такая ситуация типична при распознавании диалогов и интервью.

2. Существующие подходы

В типичной одноязычной системе распознавания речи используются: акустическая модель (АМ), лингвистическая модель (ЛМ), лексикон распознавания и декодер [5]. АМ — это множество НММ-моделей, статистически характеризующих акустические свойства фонем. ЛМ — это статистическая модель, задающая вероятности появления пар словоформ в речи. АМ предварительно обучается на большом корпусе речевых данных, а ЛМ, соответственно, — на большом корпусе текстовых данных. Лексикон распознавания представляет собой словарь, в котором словоформы содержатся в орфографическом виде и в виде фонемных транскрипций, что позволяет устанавливать соответствие между лексическим и акустическим уровнями. В декодере обработанный входной речевой сигнал сопоставляется с информацией, хранящейся в АМ и ЛМ, и определяется наиболее вероятная последовательность словоформ, соответствующая этому речевому сигналу.

Для распознавания двуязычной речи было предложено два подхода: многопроходный и однопроходный.

Многопроходный подход заключается в нахождении границ разноязычных фрагментов, идентификации языков и использовании соответствующих одноязычных систем для распознавания этих фрагментов [14]. Этот подход имеет существенный недостаток: он в значительной степени зависит от точности нахождения границ одноязычных фрагментов и от точности идентификации языка. В случае близкородственных языков для их идентификации проблематично использовать акустическую, фонотактическую и просодическую информацию.

При однопроходном подходе работает единая двуязычная система распознавания, для которой АМ, ЛМ и лексикон строятся как для единого смешанного языка. При этом нет необходимости идентифицировать язык.

Однопроходный подход применен в [1, 2] для распознавания хинди-английской речи. Использовалась англоязычная система распознавания речи *Sphinx* и прилагающаяся к ней АМ английских фонем [4]. ЛМ была обучена на смешанноязыковых текстах, соответствующих созданному небольшому смешанному речевому корпусу (213 предложений). Объем хинди-лексикона составлял 2071 слово. Слова хинди сначала были транслитерированы, а потом фонетически затранскрибированы с использованием английских фонем. Недостающие хинди-фонемы были аппроксимированы комбинациями английских фонем.

В [16, 17] однопроходный подход использован при распознавании спонтанной речи с большим словарем и переключением между китайским и английским языками. Обучение АМ происходило на 58,4 часа речи 139 дикторов. Алфавит фонем содержит английские и китайские фонемы, а также фонемы, принадлежащие обоим языкам (обозначаемые одинаковыми символами МФА — Международного фонетического алфавита (*IPA, International Phonetic Alphabet*)). Для обучения ЛМ путем машинного перевода были созданы искусственные тексты с переключением языков. Результат распознавания контрольной выборки (10 дикторов, 2,3 часа речи) — 36,9% MER (*mixed error rate* — смешанный показатель ошибки: пословной для английского языка и иероглифической — для китайского).

В [11] однопроходный подход применен к распознаванию речи с переключением между близкородственными мандаринским и тайваньским диалектами китайского языка, использующими одну и ту же систему письменности.

В [3] разработан гибридный подход: наряду со смешанными китайско-английскими моделями используется получаемая информация о границах разноязычных участков речи. В результате достигнута точность пословного (для английского языка) и послогового (для китайского) распознавания, равная 60%.

3. Предлагаемый подход

Предлагаемый однопроходный подход учитывает особенности фонетических систем русского и украинского языков, позволяющие использовать в качестве АМ акустическую модель, разработанную ранее для распознавания украинской речи [12]. Двухязычная ЛМ может быть обучена на множестве украинских и русских текстов [15], при этом она может моделировать межфразовое переключение языков. Лексикон распознавания объединяет словоформы обоих языков; транскрипции русских слов представлены украинскими фонемами.

Предложенный подход не требует создания специальных речевого и текстового корпусов для обучения АМ и ЛМ.

4. Особенности украинской и русской фонетических систем

Украинский и русский языки являются близкородственными. Одной из целей данного исследования является выяснение того, положительно или отрицательно сказывается близкородственность на результатах распознавания речи.

Поскольку в работах, посвященных речевым технологиям, обычно не проводится четкое разделение между фонемами и аллофонами, в дальнейшем будем употреблять единый термин «фонема». Обычно для синтеза и распознавания русской и украинской речи пользуются расширенным алфавитом фонем [9, 10, 13].

