

УДК 004.896+004.934

**Егорченков Алексей Владимирович**

Кандидат технических наук, доцент кафедры геоинформационных технологий, [orcid.org/0000-0003-1390-5311](https://orcid.org/0000-0003-1390-5311)  
Киевский национальный университет им. Т. Шевченка, Киев

**ПРИКЛАДНОЕ ПРИМЕНЕНИЕ РЕФЛЕКТОРНОЙ СИСТЕМЫ  
ГОЛОСОВОГО УПРАВЛЕНИЯ**

***Аннотация.** С развитием компьютерных технологий способ взаимодействия со многими электронными устройствами становится все более естественным и приближенным к реальному миру. И самым естественным способом управления является голос человека. Автором предложена уникальная как по эффективности (качественно работает со свободной речью, требует затрат на создание и настройку на предметную область во много раз меньше, чем традиционные системы), так и по универсальности (можно научить любому языку, настроить на управление любыми техническими устройствами) голосовая система управления техническими устройствами, построенная на базе теории несилового взаимодействия. Она слышит свободную речь пользователя и реагирует (так, как ее научили) на информативную составляющую этой речи. Показаны возможности применения такой системы и перспективы ее развития.*

***Ключевые слова:** теория несилового взаимодействия; система голосового управления; информационные технологии; рефлекторные системы; технические устройства*

**Постановка проблемы**

С развитием компьютерных технологий способ взаимодействия со многими электронными устройствами становится все более естественным и приближенным к реальному миру. И самым естественным способом управления является голос человека.

Для построения систем голосового управления, которая работает локально на устройстве, не требует постоянного подключения к интернету и легко адаптируется к различным языкам предлагается воспользоваться теорией несилового взаимодействия [1]. Исходя из этой теории, предложена принципиально новая модель распознавания устной речи: «устная речь» + «расчет несилового (информационного) воздействия на реакцию» + «реакция (понимание или поведение)» = РЕФЛЕКС.

**Анализ основных исследований  
и публикаций**

Несмотря на то, что первое устройство по распознаванию речи появилось в 1952 г., системы голосового управления только сейчас начинают становиться действительно востребованными. Для управления телефоном на базе операционной системы Android корпорации Google используется система Google Now. Фирма Apple в сотрудничестве с фирмой Nuance разработала систему Siri.

Вышеперечисленные системы построены на одинаковом принципе – использование «облачных» вычислительных центров и огромных объемов обучающих данных. Это требует постоянного подключения к сети Интернет. Также в данных системах отсутствует поддержка украинского языка [2; 7-11].

**Цель статьи**

Целью статьи является описание рефлекторной системы голосового управления.

**Основной материал исследования**

Теория несилового взаимодействия по-новому раскрывает и объясняет механизмы интеллектуальной деятельности человека. Исходя из гипотезы о том, что все взаимодействия в природе являются информационными (несиловыми), выдвинуто предположение о механизме взаимодействия нейронов мозга человека. И это предположение подтверждено экспериментальными исследованиями.

Отсюда появилась возможность создать технические устройства, работающие на тех же законах, что и мозг человека. И это не классические и известные в кибернетике формальные нейроны, также похожие на естественные, как бумажный кораблик похож на океанский корабль. Это более совершенные и сложные образования, реагирующие на раздражение так же, как реагируют естественные нейроны.

Отсюда технические системы, созданные на основе теории несилового взаимодействия, способны вырабатывать рефлексы на все происходящее. При этом алгоритмы (правила) выработки рефлексов те же, что и у человека [7].

Так, на основании теории несилового взаимодействия был разработан ряд рефлекторных систем голосового управления.

В общем случае система состоит из таких компонентов:

#### 1. Фонемный стенограф

Преобразование входящего оцифрованного звукового сигнала, содержащего устную речь, в набор фонем или набор слов.

#### 2. Ядро системы

Моделирование системы голосового управления техническим устройством. Содержит программную реализацию всех методов, моделей и алгоритмов системы; словарь команд; протокол работы; настройки системы.

#### 3. Обучаемая БД распознавания речи

Хранение информационной базы распознавания речи и выделения управляющего воздействия. В базе данных (БД) содержится статистика определенных входных воздействий и реакций системы.

Следует обратить внимание, что в роли фонемного стенографа могут использоваться разные «движки» (с англ. engines), что позволяет сделать систему гибче и более приспособленной к внешним условиям и разным языкам.

Ниже приведен краткий обзор таких «движков».

#### 1. Google speech Recognition API

Необходимо сделать POST запрос на сервер со звуковыми данными. Возвращается распознанный текст в виде слов. Качество распознавания удовлетворительное. Есть зависимость от контекста.

