

УДК 81.23+004.82

СТРУКТУРА МЕНТАЛЬНЫХ РЕПРЕЗЕНТАЦИЙ: ИЗВЛЕЧЕНИЕ ТЕКСТА ИЗ ПАМЯТИ, НЕЙРОННАЯ СЕТЬ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Вардан Геворгович Арутюнян

аспирант кафедры славяно-русской филологии

Балтийский федеральный университет им. И. Канта

236041, Калининград, ул. А. Невского, 14. vardan.arutyunyan89@gmail.com

Рассматривается проблема организации ментального лексикона человека с позиций коннекционистского подхода. Обосновывается положение, согласно которому представительство знаний в голове индивида опирается на специфическое построение нейронной сети и особое устройство мозга. Анализируются существующие точки зрения на структуру и характер «единиц» знания, а также предлагается концепция их функционирования. В связи с этим обсуждается вопрос извлечения текста как цельного психолингвистического образования из долговременной памяти и описывается модель этого процесса. Отстаивается гипотеза, что основным в организации ментального пространства человека является ассоциативно-семантический сетевой принцип связи. На основе имеющихся данных разрабатывается методика моделирования структур представления знаний в системах искусственного интеллекта.

Ключевые слова: ментальный лексикон; долговременная память; ассоциативно-семантические сети; ментальные репрезентации; человеческий мозг; искусственный интеллект; коннекционизм.

Проблема организации структур представления знаний / ментальных репрезентаций стала одной из центральных в когнитивной науке XXI в. При этом очевидно, что междисциплинарный подход в исследованиях – единственно релевантный, поскольку приблизиться к пониманию фундаментальных принципов работы мозга, а соответственно, и языка, мышления, сознания и других психических феноменов, невозможно с позиции какой-либо одной науки: «Синтез гуманитарного и естественно-научного знания является не факультативной декорацией и данью моде, а необходимым условием научного прогресса: кардинальные вопросы лингвистики не могут быть решены без учета фактов биологии и психологии...» [Черниговская 2009: 329–330].

В связи с этим возникает ряд важных вопросов: как представлены в человеческом мозге (долговременной памяти) ментальные репрезентации? Как они организованы и какими нейронными механизмами обеспечиваются? Что такое ментальная «единица» знания? И достаточно ли у нас оснований для того, чтобы говорить о каких-либо «единицах» в традиционном смысле? Как функционирует долговременная память и как она устроена? Каким образом про-

исходит процесс извлечения *текста* из памяти? И возможно ли все эти психические операции, реализующиеся в мозге, моделировать в системах искусственного интеллекта? Все перечисленные вопросы будут обсуждаться в данной статье.

Первое положение, которое мы развиваем, заключается в идее связности и взаимообусловленности ментальных феноменов: мышление, язык, память и другие высшие психические функции представляют собой континуум. Кроме того, для того чтобы выяснить, как они устроены *на самом деле*, необходимо учесть принципы организации биологического субстрата, в котором они реализуются. Известно, что мозг представляет собой огромную и сложную нейронную сеть, в которой при выполнении когнитивных задач нейроны объединяются во взаимодействующие специфические сетевые ансамбли и функциональные системы; связи же между ними устанавливаются в основном по ассоциативному принципу [Анохин 1968; Шульговский 2000].

Именно поэтому мы убеждены, что кортикальное представительство языка, как и знаний вообще, – это *сетевое представительство*. Человеческий мозг является единой нейронной се-

тью, лишенной каких-либо модулей или подмодулей. Соответственно, мы считаем, что в дискуссии о специфике организации ментального лексикона явный приоритет за сторонниками односистемного подхода (о двусистемном и односистемном подходе см. работы: [Bybee 1995; Caramazza et al. 2001; Chernigovskaya, Gor 2000; Plunkett, Bandelow 2006; Prasada, Pinker 1993; Ullman et al. 1997] и др.).