Система фонем украинского языка включает все фонемы русского языка. При этом фонотактические различия языков незначительны. В первую очередь различия касаются частоты встречаемости звонких взрывных заднеязычных фонем (украинские [г] и [г'] встречаются в речи значительно реже, чем русские [г] и [г']), а также частоты встречаемости ударной гласной переднего ряда среднего подъема (украинская [е], русская [э]) в позиции после мягких согласных (укр. «*суттєво*», рус. «*комитет*»).

В русской речи украинских билингвов наблюдается фонотактическая интерференция, которая наиболее сильно сказывается на замене взрывного [г] щелевым, а также на ослаблении редукции безударных гласных.

Данное исследование опирается не только на эти общие сведения, но также и на наш опыт разработки систем синтеза украинской и русской речи [16].

Для синтеза украинской речи нами использован алфавит из 59 фонем: 12 гласных (6 ударных и 6 безударных), 45 согласных (22 пары «твердая-мягкая» и непарная [й]), пауза и гортанная смычка.

Алфавит, используемый нами для синтеза русской речи, значительно меньше, поскольку отсутствуют [г], [г'] («южнорусские»), [ц'], [ч], [дз], [дз'], [дж], [дж']. В то же время, каждая из русских фонем совпадает с одной из украинских фонем или близка к одной из них. Кроме того, на произнесении русских фонем сказывается украинский акцент, что еще более приближает русские фонемы к украинским.

Мы предполагаем, что билингвы в Украине, говоря на украинском и русском языках, пользуются одним и тем же набором фонем. Это, в свою очередь, дает основание предполагать, что при распознавании русской речи можно использовать украинскую акустическую модель.

5. Речевые и текстовые данные

Проведенное исследование базируется на данных Акустического корпуса украинской эфирной речи (АКУЕМ) [14], в котором представлена украинская и русская речь, прозвучавшая в украинском телеэфире. Количественные характеристики корпуса АКУЕМ приведены в Таблице 1. На обоих языках говорят 160 дикторов корпуса.

Таблица 1. Количественные характеристики речевого корпуса украинской и русской теле- и радиоэфирной речи

	Украинская речь	Русская речь
Длительность акустических записей	116 часов	190 часов
Общее количество словоформ	962 504	1 721 606
Количество уникальных словоформ	69 500	83 500
Количество дикторов, длительность речи которых превышает 10 сек.	1 723	2 781

Для проверки принятого подхода в качестве диктора был выбран билингв, имеющий опыт публичных выступлений. Были отобраны фрагменты его спонтанной речи на украинском и русском языках, записанной в телестудии во время трансляций двух политических ток-шоу. Все речевые фрагменты были сегментированы на моноязычные участки длительностью до 15 сек. с границами, соответствующими границам синтагм. Таким образом, исследовалась речь с межфразовым языковым переключением. Объем украинского речевого материала составил 1693 слова (856 уникальных словоформ), русского — 823 слова (503 уникальные словоформы).

Текстовый материал, использованный в исследовании, состоял из всех текстов (на украинском и русском языках) корпуса АКУЕМ. Объем текстового материала — 2,7 млн. словоформ.

6. Система распознавания украинской речи

Базовой для исследования является дикторонезависимая система распознавания подготовленной и спонтанной украинской речи [12], разработанная с помощью инструментария НТК [18]. В качестве АМ используется набор контекстно-независимых скрытых Марковских моделей, обученных на материале украинской речи корпуса АКУЕМ. Помимо акустических моделей 54 фонем украинского языка используются особые модели для паузы, звучащих пауз гезитации («э-э-э», «а-а-а», «м-м»), вдохов/выдохов, чмоканья, кашля, смеха и плача.

ЛМ представляет собой биграммную модель языка, заданную вероятностями пар словоформ. ЛМ обучена на украиноязычных текстах корпуса АКУЕМ (20 Мб) и текстах из Интернета (400 Мб).

Лексикон распознавания содержит 116 тыс. словоформ. Часто употребляемые в речи словоформы, в том числе числительные, имеют от 1 до 10 фонемных транскрипций, остальные — от 1 до 3. В среднем на одну словоформу в лексиконе распознавания приходится 1,5 фонемных транскрипций, отражающих каноническое и спонтанное произнесение. Для основной части словоформ (92%) транскрипции порождены полностью автоматически, в 7% словоформ вручную были расставлены знаки ударения (фамилии, географические названия и т. п.). Для остальных, наиболее частотных словоформ (1% от общего

объема лексикона), автоматически порожденные фонемные транскрипции были дополнены транскрипциями, написанными экспертом.