#### 2. Nuance Dragon NaturallySpeaking

Один из пионеров в области распознавания речи. Коммерческая программа с закрытым исходным кодом, и заявленной точностью распознавания 99%. Используется с несколькими языками распознавания. Предоставляется SDK для встраивания в свои системы. Есть Dragon Mobile SDK для встраивания в мобильные приложения. Процесс распознавания осуществляется на сервере [3].

#### 3. Yandex speech kit

Технологии аналогично Google. На настоящий момент поддерживаются только два языка распознавания: русский и турецкий. Заявленная точность распознавания 82% [4].

#### 4. Microsoft Speech API

Распознавание осуществляется на компьютере. Нет кроссплатформенности. На настоящий момент нет поддержки русского и украинского языков.

#### 5. CMU Sphinx

На данный момент лидер среди свободного софта по распознаванию речи. На специфической задаче уровень распознавания до 81.38%. Есть возможности по фильтрации шумов [5; 6].

#### 6. Фонетический стенограф на основе Julius speech recognition tool [12]

На выходе генерирует фонемы. Например, «п ь й! А м а» для слова «Прямо». Показывает хорошие результаты (с точки зрения стабильности повторов фонем) для русского и украинского языков, в том числе и в неконтролируемом звуковом окружении.

Поскольку в РСГУ не нужны словари, не надо анализировать слова (грамматику) и создавать модели интеллектуального анализа текста, такие системы имеют относительно традиционный ряд преимуществ: многодикторность, вариативность произношения команд (система «поймет» как «выключи звук», так и «отключи, пожалуйста, громкость»), обработка команд на устройстве в реальном времени, работа в условиях неконтролируемой акустической среды (в условиях шума), простота алгоритмов определения реакции. РСГУ проще и дешевле разработать, чем существующие на рынке аналогичные системы [7].

На сегодняшний день реализовываются такие РСГУ:

1. Система голосового управления телевизором (*prototun*). В качестве прикладного примера и для отработки алгоритмов распознавания был разработан прототип голосового управления телевизором. Включает в себя набор команд по переключению каналов, управлению громкостью, выбору канала. Может переспросить пользователя, если команда была непонятна. Система обучаемая.

Пример применения:

Диктор в свободной форме озвучивает необходимые для него действия системы. Например, «Найди-ка мне новости». Или «Отключись через 20 минут». Или «Сделай громче звук, пожалуйста». Программная платформа управления голосом передает необходимую команду на техническое устройство или озвучивает диктору информацию, затребованную в его команде. При обучении диктор сам выполняет необходимое действие и у системы вырабатывается рефлекс на подобное обращение. Если диктор говорит «по-разному», то вырабатывается устойчивый рефлекс именно на информативную часть голосовой команды.

2. Система голосового управления телефоном на базе Android. Находится в стадии разработки. Цель – управление телефоном исключительно голосом в различных условиях, в первую очередь при движении в автомобиле за рулем. Включает в себя 13 команд. Позволяет принять/отклонить вызов,

включить/выключить звук, включить/выключить громкую связь, отправить СМС с заранее определенным текстом. В данном проекте БД состоит из двух частей: определение, что пользователь произнес имя программы, и определение непосредственно команды. Точность реакции в контролируемом звуковом окружении около 95%.

Рассмотрим обобщенно алгоритм работы системы на примере СМС-сообщения. При наступлении соответствующего события – входящее СМС-сообщение, запускается программа распознавания, которая слушает, что говорит пользователь. Если программа определяет, что пользователь произнес ее имя, подключается вторая часть БД и идет непосредственное определение команды, которое учитывает контекст команды (например, при входящем СМС-сообщении более вероятно команда «Ответить на сообщение», чем «Принять звонок»). После того, как команда будет

определена, следует ее выполнение, например, отправка СМС с заранее определенным текстом. После этого система еще ждет некоторое время и затем отключается до следующего входящего события.

3. Система голосового управления терминалами. Цель – голосовое управление терминалами в условиях неконтролируемого звукового окружения. Причем звуковое окружение будет меняться в зависимости от места расположения терминала. В данном проекте планируется оснащение терминала датчиком физического присутствия человека и система подавления шумов. В отличие от проекта системы голосового управления телефоном, в данном проекте необходимо надежное распознавание числительных. Обобщенно взаимодействие пользователя с заправочным терминалом показано на рисунке.



Рисунок – Схема взаимодействия пользователя и терминала

### Выводы и перспективы для дальнейших исследований

Теория несилового взаимодействия является удобным инструментом для построения рефлекторных систем, в частности, систем голосового управления, что было показано в результате практической разработки таких систем.

Это уникальная как по эффективности (качественно работает со свободной речью, требует затрат на создание и настройку на предметную область во много раз меньше, чем традиционные

системы), так и по универсальности (можно научить любому языку, настроить на управление любыми техническими устройствами) система. Она слышит свободную речь пользователя и реагирует (так, как ее научили) на информативную составляющую этой речи [7].

