

**Крылов В.С.**

## **Компьютерная паралингвистика как основа приобретения специальных профессиональных навыков в области информационных технологий**

*Крылов Владимир Сергеевич, кандидат биологических наук, доцент  
кафедра информационно-компьютерных технологий*

*Республиканское высшее учебное заведение Крымский инженерно-педагогический университет  
г. Симферополь, Крым*

**Аннотация.** Информационные технологии (ИТ) трансформировали общество в информационное. Сформировалась глобальная инфраструктура, обеспечивающая передовые услуги за счет организации связи между людьми, и даже между вещами физическими или виртуальными. Широкий круг исследователей получил возможность использовать ранее недоступные инструменты и технологии для экспериментальных и теоретических исследований. ИТ, методы решения проблем организации и функционирования глобальной информационной структуры – являются источником моделей для изучения взаимодействия сложившихся в процессе эволюции вербальных (речевых) и невербальных каналов коммуникации. На основе этих технологий возникают новые интегрированные направления исследований. Например, компьютерная паралингвистика предлагает модели, которые позволяют значительно усовершенствовать системы анализа и обработки естественной речи и текстов, усовершенствовать поисковые системы. Включение в учебный процесс достижений компьютерной лингвистики, позволяет будущему специалисту приобрести навыки успешной конкуренции на рынке труда.

**Ключевые слова:** компьютерная паралингвистика, информационные технологии, учебный процесс

Первоначально информационное пространство развивалось как совершенствование сети интернет. Обмен текстами повлек за собой появление новых языковых особенностей, лингвистические характеристики которых определяются соответствующим каналом коммуникации в информационном пространстве. Такими, как электронная почта, мгновенный обмен сообщениями в чатах, СМС сообщения, твиттер и так далее. Просто текстовых сообщений для передачи, например, эмоционального состояния коммуниканта оказалось недостаточно. В результате, спонтанно возникли смайлики. Текст сообщения стал дополняться кодами, не имеющими прямого отношения к алфавиту и грамматике сообщения. Такие сообщения представляют собой единое целое для передачи смысла и уже не являются просто суммой, объединением разнородных элементов. То есть, сообщения объединяются в единую структуру представленную компонентами: вербальной (речевой) и невербальной (не речевой). Последняя относится к другим знаковым системам.

На основе возможности использования ранее недоступных инструментов и технологий возникли новые интегрированные направления исследований. Компьютерная паралингвистика определяется как междисциплинарное направление, задача которого обобщать и развивать методы исследования на основе интегрированных моделей взаимодействия вербальных (речевых) и невербальных каналов коммуникации. Результатом должны стать решения прикладных задач для совершенствования систем взаимодействия пользователей как с отдельными техническими устройствами, так и с информационной реальностью в целом [13].

**Цель настоящей** статьи предложить широко используемые информационные технологии, методы функционально-ориентированного построения сложных программных систем, опыт разработки платформ программирования, как основу моделей компьютерной паралингвистики для решения прикладных задач. Например, визуальное программирование является реальным примером модели, в которой объединяются вербальные (язык программирования) и невербальные (манипулирование графическими объектами) методы

программирования. Включение в учебный процесс достижений компьютерной лингвистики, позволит будущему специалисту приобрести навыки успешной конкуренции на рынке труда.

Язык, как система фонетических, лексических и грамматических средств, порождает естественную информационную реальность, через которую обеспечиваются коммуникации в обществе. Внедрения информационных технологий (ИТ) в самые различные сферы деятельности людей, стремительное развитие сети интернет первоначально рассматривалось как дополнение или расширение естественного информационного пространства. Совершенствование технических и программных средств из дополнения естественного информационного пространства превратилось в самостоятельную информационную реальность, а общество незаметно для себя стало информационным. В освоении возникшей информационной реальности задействованы интуитивные механизмы выработки адаптивного поведения, которые сложились в результате эволюции и из индивидуального опыта. Эти интуитивные механизмы обеспечивают процесс освоения ребенком языка его социального окружения, процесс формирования пиджин и креольских языков [10], формирование невербального адаптивного поведения, как в материальной, так и в виртуальной реальностях [5]. Они имеют не только внешнее сходство между собой, но и вполне укладываются в общую стратегию адаптации к новому окружению, новому пространству, как социальному, так и информационному и материальному. Каждый из этих процессов зависит друг от друга. Вербальную часть в реальной речи невозможно отделить от невербальной, не утратив смысла. Таким образом, для целостного понимания механизмов адаптации, в том числе нужны и невербальные средства в языковой коммуникации, которые изучаются в специальном разделе – паралингвистика [13].