Точность пословного распознавания контрольной выборки описанной системой — 87,71 %.

7. Экспериментальная система

Наиболее трудоемким и длительным процессом при разработке системы распознавания речи является создание и обучение АМ.

Исходя из гипотезы, что все фонемы русской речи можно моделировать украинскими фонемами, в качестве АМ для распознавания речи с языковым переключением была выбрана украинская АМ.

Лексикон для распознавания речи с языковым переключением содержит 56753 украинские словоформы и 58058 русских словоформ. Количество словоформ, совпадающих по орфографическому написанию, — 4952 (например, «новенький», «проводили»); совпадающих и по написанию, и по звучанию — 257 (например, «думаю», «народ»).

Фонемные транскрипции русских словоформ были порождены транскриптором, используемым для синтеза речи. Были изменены обозначения фонем: «и» на «і», «э» на «е», «г» на «г». В Таблице 2 приведены примеры фонемных транскрипций русских словоформ.

Таблица 2. Примеры словоформ и их фонемных транскрипций

Примеры словоформ	Фонемные транскрипции на базе украинского алфавита фонем
думаю	д У м а й у
государственный	г а с у д А р с т в' е н и й
обстоятельства	а п с т а й А т' е л' с т в а

ЛМ для данного исследования была обучена на всех текстах (украиноязычных и русскоязычных) корпуса АКУЕМ. Поскольку речевые сегменты контрольной выборки также содержатся в этом корпусе, OOV = 0 (out-of-vocabulary rate — процентное отношение количества словоформ контрольной выборки, отсутствующих в ЛМ, к общему количеству словоформ контрольной выборки).

На рис. 1 представлена конфигурация экспериментальной системы распознавания слитной спонтанной украинско-русской речи с языковым переключением.

Было проведено девять экспериментов для оценки работоспособности предложенного подхода к распознаванию речи с языковым переключением. Оценивалась пословная точность распознавания. Поскольку все речевые сегменты контрольной выборки распознаются независимо друг от друга, их общее множество можно рассматривать как речь с языковым переключением.

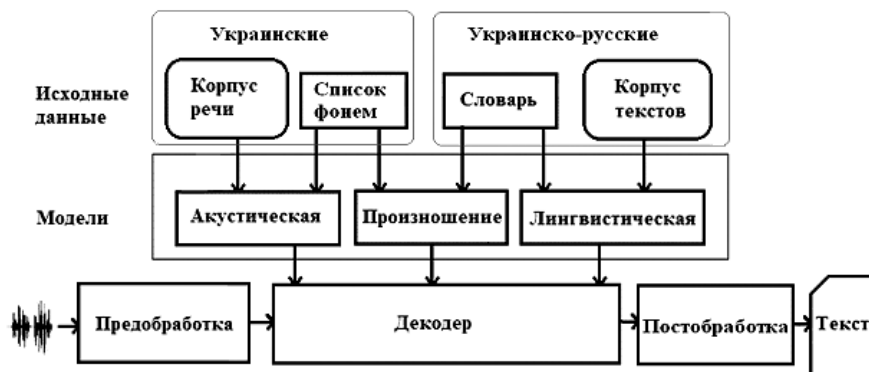


Рис. 1. Конфигурация системы распознавания украинско-русской речи

Во всех экспериментах использовалась украинская АМ. Распознавались:

1. все украинские речевые сегменты контрольной выборки;
2. все русские речевые сегменты;
3. все речевые сегменты.

Для распознавания украинской речи были разработана ЛМ, обученная на всех украинских текстах корпуса АКУЕМ. Лексикон распознавания включал только украинские словоформы из общего лексикона. Аналогично, для распознавания русской речи были разработаны ЛМ, обученная на всех русских текстах корпуса АКУЕМ и лексикон распознавания, включающий только русские словоформы из общего лексикона.

Для распознавания речи с языковым переключением использовались украинско-русская ЛМ и украинско-русский лексикон.