Важно при этом, что исследования, которые ведутся в когнитивной психологии, подтверждают тот факт, что долговременная память представляет собой сложную сеть: «... семантическая память состоит из обширной сети понятий, которые составлены из единиц и свойств и соединены рядом ассоциативных связей» [Солсо 2012: 306]. Кроме того, память рассматривается как «хранилище» ментальных репрезентаций: «... [мы] будем употреблять термин “память” в ограниченном смысле, применительно к представлениям, хранящимся в долговременной памяти, включая в данное понятие и процессы (нижнего уровня), которые обеспечивают доступ к информации, записанной в этих представлениях» [Скрэгг 1983: 228]. Память – «это способность сохранять накопленный опыт и знания в виде “следов” (энграмм) когнитивных и ментальных репрезентаций как определенных структур представления знаний» [Кубрякова 1997б: 114].

Таким образом, память является семантической сложноорганизованной сетью понятий, которые связаны друг с другом ассоциативными связями. Установив общий структурный характер долговременной памяти, необходимо разобраться, *что* хранится в этой сети, *какого рода* знания представлены в ней и *какова специфика* ментальных репрезентаций.

Отсюда второе положение, которое мы отстаиваем, – отрицание каких-либо *единиц* знания или ментальных *единиц*. Обратимся, например, к определению, которое дается в Кратком словаре когнитивных терминов: «Концепт – термин, служащий объяснению *единиц* (курсив наш. – В. А.) ментальных или психических ресурсов нашего сознания и той информационной структуры, которая отражает знание и опыт человека; оперативная содержательная *единица* памяти, ментального лексикона, концептуальной системы и языка мозга (*lingua mentalis*)» [Кубрякова 1997а: 89].

Некоторые ученые считают, что в ассоциативно-семантической сети представлены именно концепты, которые и являются единицами знания: «... она [семантическая сеть] представляет собой ассоциативную организацию связей, точки пересечения которой называются узлами. Каж-

дый узел мыслится как представляющий, или репрезентирующий определенный концепт» [Панкрац 1997б: 169]; «В своей наипростейшей форме семантическая сеть есть совокупность точек, называемых узлами; каждая из них может мыслиться как представление некоторого понятия» [Скрэгг 1983: 230].

Однако с развитием *коннекционистского* подхода в когнитивной науке произошло кардинальное переосмысление данного положения: были выявлены значительные недостатки ранних версий семантических сетей и показана их упрощенность.

Коннекционизм как один из подходов в когнитивной науке возник в 1980-х гг., когда Дж. МакКлелланд и Д. Румельхарт разработали «первую коннекционистскую модель как модель интерактивной обработки информации, происходящей при активации и распространении активации по узлам нейронной сетки» [Панкрац 1997б: 87; McClelland, Rumelhart 1981]. Одно из ключевых положений, которое выдвинули ученые, – идея параллельно распределенного процесса обработки информации: «... работа модели ведется не пошагово, а параллельно, т.е. с одновременным возбуждением разных участков нейронной сетки» [Панкрац 1997б: 87].

Исследователи подчеркивают, что «мыслительные процессы протекают в мозге, состоящем из десятков миллиардов связанных между собой нейронов. Эти относительно простые нейроны, которые взаимодействуют с сотнями тысяч других нейронов, являются основой сложной обработки информации» [Солсо 2012: 48]. Известно, что скорость передачи нервных импульсов не велика: «Нервная передача – относительно медленный, подверженный помехам процесс, и некоторым нейронам требуется 3 мс для генерализации разряда» [там же: 49]. Так как же происходят сложные когнитивные процессы при таких маленьких скоростях? Коннекционисты дали убедительный ответ – за счет параллельной обработки информации.

Кроме того, было показано, что сами понятия, *единицы знания, не содержатся* в сети: «...в коннекционистских моделях сами паттерны не хранятся; хранится сила связи между единицами, которая позволяет восстановить эти паттерны» [там же: 323]; «Репрезентация какого-либо концепта здесь не является постоянной и “записанной” в самой сети – она возникает благодаря активации и возбуждению взаимосвязанных элементов. Будучи активизированной, каждая единица способна возбуждать или же гасить возбуждение других связанных с нею единиц, а весь активизированный участок сетки соответствует необходимой модели

или структуре знания» [Панкрац 1997а: 87]; «... концепты не содержатся в сети представления знаний, но их значения могут быть сконструированы при помощи их позиций в сети» [Kintsch 1988: 165].