Для объяснения способности освоения языка Н. Хомским была предложена гипотеза о существовании универсальной грамматики (УГ). УГ определялась как общий набор правил, который даёт зна-

чально при рождении как знание о языке на генетическом уровне. С целью формализации описания УГ было развито представление о порождающей грамматике, свойства которой производны от нее. В рамках порождающей грамматики формулируется система правил, при помощи которых можно определять комбинации слов, составляющих грамматически правильные предложения [6].

Размытость в определении понятия универсальной грамматики, гипотеза о её существовании породили невероятное количество лингвистических теорий. Доминировавшие в прошлом веке достаточно подробно представлены в [3]. В то же время ни одной из теорий, опираясь на представление об УГ, так и не удалось на достаточно приемлемом уровне составить представление о том, как ребенок овладевает конкретным языком.

В работе [8] была поставлена цель определить экспериментально какая модель грамматики лучше всего представляет речь маленьких детей: основанная на абстрактных правилах или основанная на конкретных словах и фразах и некоторых локальных абстракциях ассоциированных с этими словами и фразами.

По записям речи детей составлялся корпус для каждого ребенка. С помощью специальных статистических методов компьютер выводил грамматику, порождающую высказывания в корпусе. Специальная программа выводила два типа грамматики: полностью абстрактная, и «основанная на использовании» или лексико-специфическая. После вывода обе грамматики проверялись на тестовом корпусе. Как оказалось, лексико-специфическая грамматика более адекватно соответствовала данным тестового корпуса.

Следует отметить, что в этом исследовании регистрировалась только вербальная часть общего поведения ребенка. Кроме того, лексико-специфическая грамматика, так как она определена исследователями в данной работе, по классификации Н. Хомского относится к контекстно-зависимой грамматике. То есть, относится к грамматике с менее жесткими правилами, определяющими синтаксис языка, чем контекстно-свободная грамматика. Это означает, что аппроксимация синтаксиса реальной речи ребенка более точно отражается формальным языком с контекстно-зависимой грамматикой [1, 2].

Развитие интернета повлекло за собой появление новых языковых особенностей, лингвистические характеристики которых определяются соответствующим каналом коммуникации в информационном пространстве. Синтаксис, грамматика таких языков значительно отличается от синтаксиса, грамматики базового языка коммуникантов. Они определяются как контактные языки подобные пиджин и креольским языкам [9, 10].

В каналах мгновенных текстовых сообщений интернет чатов, твиттере, форумов, СМС и других идет активное словообразование через образования новых аббревиатур. В тексты обменов активно включаются смайлики, визуальные и аудио элементы. Текст сообщения становится поликодовым, в котором все элементы не являются просто суммой, объединением, а интегрированы единое целое для передачи смысла. Такие тексты определяются как креализованные [9].

То есть тексты, состоящие из двух объединённых в единую структуру частей: вербальной (речевой) и невербальной, которая относится к другим знаковым системам. Обе эти части зависят друг от друга. Вербальную часть невозможно отделить от невербальной. Невербальные средства в языковой коммуникации изучаются в специальном разделе паралингвистика [9, 11].

С невербальным поведением связывают такие каналы коммуникации людей с их окружением, как интонация, визуальные сигналы, тактильную чувствительность или касание (прикосновение), использование пространства, использование времени, физический внешний вид тела, использование предметов материальной культуры, а так же ольфакторные или обонятельные сигналы [11].

Невербальное поведение представляет собой сложное явление, которое характеризуется зависимостью от окружающего пространства, контекста, важностью особенностей движений, автоматизмом. То есть продуцирование и восприятие происходят в основном вне осознания. Оно является предметом исследования широкого спектра дисциплин, включая биологию, антропологию, социологию, коммуникации, а также социальную и экспериментальную психологию [5, 11].

