

РАЗРАБОТКА КОМПЬЮТЕРНОЙ СИСТЕМЫ ТЕСТИРОВАНИЯ КОГНИТИВНЫХ СПОСОБНОСТЕЙ РЕСПОНДЕНТОВ НА ОСНОВЕ ПРЕДЛОЖЕНИЙ С ЭЛЛИПСИСАМИ

К. А. Найденова¹ [0000-0003-2377-7093], Е. С. Булыкина² [0009-0009-6018-8360],
В. А. Пархоменко³ [0000-0001-7757-377X], А. В. Щукин⁴ [0000-0002-9534-824X],
Т. А. Мартирова⁵ [0000-0003-0000-6608]

^{1,5} Военно-медицинская академия, Санкт Петербург,

^{2,3,4} Санкт-Петербургский политехнический университет Петра Великого,
Санкт Петербург;

¹ ksennaidd@gmail.com, ² bulykina06@gmail.com,

³ vladimir.parkhomenko@spbstu.ru, ⁴ alexander.schukin@spbstu.ru,

⁵ martta462@yandex.ru

Аннотация

Когнитивные способности охватывают все высшие интеллектуальные функции человека, в том числе устную и письменную речь. Использование эллипсисов связано с умением четко формулировать мысль, избегая избыточности её выражения. Тестирование умения использовать эллипсисы может быть основой, с одной стороны, для оценки самой когнитивной операции порождения эллипсисов в письменной речи и обнаружения нарушений этой операции. Это может послужить в медицине для раннего обнаружения деменции. С другой стороны, такое тестирование применимо для обучения речевым технологиям, в том числе на иностранных языках. С научной точки зрения, важно определить степень влияния обучения на выполнение заданий теста.

В статье описаны разработка компьютерной системы дистанционного online тестирования выполняемой человеком операции преобразования полного предложения в предложение с эллипсисом определенного типа, принципы оценки результатов тестирования испытуемых. Особенностью системы является создание нового типа заданий, связанных с обработкой естественного языка. Сложность разработки определяется необходимостью автоматизированного

формирования банка пар предложений на естественном языке без эллипсисов и с эллипсисами, что влечет включение в систему средств синтаксического анализа текстов и алгоритмов автоматизированного формирования эллипсисов в предложениях и/или восстановления полного предложения из эллиптического. В заключении обсуждены перспективы развития и применения предлагаемой системы.

Ключевые слова: эллипсис, обработка естественного языка, on-line тестирование, когнитивные способности

ВВЕДЕНИЕ

Настоящая работа основана на рассмотрении когнитивных механизмов порождения эллипсисов и формулировке правил генерации некоторого типа эллиптических предложений в русском языке. Некоторые из этих правил сформулированы на основе анализа текстов планиметрических задач и частично из предложений художественной литературы (300 предложений) [1].

Эллиптическими называются конструкции, которые содержат опущенный, но однозначно восстанавливаемый элемент в предложении (слово или сочетание слов) [2].

Явление эллипсиса связывают с желанием избежать избыточности в речи или письменном тексте. Однако понятие избыточности трудно формализуемо. Автор работы [3] связывает исследование избыточности с изучением функциональной нагрузки различных элементов и структур текста. В работе проведено оригинальное исследование, когда из текста удалялись различные элементы и затем наблюдались реакции носителей языка при восстановлении пропусков.

Общепринятой типологии эллиптических конструкций на данный момент не существует [4]. В конкретных примерах между эллипсисом и сходными явлениями неполноты, имплицитности, зевгмы может не быть четкой границы, подчеркивают Л.Ю. Максимов и В.В. Бабайцева [5]. В результате в одном и том же предложении эллипсис относят к разным типам. Так, эллипсис в предложении «Один ведерком черпает, другой – шапкой, третий – горстями» относят к зевгме [3], к глагольному эллипсису [4], анафорическому или сочинительному эллипсису [2] и к грамматическому эллипсису [6].

Основание эллипсиса мы связываем с мыслительными процессами, реализуемыми при образовании предложений. Важным является изучение связи

между когнитивными операциями, то есть процессами манипулирования знаниями, и синтаксисом различных языков. Начало этому направлению работ было положено Р. Джекендоффом [7–9].

С нашей точки зрения, классификация эллипсисов может быть произведена на основе знаний, которые участвуют в процессе их образования:

1. однозначно восстанавливаемый эллипсис в силу того, что опускаемая часть уже есть в тексте в том же или ближайшем предложении;
2. однозначное восстановление на основе общих знаний, традиций, принятых соглашений, подразумеваемого контекста;
3. многозначное восстановление – по контексту может быть несколько возможностей восстановления;
4. невозможность восстановления.

Приведем некоторые примеры из дневников драматурга Е. Шварца [10].

Восстановление однозначное: *Иногда Милочка здоровалась со мной приветливо, иной раз невнимательно, как бы думая о другом, то дружески, а вдруг – как с малознакомым.*

Я словно заново научился ходить и смотреть, а главное – говорить.

Восстановление однозначное на основе традиционных знаний:

Был вчера в Художественном. Я живу на Суворовском.

Восстановление многозначное с синтаксически разными фрагментами, но семантически эквивалентными и зависимыми от контекста:

В домах живут. (Люди – общезначимо, у Шварца, семьи военных, театральные деятели).

Невозможно восстановить – из-за противоречия логического, незнания контекста, когнитивных ошибок.

Такой пример дает Умберто Эко в [11]: *Charles makes love with his wife twice a week, so does John* (Чарльз занимается любовью со своей женой дважды в неделю, Джон – тоже).

Далее мы будем рассматривать только два типа эллипсисов, которые однозначно разрешимы на основе анализа предложения, содержащего эллипсис, и, возможно, соседних предложений. Мы будем называть их «глагольный эллипсис» (the VPE) и «именной эллипсис», структура которых основана на именных, глагольных и предложных группах (Noun Phrases, Verb Phrases, Prepositional

Phrases). В литературе эти типы эллипсисов известны: первый – как эллипсис глагольной группы (исключение глагола, как изолированного, так и в составе глагольных групп и целых клауз), второй – как эллипсис составляющих именной группы с «сохранением представителя» [4].

Когнитивный подход предполагает, что:

1. полное предложение описывает некоторую когнитивную ситуацию, например, геометрическую конфигурацию, выраженную в тексте задачи по планиметрии;
2. полное предложение мысленно трансформируется в неполное эллиптическое предложение; причем в основе трансформации лежат когнитивные операции, выполняемые по определенным правилам.