Дальнейшие исследования по разработке рефлекторных систем голосового управления открывают новые горизонты для ученых и предпринимателей в развитии и управлении техническими устройствами.

## Список літератури

1. Тесля Ю.М. Введення в інформатику природи [Текст] / Ю.М. Тесля // Монографія. – К.: Маклаут, 2010. – 255 с.
2. Сайт «Інтроформатика» [Електронний ресурс]. – Режим доступу: <http://introformatika.org.ua/>.
3. Сайт «NUANCE» [Електронний ресурс]. – Режим доступу: <http://www.nuance.com/for-developers/dragon/client-sdk/index.htm>
4. Сайт «Каталог технологій Yandex» [Електронний ресурс]. – Режим доступу: <https://tech.yandex.ru/speechkit/cloud/>
5. Сайт «CMU Sphinx» [Електронний ресурс]. – Режим доступу: <http://cmusphinx.sourceforge.net/>
6. Сайт статті «Шумопониження в CMU Sphinx» [Електронний ресурс]. – Режим доступу: <http://habrahabr.ru/post/227099/>
7. Тесля Ю.М. Рефлекторная система голосового управления техническими устройствами (PCГУ) [Текст] / Ю.М. Тесля, О.Ю. Чорний // Управління розвитком складних систем. – К. – 2013. – №15. – С. 105-110.
8. Тесля Ю.М. Теорія несилової взаємодії та її застосування до побудови систем голосового управління технічними пристроями [Текст] / Міжнародна науково-технічна конференція «Комп'ютерні науки: освіта, наука, практика», 14-16 червня 2012р., Миколаїв, НУК.
9. Тесля Ю.М. The Non-Force Interaction Theory for Reflex System Creation with Application to TV Voice Control [Текст] / Ю.М. Тесля, В.В. Пилипенко, Н.Л. Попович, О.Ю. Чорний // VI Міжнародна конференція: «International Conference on Agents and Artificial Intelligence (ICAART 2014)», 6-8 мая 2014 г., м. Анже, Франція
10. Тесля Ю.М. Приложение теории несилового взаимодействия к построению системы голосового управления [Текст] / Ю.М. Тесля, О.Ю. Чорний // XI Міжнародна науково-технічна конференція «ABIA-2013», 21-23 мая 2013 г.
11. Тесля Ю.М. Система голосового управления на базе теории несилового взаимодействия [Текст] / Ю.М. Тесля, А.В. Єгорченков, Н.Ю. Єгорченкова // III міжнародна науково-практична конференція «Обчислювальний інтелект (результати, проблеми, перспективи)», Київ-Черкаси, 2015 р. – С. 261-262.
12. Пилипенко В.В. Распознавание ключевых слов в потоке речи при помощи фонетического стенографа [Текст] / В.В. Пилипенко // Речевые технологии. – 2009. – № 1. – С. 75-79.

Статья поступила в редколлегию 08.02.2016

Рецензент: д-р техн. наук, проф. А.А. Белошицкий, Киевский национальный университет им. Т. Шевченко, Киев.

#### Єгорченков Олександр Володимирович

Кандидат технічних наук, доцент кафедри геоінформаційних технологій, [orcid.org/0000-0003-1390-5311](http://orcid.org/0000-0003-1390-5311)  
Київський національний університет ім. Т. Шевченка, Київ

#### ПРИКЛАДНЕ ЗАСТОСУВАННЯ РЕФЛЕКТОРНОЇ СИСТЕМИ ГОЛОСОВОГО УПРАВЛІННЯ

**Анотація.** З розвитком комп'ютерних технологій спосіб взаємодії з багатьма електронними пристроями стає все більш природним і наближеним до реального світу. І самим природним способом управління є голос людини. Автором запропонована унікальна як за ефективністю (якісно працює з вільною промовою, вимагає витрат на створення і налаштування на предметну область у багато разів менше, ніж традиційні системи), так і за універсальністю (можна навчити будь-якій мові, налаштувати на управління будь-якими технічними пристроями) голосова система управління технічними пристроями, побудована на базі теорії несилової взаємодії. Вона чує вільну мову користувача і реагує (так, як її навчили) на інформативну складову цієї мови. Показано можливості застосування такої системи та перспективи її розвитку.