Благодаря коннекционистскому подходу стало понятно, что узлы (точки пересечения) семантической сети – не концепты или ментальные репрезентации. Это некие *признаки / свойства* понятий, которые активируются и в свою очередь возбуждают или тормозят связанные с ними узлы, а весь участок активируемой сети и является структурой знания. Более того, никакая ментальная репрезентация не существует и не хранится автономно в долговременной памяти. Тот обширный участок сети, в котором представлена структура знания, имеет связи с бесчисленным количеством других и в речемыслительной деятельности притягивает целый кластер разнородной информации. Именно поэтому мы считаем, что идея каких-либо *единиц* в ментальном пространстве нерелевантна. Это подтверждается и ранними работами в области семантических сетей: «Понятие обладает информационным содержанием лишь в силу того, что оно связано с другими узлами. Можно считать, что информация существует в отношениях. Понятие, не участвующее ни в каких отношениях, лишено содержания, и мы считаем, что доступ к нему закрыт. Это понятие, о котором ничего не известно» [Скрэгг 1983: 232].

Таким образом, коннекционистские модели отличаются от других по ряду значимых параметров:

1. Обработка любой информации происходит не пошагово, а параллельно, с активацией многих частей сети и установлением между ними ассоциативных связей.

2. В долговременной памяти содержатся не готовые образы, понятия и ментальные репрезентации, а сложная иерархическая система связей, благодаря которой возможно восстановление паттернов.

3. Модель основана не на искусственных схемах, а на представлениях об организации головного мозга, в котором нейроны при когнитивной работе соединяются во взаимодействующие сетевые ансамбли.

4. Алгоритмы функционирования коннекционистских схем подтверждены компьютерными моделями, в которых значимое место занимает принцип ассоциативности [Zock, Vilas 2004].

В связи с этим встает вопрос о представительстве *текста* в долговременной памяти человека. Каким образом подобное цельное психолингвистическое образование хранится в ассоциативно-семантической сети ментального пространства? Если даже отдельная структура представления знания существует в нейронной сети в качестве слож-

ного психического феномена, то как репрезентирован целый текст? И как происходит извлечение текста из долговременной памяти при вспоминании? Ниже мы представим нашу модель реализации данного процесса.

Современные модели ментального лексикона, разрабатываемые в рамках психо- и нейролингвистики, опираются на принцип категориальной иерархии. Экспериментальные исследования показали, что «... элементы, являющиеся точками пересечения наибольшего количества связей, составляют ядро ментального лексикона – самую активную его часть» [Золотова 2005: 43]. Было установлено, что в ядро ментального лексикона входят самые емкие и общие понятия; через ядро могут устанавливаться контакты между не связанными друг с другом словами и репрезентациями; ядро обеспечивает экономность хранения энциклопедических знаний и эффективность параллельного учета их в речемыслительной деятельности; а также ядро является «запускающим механизмом», откуда начинается поиск нужного слова [там же]. Соответственно, условная модель представляет собой иерархическую структуру с ядром в верхней части.

Полагаем, что текст, как и отдельная ментальная репрезентация, хранится распределенно по всей нейронной сети (см.: [Арутюнян 2013а]). Извлечение же текста из памяти – это сложный психологический процесс, включающий множество операций и реализующийся в несколько этапов.

Первый шаг вспоминания связан с активацией семантического ядра текста, т. е. тех участков сети, которые репрезентируют ключевые слова и понятия. Как правило, это наиболее общие и емкие структуры представления знаний, входящие в ядро ментального лексикона, которое представляет собой узлы, находящиеся на верхней ступени иерархической структуры, связи из которых распространяются вниз от самых общих / фундаментальных и наиболее активно участвующих в речемыслительной деятельности понятий к более частным (по принципу категориальной иерархии).