Напомним некоторые правила генерации предложений с глагольным и именным эллипсисами [1].

Правило 1. Если имеется ввиду одно и то же действие над несколькими объектами, то после описания этого действия над первым объектом в предложении далее это действие над другими объектами может быть описано без копирования наименования этого действия (пропускается глагол).

Правило 2. Глагол может входить в предложение в составе глагольной группы (ГГр), тогда после полного описания действия над первым объектом возможно в последующих описаниях того же действия над другими объектами опускать не только глагол, но и повторяющиеся составляющие глагольных актантов.

Правило 3. Объект может быть выражен в составе именной группы (ИГ), тогда при упоминании в предложении того же самого объекта возможен пропуск не только его наименования, но и общих повторяющихся составляющих ИГ, характеризующих объект.

Рассмотрим некоторые примеры:

Дана трапеция ABCD с основанием AD. Биссектрисы внешних углов при вершинах A и B пересекаются в точке P, а при вершинах C и D – в точке Q.

Во втором предложении, часть ИГ «биссектрисы внешних углов» и глагол «пересекаются» опущены.

В результате первой операции он должен был добиться уничтожения врага к западу от реки и выйти к ней, в результате второй – создать

плацдармы, а затем провести вторжение в глубь Германии (У. Черчилль. Вторая мировая война. Том 6. Часть 2).

В этом предложении пропущена составляющая сказуемого «должен был».

Но когнитивные правила выполнимы только с соблюдением некоторых синтаксических правил, назовем их «синтаксическими аспектами эллипсисов»:

1. каждое слово (словосочетание) может иметь подчиняющее (стержневое – pivotal, main, governing) слово и притом только одно (обратное неверно; без этого правила эллипсис был бы в принципе невозможен);
2. подчиняющее слово (например, вершина ИГ или ГГр) или словосочетание могут иметь несколько аналогичных по грамматической форме и по значению управляемых слов (subordinate) или словосочетаний;
3. слова, присутствующие в эллиптических фрагментах предложений, сохраняют грамматические функции и формы, свойственные им в соответствующих полных предложениях [12];
4. реконструкция глагольных эллипсисов не зависит от того, сколько раз пропущен глагол, и от того, сколько у глагола аналогичных подчиненных словосочетаний;
5. пропущенные слова или составляющие уже были употреблены в ближайшем или в том же самом предложении [12];
6. в эллиптической части предложения опускаются управляющие (pivotal) слова (вершины ИГ или ГГр) отдельно или с некоторыми составляющими, но зависимые от них слова остаются [12].

Модели управления для каждой лексемы и способы их морфосинтаксического оформления даны в [13].

ТИПОЛОГИЯ ЭЛЛИПТИЧЕСКИХ ПРЕДЛОЖЕНИЙ

В результате анализа текстов были выделены три типа предложений с эллипсисами.

Тип 1 охватывает случаи, когда два (и более) простых предложения с сочинительной связью входят в сложноподчиненное предложение:

В равнобедренный прямоугольный треугольник ABC с прямым углом при вершине B вписан прямоугольник MNKB так, что две его стороны MB и KB лежат на катетах, а вершина N – на гипотенузе AC.

В каждом из придаточных сочинительных предложений имеется свое подлежащее. Первое придаточное предложение без эллипсиса ограничено слева союзом (союзным словом): «если»; «так, что», «так, чтобы» и т. п. Справа в качестве ограничения употребляются запятая и сочинительный союз «а». Часто ограничителями могут быть запятые и союз «и», в особенности при множественном эллипсисе. Несколько простых предложений с сочинительной связью могут представлять самостоятельное сложносочиненное предложение. Тире, обозначающее пропуск глагола, также может отсутствовать.

Структуру предложений Типа 1 можно представить следующим образом (ИГподл, ГГр – именная группа подлежащего и глагольная группа, соответственно):

ИГ1подл ГГр1- полная, [ИГ_і подл ГГр_і -неполная, $i = \{2, \dots, n\}$].

В этой формуле ГГр включает только аргументы глагола и не включает ИГподл.

Тип 2 охватывает предложения с одним подлежащим, где сочинительной связью соединены ГГр-пы с одним и тем же глаголом, но различными актантами (прямыми и косвенными дополнениями):

Окружность касается сторон AB и AD прямоугольника ABCD и пересекает сторону DC в единственной точке F, а сторону BC – в единственной точке E.

Предложения Типа 2 имеют следующую структуру:

ИГ1подл ГГр1- полная, [ГГр_і -неполная, $i = \{2, \dots, n\}$].

Глаголы в предложениях Типов 1 и 2, как правило, употребляются в изъявительном наклонении, но встречаются оба типа предложений с глаголом в страдательном залоге.

Мы объединили в Тип 3 все предложения, отличные от предложений Типов 1 и 2, включающие неопределенно-личные, побудительные, побудительно-инфинитивные предложения и предложения, где эллипсис появляется в деепричастном и причастном оборотах:

Вершину C треугольника ABC примите за начало координат, а ось проекций – за ось абсцисс.

В треугольник ABC вписана окружность радиуса R, касающаяся стороны AC в точке D, стороны AB – в точке E и стороны BC – в точке F.

Относительная частота встречаемости предложений Типов 1, 2 и 3 в текстах

планиметрических задач дана на Рис. 1.



Рис. 1. Частота встречаемости предложений Типов 1, 2, 3 в задачах планиметрии

Частота встречаемости различных типов предложений зависит от типов текстов. В книге Клейна [14] встретилось 14 предложений с именными и глагольными эллипсисами, из них одно предложение Типа 2, 7 предложений Типа 1 и 6 предложений Типа 3, в числе которых 2 предложения – неопределенно-личные и 4 – побудительные. Из 100 задач в [15] 30 содержат эллипсисы, из них 16 предложений Типа 1, 11 – Типа 2 и 3 – Типа 3.

Мы также сравнили относительные частоты встречаемости предложений Типов 1, 2 и 3 в геометрических текстах и художественной литературе (Рис. 2).