**Ключові слова:** теорія несилової взаємодії; система голосового управління; інформаційні технології; рефлекторні системи; технічні пристрої

#### Igorchenkov Oleksii

PhD, associate professor of the department of geoinformatics, [orcid.org/0000-0003-1390-5311](http://orcid.org/0000-0003-1390-5311)  
Taras Shevchenko National University of Kyiv, Kyiv

#### APPLIED APPLICATION OF REFLEX SYSTEM OF VOCAL MANAGEMENT

**Abstract.** With development of computer technologies the method of co-operating with many electronic devices becomes more natural, close to the real world. And by the most natural method of management, there is voice of man. An author is offer unique both on efficiency (qualitatively works with free speech, requires expenses on creation and tuning on a subject domain in oftentimes less than, than traditional systems) and on universality (it is possible to teach to any language, to influence on a management any technical devices) vocal control system by technical devices, built on the base of theory of non-violent interaction. She hears free speech of user, and reacts (because she was taught) on the informing constituent of this speech.

Possibilities of application of such system and prospect of her development are shown. An example of application of the reflex system of vocal management is made technical devices: system of vocal management by television, system of vocal management by a telephone on a base Android and system of vocal management by terminals. The results of experiment showed the high level of efficiency and competitiveness of the investigational system on the row of indexes : does not require limitations in a language; uses an universal algorithm that can be used for creation of many devices of recognition, and not only speeches; studies easily; does not react on noise; fishes for useful information in the stream of free speech; allows naturally to add new commands; allows to reprogram behavior of the system voice; studies on own errors; executes fixing, analysis and correction of errors of recognition; executes vocal authentication of speaker; tuned under a manner and habits of speaker.

**Keywords:** *theory of non-violent interaction system of vocal management; information technologies; reflex systems; technical devices*

#### References

1. Teslya, Y. M. (2010). *Introduction to informatics of nature*. Kyiv, Ukraine: Maklout.
2. Sait «Introformatika» [Site “Introformatika”]. [introformatika.org.ua](http://introformatika.org.ua). Retrieved from <http://introformatika.org.ua/> [in Russian].
3. Sait «NUANCE» [Site “NUANCE”]. [nuance.com](http://www.nuance.com). Retrieved from <http://www.nuance.com/for-developers/dragon/client-sdk/index.htm>. [in English].
4. Sait «Katalog technology Yadex» [Site “Product technology Yandex”]. [tech.yandex.ru](http://tech.yandex.ru). Retrieved from <https://tech.yandex.ru/speechkit/cloud/> [in Russian].
5. Sait «CMU Sphinx» [Site “CMU Sphinx”]. [cmusphinx.sourceforge.net](http://cmusphinx.sourceforge.net). Retrieved from <http://cmusphinx.sourceforge.net/> [in English].
6. Sait statti «Shumoponyzhenye v CMU Sphinx» [Site of an article “Noise in CMU Sphinx”]. [habrahabr.ru](http://habrahabr.ru). Retrieved from <http://habrahabr.ru/post/227099/> [in Russian].
7. Teslia, Y., & Chorny, O. (2013). *Reflex voice control technical devices (RVCTD). Management of development of complex systems*. Kyiv, Ukraine: 15, 105 – 110.
8. Teslia, Y.M. (2012). *The theory of non-forcible interaction and its application to construction of technical devices Voice Control. Proceedings from KNTP'12: The International Scientific and Technical Conference “Computer science: education, science, practice”*. Mykolayiv: NUK [in Ukrainian].
9. Teslia, Y.M., Pilipenko, V.V., Popovich, N.L., & Chorny, O.Y. (2014). *The Non-Force Interaction Theory for Reflex System Creation with Application to TV Voice Control. Proceedings from ICAART'14 The sixth international conference: «International Conference on Agents and Artificial Intelligence (ICAART 2014)»*. Angers, France.
10. Teslia, Y.M., & Chorny, O.Y. (2013). *Prilozhenie teorii nesilovogo vzaimodejstviya k postroeniju sistemy golosovogo upravlenija [Application of the theory of non-force interaction to build a voice control system]. Proceedings from ABIA'13: XI Mizhnarodna naukovno-tehnichna konferentsiia «ABIA'13» - The Ninth International Scientific and Technical Conference “ABIA'13”*.
11. Teslia, Y.M., Iegorchenkov, O.V., & Yehorchenkova, N.Y. (2015). *Voice control system based on the theory of non-force interaction. Proceedings from OIRPP'15: The Third International Scientific and Practical Conference “Computational intelligence (results, problems and perspectives)”*. (pp.261-262). Kyiv – Cherkassy.
12. Pylypenko, V. (2009). *Recognition of keywords in the flow of speech using phoneme-by-phoneme recognizer, Rechevyte tekhnologii*, 1, 75-79 [in Russian].

#### Ссылка на публикацию

- APA Iegorchenkov, A.V. (2016). *Applied application of reflex system of vocal management (2016). Management of Development of Complex Systems*, 25, 103 – 107.
- ГОСТ Егорченков А.В. *Прикладное применение рефлекторной системы голосового управления [Текст] / А.В. Егорченков // Управление развитием сложных систем. – 2016. – № 25. – С. 103 – 107.*