Проиллюстрируем это на простом примере: «*Женищина подвинула стул*». Несмотря на то что в данном случае речь идет о *женщине*, изначально будет активирована та часть сети, в которой представлено понятие *человек* (общее), которое, в свою очередь, возбудит связанную с ним область – *женщина* – и затормозит остальные – *мужчина, мальчик, девочка* и т.д. Необходимо отметить, что параллельно будут активированы другие области сети: те, которые будут конструировать концепты *делать* (общее) и *предметы*

(общее). При этом они активируют связанные с ними ментальные репрезентации *подвинула* и *стул*, затормозив все остальные. Следует отме-

тить, что все активированные участки сети сразу будут соединяться друг с другом ассоциативно-семантическими сетями (рис. 1).

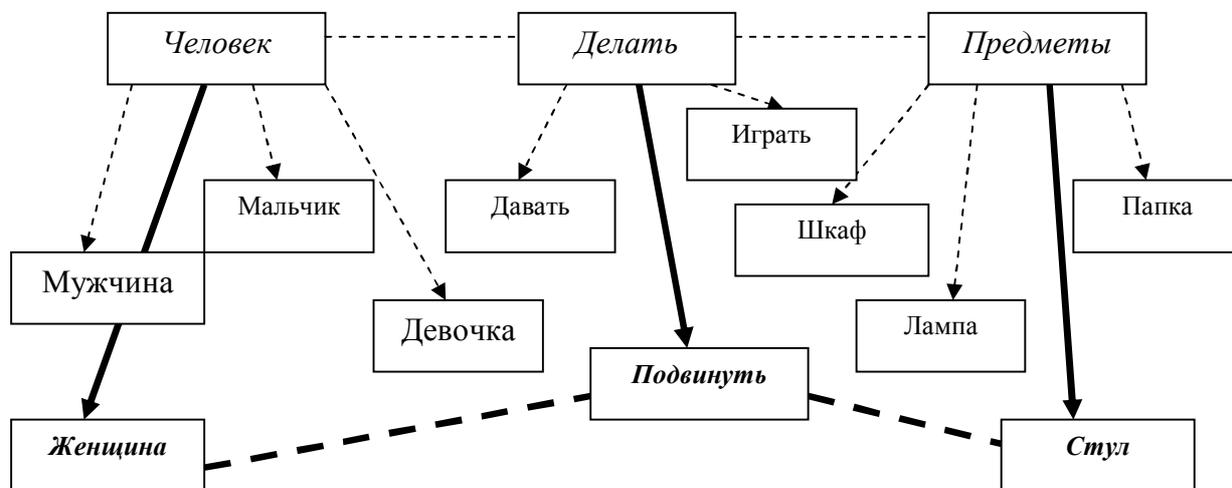


Рис. 1. Модель извлечения текста из долговременной памяти (на примере одного предложения)

Второй этап включает разрастание сети: активируется все больше узлов и частей сети, связываются друг с другом частные понятия. Здесь может происходить процесс *замены деталей текста*. Понятия, входящие в семантическое ядро текста, практически не подвержены этому процессу, но частные ментальные репрезентации подвержены. Нам кажется, что это можно объяснить *теорией распространения активации*, выдвинутой А. Коллинзом и Э. Лофтус [Collins, Loftus 1975]. Она гласит, что сила и расстояние ассоциативной связи между элементами различна: между понятиями распространяется активация и те концепты, которые расположены рядом (*огонь – пожар*), связаны сильнее, чем другие (*огонь – карандаш*).

Таким образом, чем ближе будут находиться частные ментальные репрезентации, тем больше вероятность их замены. Например, если в тексте говорилось о том, что *летом* лил сильный дождь, то позже при вспоминании *лето* может замениться на *осень*, поскольку *сильный дождь* связан более сильными ассоциативно-семантическими связями с *осенью*, а не с *летом*¹.