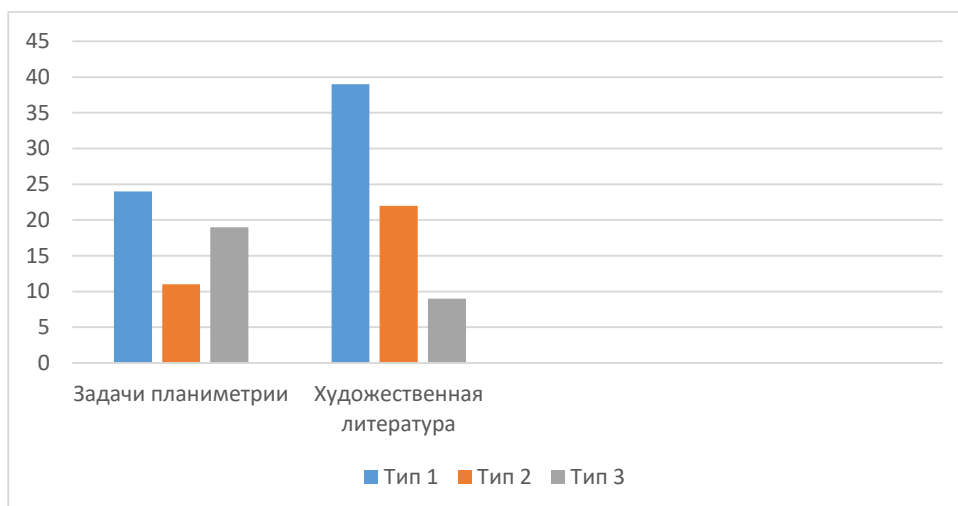


Рис. 2. Сравнение частот встречаемости предложений Типов 1, 2 и 3 в планиметрических и художественных текстах

ЦЕЛИ СОЗДАНИЯ СИСТЕМЫ

Первоначальная идея создания системы тестирования когнитивной функции (КФ), связанной с формированием эллиптических предложений, возникла из потребности разработки тестов когнитивных нарушений (КН) для ранней диагностики деменции.

По приблизительным оценкам, в РФ насчитывается около двух миллионов пациентов с деменцией, и этот показатель растёт как среди лиц старческого возраста, так и среди трудоспособного населения. Проблема сохранения и восстановления КФ является междисциплинарной и становится одной из важнейших в современной медицине [16].

Одним из самых распространенных способов определения КН является тест Мини-Ког (Mini-Cog) [17]. Он состоит из трех заданий: 1) запомнить и повторить три названных слова, например, «лимон, ключ, шар». Слова могут быть использованы любые, общеупотребительные, хорошо знакомые пациенту; 2) нарисовать часы (большой круглый циферблат) со стрелками, показывающими определенное время; 3) вспомнить три слова, которые были названы в первом задании. Для оценки результата за каждое правильное слово присваивают 1 балл. Если пациент не назвал три слова или назвал менее трех слов, предполагают КН. В таком случае необходимо более детальное обследование КФ. Однако этот тест обладает низкой чувствительностью (76%) и специфичностью (73%) [17].

КФ формирования эллипсиса опирается на удержании в памяти уже сказанного (написанного), то есть с планированием речи, и её тестирование может быть использовано для ранней диагностики деменции или других когнитивных нарушений.

Еще одна цель разработки системы тестирования способности к формированию эллипсисов связана с изучением, насколько эта способность является результатом обучения. Отсюда появилось предложение использовать сначала примеры различной сложности по типам эллипсисов, а затем давать задания преобразовать полное предложение в эллиптическое по аналогии с представленным примером. Задания могут строиться как из однотипных примеров, так и с примерами возрастающей сложности.

Наконец, еще одна цель непосредственно связана с обучением речевым технологиям. Изучение КФ формирования эллипсисов может быть использовано

в разработке автоматизированного преобразования предложений без эллипсисов в эллиптические, что может быть применимо при автоматической обработке естественно-языковых текстов с целью их сжатия или реферирования.

КОНЦЕПЦИЯ ТЕСТИРОВАНИЯ

Респонденты разбиваются на 2 группы. Одной группе предлагаются 2 теста – один за другим. Первый заключается в том, чтобы по примеру одного предложения с эллипсисом перефразировать второе предложение без эллипсиса, используя эллипсис как в первом предложении (10 заданий). Второй тест заключается в нескольких заданиях (до 10): составить текст планиметрической задачи по рисунку и/или краткому формальному описанию.

Вторая группа респондентов работает только со вторым тестом (то есть без предварительного обучения).

Пример задания в первом тесте:

Преобразуйте в каждом задании второе предложение по аналогии с первым предложением.

Пример задания в первом тесте:

- 1. Даны окружность радиуса R и касательная к ней; постройте квадрат так, чтобы две его смежные вершины лежали на касательной, а две другие – на окружности, и вычислите сторону квадрата.*
- 2. В правильную четырехугольную пирамиду вписан куб так, что четыре его вершины принадлежат боковым ребрам пирамиды, а остальные четыре его вершины принадлежат плоскости её основания.*

Пример задания во втором тесте:

Преобразуйте данное описание планиметрической задачи в текст.

«ABC – треугольник.

Дан угол B.

Дан угол C.

Проведена биссектриса угла BAC.

O – окружность проходит через точки A, B и C.

Биссектриса пересекает окружность в точке E.

Биссектриса пересекает сторону BC в точке D.

Отношение AE:DE = ?»

Далее описана автоматизированная система, предназначенная для дистанционного тестирования КФ формирования эллиптических предложений респондентом. Реализованная версия системы обеспечивает проведение экспериментального тестирования для анализа ответов респондентов и выработки методов и шкал для их оценки. Система также обеспечивает создание различных версий заданий и накопление банка предложений.

КОНЦЕПЦИЯ РАЗРАБОТКИ СИСТЕМЫ ТЕСТИРОВАНИЯ

Выбор on-line тестирования

Несмотря на то, что интернет-технологии давно стали неотъемлемой частью нашей жизни, некоторые аспекты онлайн-взаимодействия до сих пор вызывают дискуссии. А.А. Реттих [18] пришла к заключению, что онлайн-исследование является отдельным направлением психологии, так как обладает характерными чертами и особенностями.

Рассмотрим некоторые преимущества онлайн-формата, выделяемые в статье Е.Ю. Бруннер [19]:

- расширение выборки испытуемых и их географии;
- лёгкость корректировки теста;
- автоматизированные сбор и обработка данных;
- стандартизация исследования за счёт однообразности инструктажа;
- контроль времени проведения;
- изменение порядка вопросов и ответов;
- проведение тестирования в привычных для испытуемого условиях и удобное время.

Главным недостатком онлайн-формата является большая вероятность недостоверных ответов из-за анонимности пользователя. Эта проблема может быть решена с помощью внедрения системы единого входа: можно установить личность пользователя и уточнить достоверность его ответов.