Существует общая стратегия адаптации к незнакомой окружающей среде. Первоначально при попадании в незнакомую среду, в том числе и социальную, независимо от ее сложности, в течение достаточно короткого времени определялся конечный объем паттернов или алфавит поведения. Затем эти паттерны становятся основой для сложного поведения в заданной среде. После отбора алфавита паттернов поведения за счет процесса перекомбинаций происходит формирование всех возможных способов поведения в среде. Поведение формируется как построение предложения из «слов» составленных из алфавита паттернов поведения. Далее определяются правила формирования адаптивного поведения в среде. Эти правила объединяются в целостную грамматику правильно построенных «предложений», которые соответствуют адаптивному поведению. В дальнейшем адаптивное поведение в среде или социальном окружении строится как правильно построенное, в соответствии с «грамматикой», предложение. Дисфункциональные формы поведения являются не адаптивными или ошибочным с точки зрения определенной «грамматики» «предложениями поведения».[5]

Процесс освоения ребенком языка его социального окружения, процесс формирования пиджин и креольских языков, формирование невербального адаптивного поведения как в материальной, так и в виртуальной реальностях имеют не только внешнее сходство, но и вполне укладываются в общую стратегию адаптации к новому окружению, новому пространству, как социальному, так и информационному. То есть модель этого процесса, как системы программ, должна предоставлять возможность в едином ключе, на единой основе обеспечивать интегрированную реализацию целостного поведения, моделирование паралингвистических факторов в речевой активности. Этому подходу полностью соответствует современное пред-

ставление о платформах программирования. Платформы – это программные системы, которые позволяют реализовать на общей основе разные языки программирования, технологии визуального программирования, интегрировать в единое целое разрабатываемые сложные программные системы. Например, такие платформы как .NET фирмы Microsoft или широко распространенная платформа разработки программного обеспечения приложений для мобильных телефонов Android и другие.

Такой подход в определении интегрированных моделей компьютерной паралингвистики на основании особенностей выбранной платформы программирования, имеет не только научное или прикладное значение для развития и совершенствования существующих информационных систем. Он позволяет формулировать различной сложности задания в учебном процессе подготовки специалистов в области информационных технологий. Выполнение заданий стимулирует развитие навыков и умений приобретать дополнительные знания, осваивать как можно больше востребованных информационных технологий, умение разрабатывать программные продукты для разных отраслей, сфер и направлений практического их применения.

Одно из заданий формулировалось следующим образом. Оно заключалось в разработке анализатора

упрощенного креализованного языка, для которого задан некоторый ограниченный набор слов и смайликов. Такие языки распространены в рекламной деятельности. Смайки в данном случае имитировали картинки или изображения, сопровождающие рекламные тексты. Результат анализа заключался в определении условного эмоционального состояния того, кто генерировал фразы языка. Цель задания состояла в приобретении навыков работы с выбранной платформой программирования, умения на практике применить теоретические знания из теории формальных языков и методов разработки трансляторов, наработка практических навыков приобретения в короткий срок и на достаточном уровне знаний в различных отраслях деятельности для разработки прикладных программ.

Включение в учебный процесс достижений прикладных исследований новых направлений, таких как компьютерная лингвистика, позволяет будущему специалисту приобрести навыки эффективного решения поставленных заказчиком задач. Именно на таком материале, на основе приобретенного опыта, молодой специалист получает необходимый набор профессиональных навыков успешной конкуренции на ИТ-рынке труда [4].