8. Экспериментальные результаты

Результаты экспериментов представлены в Таблице 3. Во всех экспериментах в качестве акустической модели использовалась АМ, обученная на украинской речи корпуса АКУЕМ. Обозначения лингвистических моделей:

- UKR_LM — ЛМ, обученная на всех украиноязычных текстах корпуса АКУЕМ;
- RUS_LM — ЛМ, обученная на всех русскоязычных текстах корпуса АКУЕМ;
- UKRUS_LM — ЛМ, обученная на всех текстах корпуса АКУЕМ.

ukr_lex — лексикон распознавания, включающий словоформы украинского языка;
 rus_lex — лексикон распознавания, включающий словоформы русского языка;
 ukrus_lex — лексикон распознавания, включающий словоформы обоих языков.

В «ukrus_lex» фонемные транскрипции русских словоформ представлены украинскими фонемами.

Таблица 3. Результаты распознавания украинской, русской и украинско-русской речи

Речевые сегменты	Количество слов	Точность распознавания, %		
		UKR_LM, ukr_lex	RUS_LM, rus_lex	UKRUS_LM, ukrus_lex
украинские	1693	85,17	10,99	83,70
русские	823	13,83	85,80	73,54
украинские + русские	2516	61,82	35,36	80,37

9. Обсуждение результатов

Результаты экспериментов свидетельствуют о том, что украинская АМ обеспечивает хорошие результаты как при распознавании моноязычной речи (украинской — 85,17%, русской — 85,80%), так и речи с языковым переключением (80,37%). Следует, однако, учесть, что в контрольной выборке представлена речь одного диктора, и этот диктор входит в число дикторов корпуса АКУЕМ. Кроме того, все слова контрольной выборки представлены в ЛМ. Однако, результаты экспериментов достаточно показательны, поскольку были проведены с целью сравнения эффективности различных конфигураций системы распознавания на одних и тех же моно- и двуязычных речевых сегментах.

Эксперименты по распознаванию русской речи системой, ЛМ и лексикон которой настроены на украинский язык (точность 13,83%) и украинской речи системой, ЛМ и лексикон которой настроены на русский язык (точность 10,99%), были проведены с целью оценки, какой должна быть точность идентификации языка в двухпроходной системе, когда на первом проходе идентифицируется язык, а на втором происходит распознавание с использованием моделей идентифицированного языка. Предварительные расчеты показывают, что для достижения 80,37% точности распознавания смешанной украинско-русской речи точность предварительной идентификации языка должна быть выше 95%.

Наиболее перспективным направлением дальнейших исследований является разработка кросс-язычной АМ, включающей модели как украинских, так и русских фонем. В связи с этим в дальнейшем предполагается проведение акустико-фонетического экспериментального исследования украинско-русской речи билингов с целью уточнения алфавита фонем, на основе которого должна быть создана кросс-язычная АМ.

Результаты распознавания украинско-русской речи могут быть улучшены на стадии постобработки. Ответ распознавания в виде смешанного текста можно сглаживать с помощью двуязычных словарей, в которых представлены одинаково звучащие, но имеющие разную орфографию слова обоих языков. Например, «лінія» — «линия», «ера» — «эра».

10. Заключение

В статье предложен подход к распознаванию речи с языковым переключением между близкородственными украинским и русским языками.

Использование акустической модели, разработанной ранее и обученной на украинской речи, для распознавания украинско-русской речи, на первых порах позволило избежать трудоемкой и длительной процедуры разработки специальной двуязычной акустической модели. Точность распознавания двуязычной речи при этом отличается от точности распознавания монологичной речи незначительно — на 3–7%.

Разработанные лингвистическая модель и лексикон распознавания являются двуязычными. Обучение двуязычной лингвистической модели значительно проще, чем одноязычной модели, поскольку не требует предварительного разделения множества текстов по языковому признаку. Лексикон распознавания, обеспечивающий распознавание двуязычной речи, содержит как украинские, так и русские словоформы, причем фонемные транскрипции русских словоформ представлены украинскими фонемами.

Предложенный подход не требует предварительной идентификации языка.