Обоснование выбора действий пользователя

Для максимального комфорта пользователя следует реализовать формат с минимальными действиями, которые бы имитировали прохождение тестирования в очном формате на бумажных вариантах теста: так как полная запись нового

эллиптического предложения скорее всего покажется пользователю нудным занятием, поставим ему задачу отметить (вычеркнуть) языковую единицу, без которой предложение не потеряет своего смысла. К тому же, такой вариант избавит нас от проверки возможных орфографических или пунктуационных ошибок при вводе. Также нашей системе необходима функция экспорта результатов проверки для последующей обработки в сторонней программе.

Формирование требований к системе

1. Наличие типа вопроса «Вычеркивание», позволяющего взаимодействовать с элементами предложения через клик.
2. Наличие возможности отслеживать время, затраченное на решение каждого вопроса.
3. Наличие возможности экспортировать результаты для дальнейшей обработки в сторонних программах.
4. Наличие однозначной идентификации пользователей.

Обзор on-line ресурсов для проведения тестирования

Для проведения тестирования можно использовать существующие on-line-сервисы для составления опросов. В таблице 1 приведен результат анализа сервисов, предоставляющих возможность организации on-line тестирований.

Таблица 1. Сравнение сервисов для реализации тестирования

Показатель	Google Form [20]	Questionstar [21] и Online Test Pad [22]	Moodle Quiz
Наличие типа вопроса «Вычеркивание»	Нет	Нет	Нет
Возможность отслеживать время ответа на каждый вопрос	Нет	Нет	Нет
Возможность экспортиро-	Есть	Есть	Есть

вать результаты			
Встроенный анализ результатов	Есть	Ограничен	Есть
Идентификация пользователей	Только ввод данных, без проверки	Только ввод данных, без проверки	Есть возможность подключения системы единого входа
Возможность дополнить функционал	Исходный код недоступен, но есть возможность дописать функционал с помощью Google App	Исходный код недоступен, нельзя дополнить функционал	Доступен исходный код, можно дополнить функционал

ОПИСАНИЕ СИСТЕМЫ

Ни один из существующих on-line сервисов не соответствует в полной мере перечисленным выше требованиям к нашей системе. Таким образом, возникла необходимость прибегнуть к самостоятельной разработке веб-сервиса для тестирования.

Разработка веб-приложения с нуля освобождает нас от погружения в код работающей системы, также его всегда можно скорректировать под новые требования, устранить неточности алгоритма обработки, однако в данном случае остаётся проблема аутентификации пользователей через какую-либо систему единого входа.

Система дистанционного образования Moodle [25] способна удовлетворить следующим требованиям разрабатываемой системы:

1. Moodle обладает открытым исходным кодом, в связи с чем можно программно дополнить функционал системы.
2. Система активно используется учебными заведениями, благодаря чему можно достоверно установить личность респондента с помощью единой системы входа.

Для процедур «Вычеркивание» и «Отслеживание времени ответа» на каждый вопрос реализованы два плагина Moodle Quiz: плагин типа вопроса и плагин поведения для теста.

На Рис. 3 представлена модель работы плагинов.

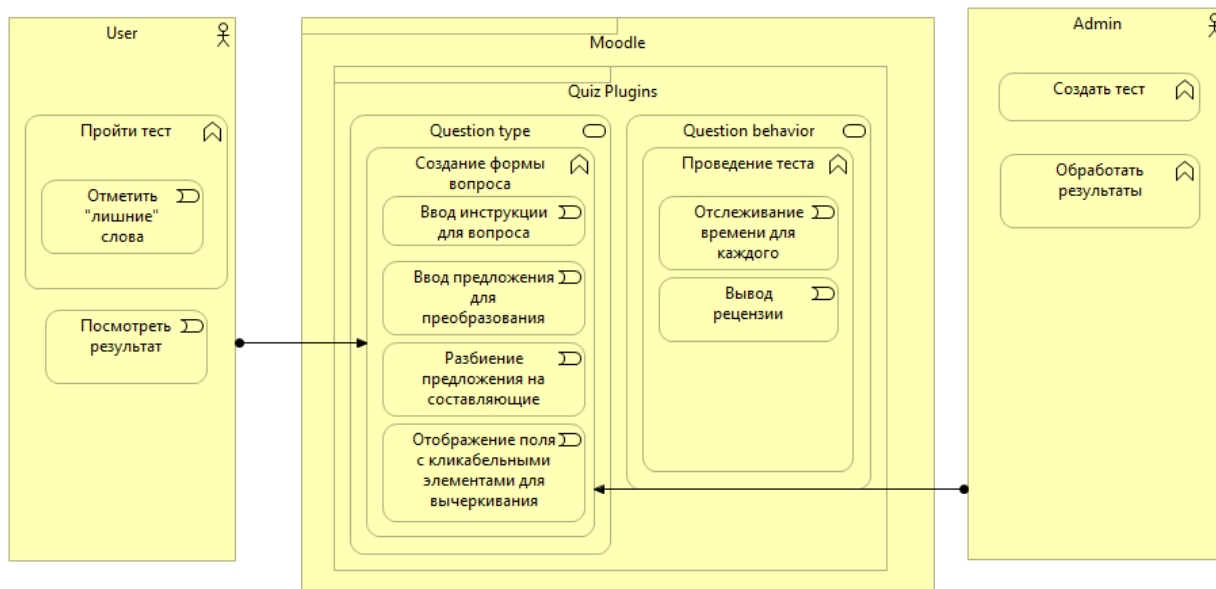


Рис. 3. Модель работы плагинов Moodle Quiz

Плагин «Question type» содержит функционал для создания формы вопроса и отображения вопроса в ходе тестирования. Вопрос включает в себя описание задания (что нужно сделать), само задание (предложение, из которого необходимо исключать слова) и ответ (какие слова респондент исключает). Под формой вопроса подразумевается веб-страница с полями для ввода данных, например, поле для ввода задания («Преобразуйте в каждом задании второе предложение по аналогии ...»). При зачеркивании, например, у элемента «~~слово~~» стилем является свойство «зачеркнутый». Плагин имитирует вычеркивание каждого отдельного слова в предложении при нажатии – динамически изменяет стиль элемента. При повторном нажатии стиль элемента возвращается к обычному, сам элемент исключается из ответа. Повторное нажатие необходимо, если пользователь передумал и решил не исключать слово из предложения.