#### ЛИТЕРАТУРА (REFERENCES TRANSLATED AND TRANSLITERATED)

1. Ахо А.В. Компиляторы: принципы, технологии и инструментарий // Ахо А.В., Лам М. С, Сети Р., Ульман Д.Д. - М. : 000 "И.Д. Вильямс", 2008. - 1184 с.  
*Aho A.V. Compilers: Principles, Techniques and Tools // Aho A.V., Lam M.S., Network, R., D. Ullman - M: 000 "I. Williams," 2008. - 1184 p.*
2. Бабин Д.Н. Об автоматной аппроксимации естественных языков. // Бабин Д.Н., Холоденко А.Б. Интеллектуальные системы, 2008, Том 12, выпуск 1-4, с. 125 -136.  
*Babin D.N. An automaton approximation of natural languages. // Babin D.N., Cholodenko A.B. - Intelligent Systems, 2008, Volume 12, Issue 1-4, p. 125 -136.*
3. Демьянков В.З. Доминирующие лингвистические теории в конце XX века // Демьянков В.З. - Язык и наука конца XX века. - М., 1999. - С.239-320.  
*Demyankov V.Z. Dominant linguistic theory in the late XX century // Demyankov V.Z. - Language and science of the end of XX century. - M., 1999. - P.239-320.*
4. Крылов В.С. Молодые специалисты способны конкурировать на рынке труда информационных технологий. // Крылов В.С. - Science and Education a New Dimension: Pedagogy and Psychology. Vol. 5, 2013 – p. 98 – 101.  
*Krylov V.S. Young professionals are able to compete in the labor market of information technologies. // Krylov V.S. - Science and Education a New Dimension: Pedagogy and Psychology. Vol. 5, 2013 - p. 98 – 101.*
5. Крылов В.С. Суеверие в мире "интернета вещей". Междисциплинарный подход // Крылов В.С. - Science and Education a New Dimension. Pedagogy and Psychology, I(7), Issue: 14, 2013., 250 – 253.  
*Krylov V.S. Superstition in the world "Internet of Things". Interdisciplinary approach // Krylov V.S. - Science and Education a New Dimension. Pedagogy and Psychology, I (7), Issue: 14, 2013., 250 – 253.*
6. Слобин Д., Грин Д. Психоллингвистика // Слобин Д., Грин Д. - М.: КомКнига, 2006. — 352 с.  
*Slobin D., Green J. Psycholinguistics // Slobin D., Green J. - M.: KomKniga 2006. - 352.*
7. Черняк Л. Интернет вещей: новые вызовы и новые технологии // Черняк Леонид - Открытые системы, № 04, 2013 - [Электронный ресурс] Режим доступа: <http://www.osp.ru/os/2013/04/13035551/>  
*Chernyak L. Internet of things: new challenges and new technologies [E-source] // Leonid Chernyak - Open systems, № 04, 2013 - Access mode: http://www.osp.ru/os/2013/04/13035551/*
8. Bannarda C. Modeling children's early grammatical knowledge/ Bannarda C., Lieven E., Tomasello M. - PNAS, 2009 vol. 106, no. 41, pp. 17284–17289.
9. Beisenova Z.S. The Study of Creolized Texts in Written Communication/ Beisenova Z.S., Bayekeyeva A.T., Sapina S.M., Dinayeva B.B., Utanova A.K. - INDIAN JOURNAL OF APPLIED RESEARCH, Volume : 4, Issue : 5, May 2014 p. 217.
10. Bickerton D. Creole Languages // Bickerton D. - Scientific American, July 1983, Vol. 249, No.1, pp. 116—122.
11. Krämer N.C. Nonverbal Communication [E-source] // Nicole C. Krämer - Human Behavior in Military Contexts, P. 150 - 188, Acces mode: <http://books.nap.edu/catalog/12023.html>
12. Lawson L. The Secret to Success with the Internet of Things [E-source]/ Lawson Loraine. – Acces mode: <http://www.itbusinessedge.com/blogs/integration/the-secret-to-success-with-the-internet-of-things.html>
13. Schuller B., Batliner A. Computational paralinguistics : emotion, affect and personality in speech and language processing / Schuller Bjoörn, Batliner Anton. – First Edition, John Wiley & Sons, Ltd , 2014, - p. 321.

**Krylov V.S. Computer paralinguistics as a basis the acquisition of special skills in the field of information technology**

**Abstract.** Information technology (IT) was transformed into an information society. Formed a global infrastructure that provides advanced services due to communication between people and even between things physical or virtual. A wide range of researchers had the opportunity to use previously inaccessible tools and technologies for experimental and theoretical studies. IT methods for solving the problems of the organization and functioning of the global information structure - are a source of models for studying the interaction of prevailing in the evolution of verbal (speech) and nonverbal communication channels. Based on these technologies there are new integrated research directions. For example, computer paralinguistics offers models, which can significantly improve the system analysis and processing of natural speech and texts, to improve search engines. Inclusion in the learning process advances of computational linguistics, allows future specialist to acquire the skills to compete successfully in the labor market.

**Keywords:** *computer paralinguistics, information technology, the learning process*