Итак, извлечение текста из долговременной памяти – это сложный психический процесс, реализующийся в несколько этапов. Во-первых, происходит активация сети в разных участках и параллельная обработка информации – установление ключевых понятий, появление связей между ними, активация узлов, связанных с этими концептами, и торможение некоторых других. Во-вторых, – разрастание сети, расширение ас-

социативно-семантических связей и зависимостей; при этом может происходить замена деталей текста.

Ранее мы обосновали, что ассоциативно-семантический сетевой принцип связи является фундаментальным в организации всего ментального пространства человека (см.: [Арутюнян 2013б]). Соответственно, данное положение в совокупности с иерархическими моделями представления знаний, а также алгоритмами извлечения текста из памяти должно лечь в основу создания *сильного искусственного интеллекта* (the strong artificial intelligence).

В настоящее время в рамках лаборатории робототехники и искусственного интеллекта Балтийского федерального университета им. И. Канта нами проводятся исследования по моделированию когнитивных процессов при помощи искусственно растущих нейронных сетей. В качестве площадки для построения и экспериментов с системой используется андроидный робот AP-600, разработанный российским научно-производственным объединением «Андроидная техника».

Для моделирования структур представления знаний (ментального лексикона в целом) нами использована описанная выше иерархическая структура организации, в которой ассоциативные связи являются фундаментальным видом контактов. Мы убеждены, что попытки создания искусственного антропоморфного интеллекта неминуемо провальны, если в алгоритмы не будут заложены принципы организации человеческого

языка.

Однако большинство исследователей не уделяет этому аспекту должного внимания: значимое место в существующих системах занимают реализация логических операций, а также принятие решений на базе алгоритмов, имеющих мало общего с функционированием мозга. Если же и предпринимались попытки моделирования языка, то в основном они сводились к тому, что в память робота загружался большой список слов (в то время как язык – это не столько слова, сколько правила и алгоритмы). В целом системы эти хороши, только они не отражают работу реального человеческого мышления: они верно выполняют некие задания, что вовсе не свидетельствует об их высоком когнитивном ранге.

Проблема заключается в том, что многие специалисты по компьютерному моделированию считают коммуникацию чуть ли не единственной

функцией языка. В то время как *главная ее функция* заключается в обеспечении мышления, а также номинации ментальных репрезентаций, благодаря чему мы можем использовать их в речемыслительной деятельности.

Модель, разрабатываемая нами, должна преодолеть недочеты других систем. В ходе наших исследований мы используем модифицированный вариант ассоциативной SOINN [Shen, Hasegawa 2009]. Входная информация с каждой системы робота поступает на соответствующую SOINN, где формируется многоступенчатая структура классов и паттернов. Кластеризованные в этом слое данные участвуют в формировании ассоциативного слоя, который содержит по одному нейрону-прототипу каждого класса из каждой системы. Наличие связи между двумя нейронами в ассоциативном слое будет означать, что они образовали ассоциацию (рис. 2).