Разбиение предложения на составляющие подразумевает выделение отдельных элементов предложения (слова, знаки препинания, числа и проч.). В результате разбиения получается список элементов, в котором каждый элемент имеет свой порядковый номер (id-идентификатор), помогающий далее однозначно определить, какое именно слово пользователь хочет исключить. Каждое

слово преобразуется в html-тег, которому назначается определённый класс (`слово`), и выводится на веб-страницу. Объекты класса «striking» отмечаются как объекты, для которых обработчик события «Нажатие» запускает функцию «Удаление» или «Восстановление» этих объектов. Таким образом получается поле, содержащее «кликабельные» элементы предложения.

Запись ответов респондента при создании нового вопроса происходит в отдельные поля, имеющие определённый id (answer_1, answer_2 и т. д.), каждый ответ автоматически оценивается, исходя из общего количества отмеченных слов (оценка для каждого отмеченного слова получается по формуле: $100\% / \text{количество отмеченных слов}$). При повторном нажатии на элемент предложения данный элемент исключается из списка ответов, остальные ответы переоцениваются. По желанию оценки могут быть изменены, например, если глагол должен быть исключён обязательно, а исключение дополнения глагола является опциональным, можно назначить оценки 75% и 25%. Сумма оценок при этом обязательно должна составлять 100%, это проверяется на сервере при внесении вопроса в базу данных. Если сумма не равна 100%, вопрос не будет внесён в базу, а на веб-странице с формой создания вопроса отобразится сообщение о необходимости следования данному условию. На Рис. 4 представлен пример автоматической оценки в случае выбора трёх слов для исключения.

Если в ходе тестирования респондент возвращается к вопросу и исправляет ответ, предыдущая попытка ответа записывается в таблицу базы данных с историей ответов, при этом новый ответ воспринимается как окончательный, и только он подлежит оценке. История ответов не отображается респонденту, но видна администратору в формате: «ответ, время подачи ответа». Ответы сохраняются на сервере только тогда, когда пользователь жмёт на кнопки «Следующий вопрос» или «Предыдущий вопрос», т.е. если респондент отмечает три слова и, не переходя к другому вопросу, решает изменить свой выбор в пользу исключения только одного слова, первый вариант ответа не будет отображён в истории.

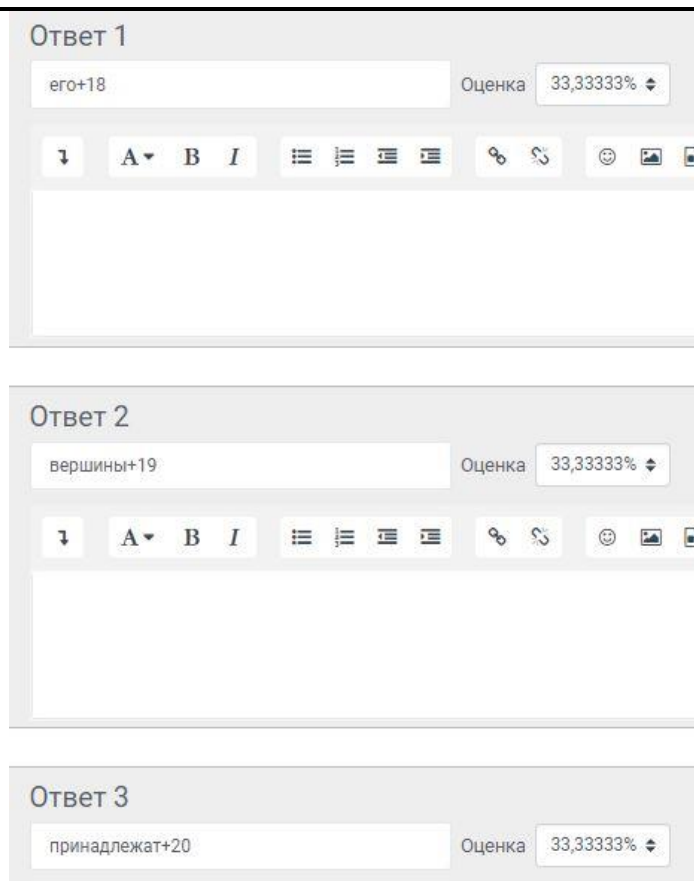


Рис. 4. Пример автоматической оценки в форме создания вопроса в случае исключения трёх слов

По завершению тестирования респонденту отображается веб-страница (документ с html-разметкой) со всеми вопросами теста, полученными баллами (в сумме за все задания и отдельно по каждому вопросу) и правильными ответами, которые были отмечены администратором при вводе заданий. Отображение верных ответов можно отключить в настройках тестирования.

Оценка ответов осуществляется следующим образом: сервер извлекает из базы данных верные ответы и проверяет их включение в ответ респондента (напомним, что ответ представляет собой строку «слово+позиция, слово+позиция ...»). Так, если на этапе создания вопроса для исключения было отмечено два слова из исходного предложения, а респондент отметил только одно, — он получит 50% от балла, назначенного в целом за этот вопрос (по умолчанию каждый вопрос даёт 1 балл, в данном случае получится 0,5 балла).

Кроме того, на веб-странице с итогами тестирования могут быть отражены рецензии (рецензия на результат, на отдельные вопросы). Рецензия на весь ре-

зультат является текстовой интерпретацией итогов, например, если итоги тестирования подразумевают конвертацию в систему оценок «удовлетворительно», «хорошо», «отлично». Рецензия на определённый вопрос может быть дана администратором или преподавателем после завершения теста или указана при создании вопроса в поле «Обратная связь». В рецензии может быть отражено всё, что захочет передать преподаватель (похвала за решение сложного вопроса, рекомендации).

Плагин типа вопроса был реализован на основе плагина «Короткий ответ», содержащегося в библиотеке типов вопроса Moodle. В ходе разработки были изменены следующие функции:

1. функция `definition_inner`, отвечающая за вывод формы создания нового вопроса на веб-страницу. В вывод было добавлено новое поле для предложения, из которого респонденту предстоит исключать слова;
2. функция `validation`, проверяющая ответы;
3. функция оценивания ответа респондента `grade_response` — функция, подсчитывающая сумму оценок за верные ответы;
4. функция `get_correct_answer`, возвращающая правильный ответ. Разработанная функция возвращает объект, содержащий все возможные для исключения слова, отмеченные администратором при создании вопроса. Данная функция вызывается на сервере функциями `grade_response` (подсчёт баллов) и `correct_response` (отображение верных ответов в итогах); (с респондентом эта функция не взаимодействует, она работает только на сервере);
5. функция `formulation_and_controls`, отвечающая за отображение вопроса в ходе тестирования и по его завершению. В этой функции реализовано разбиение предложения на «кликабельные» элементы;
6. функция `correct_response`, выводящая верный ответ пользователю. Поскольку была изменена функция для получения верного ответа `get_correct_answer`, возникла необходимость изменить также функцию вывода ответа на веб-страницу с итогами тестирования. Даже если был дан неверный ответ, на странице итогов отобразится верный ответ. Это полезно для анализа ошибок респондента. Если мы не хотим показывать

пользователю правильные ответы (вдруг расскажет другим), администратор может отключить отображение правильных ответов в настройках тестирования.