Литература

1. *Bhuvanagiri, K. K., Koppurapu, S. K.* An Approach to Mixed Language Automatic Speech Recognition. Proceedings of the Oriental COCODA, 2010, Nepal.
2. *Bhuvanagiri, K. K., Koppurapu, S. K.* (2012) Mixed Language Speech Recognition without Explicit Identification of Language, American Journal of Signal Processing, Vol. 2 No. 5, pp. 92–97.
3. *Chan J. Y. C., Ching, P. C., Lee, T., Cao, H.* (2009) Automatic recognition of Cantonese-English code-mixing speech, Computational Linguistics and Chinese Language Processing, vol. 14, No. 3, pp. 281–304.
4. *Derbali, M., Jarrah, M., Wahid, M. T.* (2012) A Review of Speech Recognition with Sphinx Engine in Language Detection, Journal of Theoretical and Applied Information Technology, Vol. 40, No. 2, pp. 147–155.
5. *Despres, J., Fousek, P., Gauvain, J.-L., Gay, S., Josse, Y., Lamel, L., Messaoudi, A.* The LIMSI-Vecsys Research Systems for N-Best 2008. Proceedings of the N-Best: North- and South-Dutch Benchmark Evaluation of Speech recognition Technology workshop, Soesterberg, NL, 2008.
6. *Fung, P. N., Shi, B., Wu, D., Bun, L. W., Kong, W. S.* Dealing with Multilinguality in a Spoken Language Query Translator. Proceedings of the ACL/EACL'97 Workshop on Spoken Language Translation, Madrid, Spain, 7–12 July 1997, pp. 40–43.
7. *Huang, C.-L., Wu, C.-H.* Generation of Phonetic Units for Mixed-Language Speech Recognition Based on Acoustic and Contextual Analysis. IEEE Transactions on Computers, 2007, vol. 56 no. 9, pp. 1225–1233.
8. *Imseng, D., Bourlard, H., Magimai-Doss, M.* Towards mixed language speech recognition systems. Proceedings of Interspeech, Makuhari, Japan, 2010, pp. 278–281.

9. Krivnova, O., Zinovieva, N., Zakharov, L., Strokin, G., Babkin, A. (1997) TTS Synthesis for Russian Language, Web Journal of Formal, Computational and Cognitive Linguistics, Issue 1, available at: <http://fccl.ksu.ru/ar2.htm>.
10. Lamel, L., Courcinous, S., Gauvain, J.-L., Josse, Y., Le, V.-B. Transcription of Russian conversational speech. Proceedings of the SLTU 2012 Third International Workshop on Spoken Languages Technologies for Under-resourced Languages. Cape Town, South Africa, 2012, pp. 162–167.
11. Lyu, D.-C., Lyu, R.-Y., Chiang, Y.-C., Hsu, C.-N. Speech Recognition On Code-Switching Among The Chinese Dialects. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Toulouse, France, May. 2006. Volume 1, pp. 1105–1108.
12. Lyudovyk, T. V., Pilipenko, V. V., Robeiko, V. V. Automated stenographer of Ukrainian speech (on material of acoustic corpus of Ukrainian speech) [Avtomaticheskoe raspoznavanie spontannoi ukrainskoi rechi (na materiale akusticheskogo korpusa ukrainskoi efirnoi rechi)]. Komp'uternaia Lingvistika i Intel'ektual'nye Tekhnologii: Trudy Mezhdunarodnoi Konferentsii "Dialog 2011" [Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog 2011"]. Bekasovo, 2011, pp. 478–488.
13. Lyudovyk, T., Sazhok, M. Unit Selection Speech Synthesis Using Phonetic-Prosodic Description of Speech Databases. Proceedings of the International Conference "Speech and Computer" (SPECOM'2004). St.-Petersburg, Russia, 2004. pp. 594–599.
14. Niesler, T., Willett, D. Language identification and multilingual speech recognition using discriminatively trained acoustic models. Proceedings of the ISCA Workshop on Multilingual Speech and Language Processing (MULTILING 2006), paper 004.
15. Robeiko, V., Pylypenko, V., Sazhok, M., Vasylieva, N., Radoutsky, O. Ukrainian Broadcast Speech Corpus Development. Proceedings of the 14th International Conference "Speech and Computer: SPECOM'2011". Kazan, Russia, 2011, pp. 435–440.
16. Vu, N. T., Lyu, D. C., Weiner, J., Telaar, D., Schlippe, T., Blaicher, F., Chng, E.-S., Schultz, T., Li, H. et al. A First Speech Recognition System for Mandarin-English Code-switch Conversational Speech. Proceedings of the ICASSP, Japan, 2012, pp. 4889–4892.
17. Weiner, J., Vu, N. T., Telaar, D., Metzger, F., Schultz, T. et al. Integration of language identification into a recognition system for Spoken conversations containing code-switches. Proceedings of the 3rd Workshop on Spoken Language Technologies for Under-resourced Languages, Cape Town; S. Africa, May 2012. MICA.
18. Young S. et al. The HTK Book (for HTK Version 3.4). Cambridge, UK, 2009.