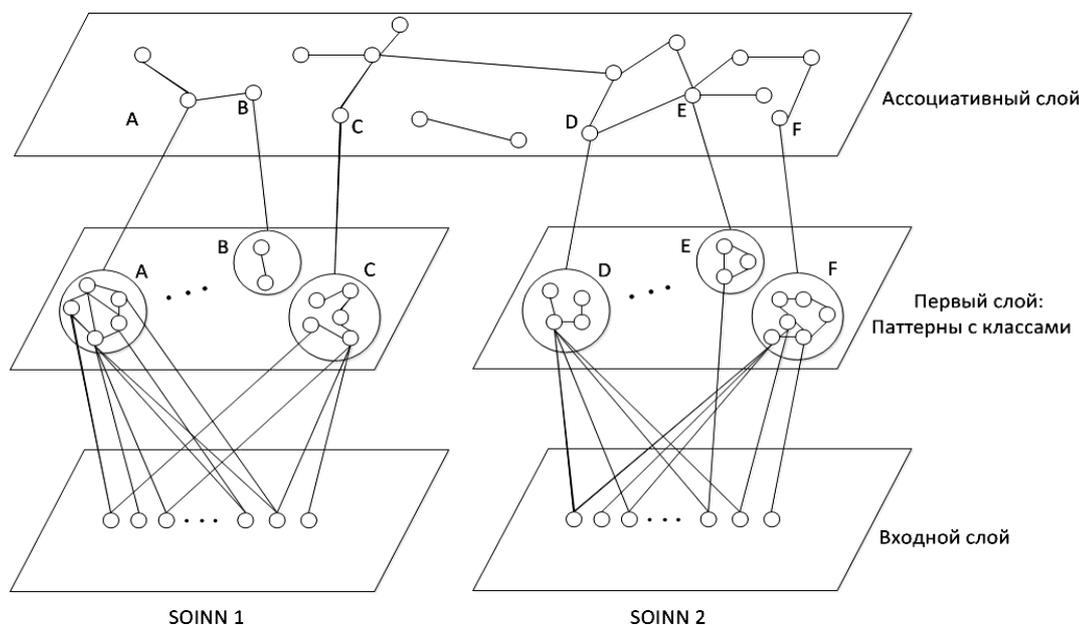


Рис. 2. SOINN с ассоциативной памятью

Таким образом, при одновременном (или с некоторой задержкой) поступлении информации на различные системы робота будут активированы определенные нейроны-прототипы. Все действующие в данный момент нейроны ассоциативного слоя будут соединены связями. Так будет строиться ассоциативная связь между, например, визуальным образом предмета (распознанным изображением телефона) и его звуковым обозначением (произнесенным в этот момент словом).

Экспериментальные исследования в области нейролингвистики показали, что ядро ментального лексикона формируется в онтогенезе в первую очередь [Золотова 2005], что придает ему статус естественного метаязыка: оно структури-

руется в первые годы жизни, а дальше происходит наращивание связей и расширение всего ментального лексикона. Поэтому мы решили пойти той же дорогой (заложить изначально в систему слова / понятия, входящие в ядро), повторив онтогенетическое развитие, а значит, реализовав модель, максимально приближенную к развитию человеческого мозга.

Обучение системы и наполнение ядра ментального лексикона будет проходить следующим образом. Роботу будут демонстрироваться различные предметы (они будут распознаваться разрабатываемой сотрудниками лаборатории системой технического зрения) и произноситься слова, обозначающие их, для построения связи между звуковым и визуальными образами. Ему

будут также давать в руки соответствующие предметы, чтобы в его памяти зафиксировались тактильные ощущения. Кроме того, прорабатывается вариант с расстановкой камер и микрофонов в помещении для непрерывного «впитывания» информации через наблюдение за находящимися в комнате людьми.

Следовательно, разрабатываемая нами модель учитывает принципы организации ментального пространства человека, опирается на функционирование мозга и представляет собой иерархическую ассоциативную конструкцию, т. е. она максимально приближена к работе человеческого мозга.

Вопрос о характере структур представления знаний является одним из самых актуальных в когнитивной науке XXI в. Мы надеемся, что результаты, которые будут получены по окончании исследований, внесут дополнительную ясность в эту проблему, а также помогут глубже понять специфику организации ментального лексикона человека.

Примечание

¹ Это было продемонстрировано в нашем собственном экспериментальном исследовании.

Список литературы

Анохин П. К. Биология и нейрофизиология условного рефлекса. М.: Медицина, 1968. 546 с.

Арутюнян В. Г. Проблема извлечения текста из памяти: ассоциативные и семантические сети (коннекционистский подход) // Альманах современной науки и образования. 2013а. №7(74). С. 12–14.

Арутюнян В. Г. Специфика организации ассоциативно-семантических сетей в человеческом сознании: обзор экспериментальных данных // *Philologia nova: лингвистика и литературоведение: сб. ст. молодых исслед.* Киров, 2013б. С. 15–20.