Для реализации динамических операций (изменение стиля для отмеченных слов, обновление полей с ответами) были разработаны специальные модули AMD (Asynchronous Module Definition), содержащие код на языке программирования Javascript. Модули включают в себя функции, обрабатывающие событие нажатия на элементы предложения, событие изменения поля ввода предложения.

Плагин поведения (Question behavior plugin) должен создавать секундомер для каждого вопроса, прекращать отсчёт времени при переходе на другой вопрос и запускаться снова при возврате на страницу вопроса.

Плагин поведения в реализованной системе является модернизацией плагина «Отложенный отзыв». Поведение «Отложенный отзыв» позволяет автоматически оценивать все вопросы по завершению теста и дополнять рецензии преподавателя после завершения. В программный код были добавлены функции, создающие в Cookie переменные для хранения времени ответа на каждый вопрос и сохраняющие время в базе данных после завершения тестирования. Код Javascript каждую секунду увеличивает значение переменной в Cookie, имя переменной передаётся при вызове функции на сервере. Необходимость использования Cookie связана с тем, что сохранение данных в Moodle происходит в функциях специального скриптового языка программирования PHP при нажатии на определённые кнопки (например, кнопка «Следующий вопрос»). Так, если пользователь закроет браузер, данные о времени не будут сохранены в базе данных, однако сохранятся в cookie браузера.

На Рис. 5 представлен интерфейс системы в ходе тестирования.

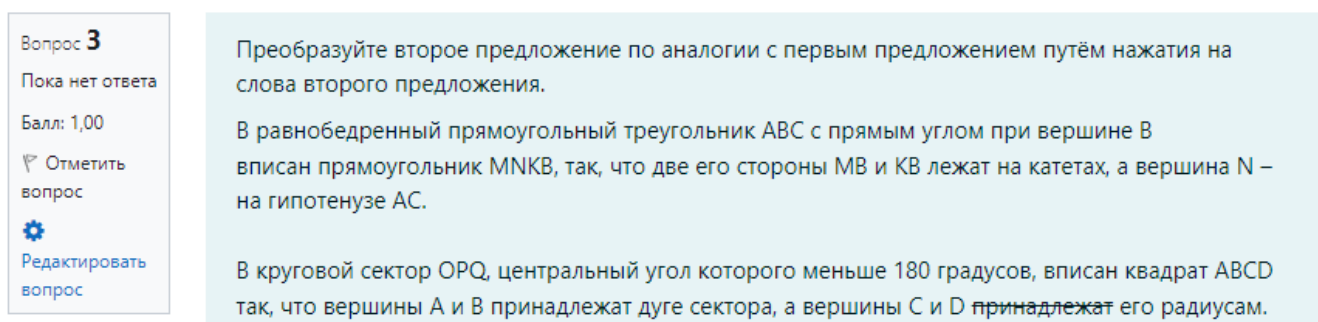


Рис. 5. Интерфейс системы

ПРОГРАММНАЯ СРЕДА РАЗРАБОТКИ

Компоненты Moodle реализованы на языках программирования PHP и JavaScript. Рассмотрим каждый из них и определим инструменты разработки.

PHP – интерпретируемый скриптовый язык программирования общего назначения. Язык применяется в разработке веб-приложений. PHP-скрипты выполняются на сервере и генерируют HTML-разметку для клиента.

Javascript – язык, используемый преимущественно для управления элементами веб-страницы. В Moodle используются AMD (Asynchronous Module Definition) модули, расширяющие возможности плагинов вопроса и поведения. Так, в нашей системе JavaScript отвечает за отображение «кликабельных» элементов предложения и меняет их стиль при взаимодействии с пользователем.

Для разработки плагина использовались:

1. PHPStorm – IDE (Integrated Development Environment), упрощающая процесс разработки благодаря автозавершению кода и автоматической проверке синтаксиса [23].
2. XAMPP – кроссплатформенная сборка веб-сервера, содержащая Apache, MySQL, интерпретатор скриптов PHP. XAMPP устанавливается автоматически при установке Moodle с помощью исполнительного файла.
3. Xdebug – средство профилирования и отладки PHP скриптов [24].

Направления развития системы

В дальнейшем в системе планируется разработать блок автоматизации формирования заданий посредством автоматической генерации эллиптических предложений из полных. Подсистема автоматизации должна включать синтаксический анализ предложений заданных типов, которые возможно преобразовать с помощью эллипсиса, искать глагольные и именные группы и элиминировать глагол и повторяющиеся фрагменты глагольных актантов. Далее эксперт должен проверить предложения после преобразований и подтвердить внесение их в базу заданий. Затем предложения и исключенные слова с помощью программы преобразуются в формат Moodle XML и загружаются в систему тестирования.

С пополнением банка предложений с эллипсисами и их анализом предполагается использовать систему для разработки алгоритмов восстановления эл-

липсисов. Эти задачи связаны с усовершенствованием синтаксических анализаторов русского языка или их расширением системами синтаксической предобработки или/и постобработки предложений. Таким образом, предлагаемая система может развиваться со временем в специализированную многофункциональную систему обработки естественно-языковых текстов с элементами обучения.

В настоящее время система подготовлена к апробации: проведение on-line тестирования на представительной выборке (выборках) респондентов на заданиях с предложениями с эллипсисами различной сложности, формирование и формализация оценок результатов тестирования для различного применения: когнитивное тестирование с целью ранней диагностики умеренных КН, с целью обучения владением эллипсисами в письменной речи, краткости и точности изложения.

ЗАКЛЮЧЕНИЕ

Представлены проектирование и разработка компьютерной системы тестирования когнитивных способностей респондентов на основе предложений с эллипсисами. Рассмотрены как вопросы формирования стимульного материала, так и проектирование системы в целом, выбор подходящих технологий для реализации системы. В качестве базы для реализации выбрана Moodle, так как эта система поддержки дистанционного образования распространяется с открытым исходным кодом и получила широкое распространение как в мире в целом, так и в России. Выбран формат разработки в виде плагинов к модулю Quiz. Данный подход позволил существенно сэкономить усилия на разработку вопрос-ответной системы тестового типа.

Исходный код плагинов размещен в открытом доступе в интернете (<https://github.com/R-D4S/Plugins-for-Ellipsis-Test>). Работа по реализации системы выполнена Е.С. Булыкиной. В настоящее время завершены внедрение и наладочные работы в Moodle СПбПУ, планируется апробация разработанных плагинов и системы в целом.

СПИСОК ЛИТЕРАТУРЫ

1. *Naidenova X.* Cognitive elements in forming and understanding sentences // CEUR Workshop Proceedings. 2021. V. 2910. P. 65–73. URL: <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-2910/short6.pdf>.
 2. *Падучева Е.В.* О семантике синтаксиса. М.: URSS, 2019. 296 с.
 3. *Грудева Е.В.* Избыточность текста, редукция, эллипсис (на материале русского языка) // Автореферат дисс. на соискание ученой степени к. филологических наук. СПб. 2008. С. 1–41.
 4. *Путинцева А., Ковригина Л., Шилин И.* Автоматическая классификация эллиптических конструкций в русской спонтанной речи // CEUR Workshop Proceedings. 2018. V. 2233. P. 120–138. URL: https://ceur-ws.org/Vol-2233/Paper_16.pdf.
 5. *Бабайцева В.В., Максимов Л.В.* Современный русский язык. Учеб. для пед. институтов по спец. № 2101. Часть 3. Синтаксис. Пунктуация. М.: Просвещение, 1987. 256 с.
 6. *Кобзарева Т.Ю., Епифанов М.Е., Лахути Д.Г.* Восстановление грамматических эллипсисов при синтаксическом анализе // Труды 14-ой национальной конференции по искусственному интеллекту с международным участием. Т. 1. Казань, Татарстан, 24–27 сентября, РИЦ «ШКОЛА», 2014. С. 108–116.
 7. *Jackendoff R.* Semantic and cognition. Cambridge: MIT Press, 1983. 283 p.
 8. *Jackendoff R.* Information is in the mind of the beholder // Linguistic and Philosoph. 1985. V. 8. No. 1. P. 23–33.
 9. *Jackendoff R.* Grammar as evidence for conceptual structure // Hall, Bresnan and Miller (Eds.). Linguistic theory and psychological reality. Cambridge, UK, MIT Press: 1978, P. 201–228.
 10. *Шварц Е.* Московская телефонная книга. М.: Изд-во АСТ, 2016. 368 с.
 11. *Эко У.* Роль читателя. Исследования по семиотике текста. М.: Изд-во АСТ, 2015. 640 с.
 12. *Гвоздев А.Н.* Современный русский литературный язык. Часть II. Синтаксис. М.: Учпедгиз, 1958. 352 с.
 13. *Mel'cuk I.* Lexical functions: a tool for the description of lexical relations in the Lexicon // L. Wanner (Ed.), Lexical functioning, lexicography and NLP. Amsterdam /Philadelphia: Benjamins, 1996. P. 37–102.
-

14. Клейн Ф. Элементарная математика с точки зрения высшей. Геометрия. Второе издание. М.: Наука, 1987. 416 с.

15. Штейнгауз Г.: Задачи и размышления. М.: Мир, 1974. 399 с.

16. Курбанова М., Галаева А., Стефановская Е. и др. Современные методы диагностики когнитивных нарушений // Российский семейный врач. 2020. Т. 24. № 1. С. 35–44.

17. Захаров В.В., Громова Д.О. Современные подходы к ведению пациентов с умеренными когнитивными нарушениями // Журнал неврологии и психиатрии 2017. № 3. С. 107–112.

18. Реттих А.А. Онлайн-исследование как новый этап развития психологии // Научный лидер. 2021. № 16 (18). С. 67–69.

19. Бруннер Е.Ю.: Организация психолого-педагогических и социологических исследований онлайн // Гуманитарные науки. 2021. № 3. С. 114–123.

20. Google Forms. URL: <https://docs.google.com/forms>

21. Questionstar. URL: <https://www.questionstar.ru/>

22. Online Test Pad. URL: <https://onlinetestpad.com/>

23. PHPStorm. URL: <https://www.jetbrains.com/phpstorm/>

24. Xdebug. URL: <https://xdebug.org/>

25. Moodle. URL: <https://moodle.org/?lang=ru>

DEVELOPING COMPUTER SYSTEM FOR TESTING COGNITIVE RESPONDENT'S ABILITIES BASED ON ELLIPTICAL SENTENCES

Xenia Naidenova^{1[0000-0003-2377-7093]}, Elena Bulykina^{2[0009-0009-6018-8360]},
Vladimir Parkhomenko^{3[0000-0001-7757-377X]}, Alexander Schukin^{4[0000-0002-9534-824X]},
Tat'yana Martirova^{5[0000-0003-0000-6608]}

^{1,5}Military medical academy, Saint Petersburg,

^{2,3,4}State Peter the Great Saint Petersburg Polytechnic University, Saint Petersburg,

¹bulykina06@gmail.com, ²ksennaidd@gmail.com,

³vladimir.parkhomenko@spbstu, ⁴alexander.schukin@spbstu,

⁵martta462@yandex.ru

Abstract

Cognitive abilities cover all higher human intellectual functions, including oral

and written speech. The use of ellipses in writing is associated with the ability to clearly formulate a thought, avoiding the redundancy of its expression. Testing the ability to use ellipses can be the basis, on the one hand, for assessing the cognitive operation of generating ellipses in writing and detecting violations of this operation. This can serve as the basis for the early detection of dementia in medicine. On the other hand, such testing can be used to teach speech technologies, including in foreign languages. From a scientific point of view, it is important to determine the influence degree of training on the performance of text tasks. The article describes the development of a computer system for remote on-line testing the operation of converting a complete sentence into a sentence with an ellipsis of a certain type, the principles of evaluating the results of testing the subjects. The concept of the system is discussed. The structure of system and software development environment are described. A feature of the system is the creation of a new type of tasks related to natural language processing. The complexity of the development is determined by the need for automated formation of a bank of sentences' pairs in natural language without ellipses and with ellipses of a certain type, which entails the inclusion, in the software environment, of text parsing tools and algorithms for automated formation of ellipses and / or restoration of complete sentences from elliptical ones. In conclusion, the prospects for the development and application of the proposed system are discussed.

Keywords: *ellipsis, natural language processing, on-line testing, cognitive abilities*

REFERENCES

1. *Naidenova X.* Cognitive elements in forming and understanding sentences // CEUR Workshop Proceedings. 2021. V. 2910. P. 65–73. URL: <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-2910/short6.pdf>.
2. *Paducheva E.V.* O semantike sintaksisa. M.: URSS, 2019. 296 s.
3. *Grudeva E.V.* Izbytochnost' teksta, reduktsiya, ellipsis (na materiale russkogo yazika) //Avtoreferat na soiskanie uchenoy stepeni k. filologicheskikh nauk. SPB. 2008. C. 1–41.
4. *Putintseva A., Kovrigina L., Shilin I.* Automatic Classification of Elliptical Constructions in Russian Spontaneous Speech // CEUR Workshop Proceedings. 2018. V. 2233. P. 120–138. URL: https://ceur-ws.org/Vol-2233/Paper_16.pdf.

5. Babaytseva V.V., Maksimov L.V. *Sovremennyy ruskiy yazik. Ucheb. dlya ped. Institutov po spets. No 2101. Chast' 3. Sintaksis. Punktuatsiya*. M.: Prosveschenie, 1987. 256 s.

6. Kobzareva T.Yu., Epifanov M.E., Lakhuti D.G. *Vosstanovlenie grammaticheskikh ellipsisov pri sintaksicheskom analize // Trudy 14-oy natsional'noy konferentsii po iskusstvennomu intellektu s mezhdunarodnym uchastiem. T. 1. Kazan', Tatarstan, 24–27 Sentyabrya, Ritz «Shkola», 2014. S. 108–116.*

7. Jackendoff R. *Semantic and cognition*. Cambridge: MIT Press, 1983. 283 s.

8. Jackendoff R. *Information is in the mind of the beholder. // Linguistic and Philosoph.* 1985. V. 8. No. 1. P. 23–33.

9. Jackendoff R. *Grammar as evidence for conceptual structure // Hall, Bresnan and Miller (eds). Linguistic theory and psychological reality. Cambridge, UK, MIT Press: 1978. P. 201–228.*

10. Shvarts E. *Moskovskaya telefonnaya kniga*. M.: AST, 2016. 368 s.

11. Eko Y. *Rol' chitatelya. Issledovaniya po semiotike teksta*. M.: AST, 2015 640 s.

12. Gvozdev A.N. *Sovremennyy ruskiy literaturnyy Yazik. Chast' II. Sintaksis*. M.: Uchpedgiz, 1958. 352 s.

13. Mel'cuk I. *Lexical functions: a tool for the description of lexical relations in the Lexicon // L. Wanner (Eds.) Lexical functioning, lexicography and NLP. Amsterdam /Philadelphia: Benjamins, 1996. P. 37–102.*

14. Klein F. *Elementarnaya matematika s tochki zreniya vysshey. Geometriya. Vtoroe izdanie*. Moskva: Nauka, 1987, 416 s.

15. Shteingauz G. *Zadachi i rasmyshleniya*. Moskva: Mir, 1974, 399 s.

16. Kurbanova M., Galaeva A., Stefanovskaya E. *at al. Sovremennye metody diagnostiki kognitivnykh narusheniy // Rossiyskiy semeinyy vrach. 2020. T. 24. No. 1. S. 35–44.*

17. Zakharov V.V., Gromova D.O. *Sovremennye podkhody k vedeniyu patsientov s umerennymi kognitivnymi narusheniyami // Zhurnal nevrologii i psikhatrii 2017. No 3. S. 107–112.*

18. Rettikh A.A. *Onlayn-issledovanie kak novyy etap rasvitiya psikhologii // Nauchnyy lider. 2021. 16 (18). S. 67–69.*

19. Brunner E.Yu. *Organizatsiya psikhologo-pedagogicheskikh i sotsiologicheskikh issledovaniy onlayn // Gumanitarnye nauki. 2021. No. 3. S. 114–123.*

20. Google Forms. URL: <https://docs.google.com/forms>

21. Questionstar. URL: <https://www.questionstar.ru/>
 22. Online Test Pad. URL: <https://onlinetestpad.com/>
 23. PHPStorm. URL: <https://www.jetbrains.com/phpstorm/>
 24. Xdebug. URL: <https://xdebug.org>
-

СВЕДЕНИЯ ОБ АВТОРАХ



НАЙДЕНОВА Ксения Александровна – к. т. н., старший научный сотрудник, Военно-медицинская академия, Санкт Петербург.

Xenia Aleksandrovna NAIDENOVA – PhD, Senior Researcher, Military medical academy, Saint Petersburg

email: ksennaidd@gmail.com

ORCID: 0000-0003-2377-7093



БУЛЫКИНА Елена Сергеевна – студентка 4 курса ФГАОУ ВО «СПбПУ», г. Санкт-Петербург

Elena Sergeevna BULYKINA – 4th year student of Peter the Great St. Petersburg Polytechnic University, St. Petersburg

email: bulykina06@gmail.com

ORCID: 0009-0009-6018-8360



ПАРХОМЕНКО Владимир Андреевич – старший преподаватель, Санкт-Петербургский политехнический университет Петра Великого, Санкт Петербург

Vladimir Andreevich PARKHOMENKO – Senior Lecturer, Peter the Great Saint Petersburg Polytechnic University, Saint Petersburg

email: vladimir.parkhomenko@spbstu.ru

ORCID: 0000-0001-7757-377X



ЩУКИН Александр Валентинович – к. т. н., доцент, Санкт-Петербургский политехнический университет Петра Великого, г. Санкт-Петербург.

Alexander Valentinovich SCHUKIN – PhD, associate professor, Peter the Great Saint-Petersburg Polytechnic University, Saint-Petersburg,

email: alexander.schukin@spbstu.ru

ORCID: 0000-0002-9534-824X



МАРТИРОВА Татьяна Александровна – младший научный сотрудник, Военно-медицинская академия, Санкт Петербург

Tan'yana Aleksandrovna MARTIROVA – Researcher, Military medical academy, Saint Petersburg

email: martta462@yandex.ru

ORCID: 0000-0003-0000-6608

Материал поступил в редакцию 15 января 2023 года