Золотова Н. О. Ядро ментального лексикона как естественный метаязык: дис. ... д-ра филол. наук. Тверь, 2005. 306 с.

Кубрякова Е. С. Концепт // Краткий словарь когнитивных терминов. М.: МГУ им. М.В. Ломоносова, 1997а. С. 89–93.

Кубрякова Е. С. Память // Краткий словарь когнитивных терминов. М.: МГУ им. М.В. Ломоносова, 1997б. С. 114–118.

Панкрац Ю. Г. Коннекционизм // Краткий словарь когнитивных терминов. М.: МГУ им. М.В. Ломоносова, 1997а. С. 87–89.

Панкрац Ю. Г. Семантические сети // Краткий словарь когнитивных терминов. М.: МГУ им. М.В. Ломоносова, 1997б. С. 169–170.

Скрэгг Г. Семантические сети как модели памяти // Новое в зарубежной лингвистике. М.: «Радуга», 1983. Вып. 12: Прикладная лингвистика. С. 228–271.

Солсо Р. Когнитивная психология. СПб.: Питер, 2012. 589 с.

Черниговская Т. В. Человеческое в человеке: сознание и нейронная сеть // Проблема сознания в философии и науке. М., 2009. С. 325–360.

Шульговский В. В. Основы нейрофизиологии. М.: Аспект Пресс, 2000. 277 с.

Bybee J. Regular Morphology and the Lexicon // *Language and Cognitive Processes*. 1995. №10. P. 425–455.

Caramazza A., Costa A., Nuria S.-G. et al. Regular and irregular morphology and its relationship with agrammatism: evidence from two Spanish-Catalan bilinguals // *Brain and Language*. 2001. №91. P. 212–222.

Chernigovskaya T., Gor K. The Complexity of Paradigm and Input Frequencies in Native and Second Language Verbal Processing: Evidence from Russian // *Language and Language Behavior*. 2000. P. 20–37.

Collins A., Loftus E. A Spreading-Activation Theory of Semantic Processing // *Psychological Review*. 1975. Vol. 82, № 6. P. 407–428.

Kintsch W. The Role of Knowledge in Discourse Comprehension: A Construction-Integration Model // *Psychological Review*. 1988. Vol. 95, №2. P. 163–182.

McClelland J., Rumelhart D. An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings // *Psychological Review*. 1981. Vol. 88, № 5. P. 375–407.

Plunkett K., Bandelow S. Stochastic approaches to understanding dissociations in inflectional morphology // *Brain and Language*. 2006. №98. P. 194–209.

Prasada S., Pinker S. Generalization of regular and irregular morphological patterns // *Language and Cognitive Processes*. 1993. №8(1). P. 1–56.

Shen F., Hasegawa O. Self-organizing incremental neural network and its applications. URL: <http://haselab.info/papers/soinn-tutorial.pdf> (дата обращения: 14.07.2013).

Ullman M. et al. A Neural Dissociation within Language: Evidence that the Mental Dictionary Is Part of Declarative Memory, and that Grammatical Rules Are Processed by the Procedural System // *Journal of Cognitive Neuroscience*. 1997. №9(2). P. 266–276.

Zock M., Bilac S. Word lookup on the basis of associations: from an idea to a roadmap // *Proc. COLLING.*, 2004. P. 89–95.

THE STRUCTURE OF MENTAL REPRESENTATIONS: TEXT EXTRACTION FROM MEMORY, NEURAL NETWORK AND ARTIFICIAL INTELLIGENCE

Vardan G. Arutyunyan

Post-graduate Student of Slavonic Russian Philology Department
Immanuel Kant Baltic Federal University

The paper explicates the problem of human mental lexicon organization in the aspect of the connectionist approach. It is substantiated that knowledge representation in human brain is based on a specific neural network and a particular brain structure. Present viewpoints on the structure and nature of knowledge “units” are analysed and a conception of their functioning is thus set forth. In this regard, the problem of extraction of text as a whole psycho-linguistic organization is discussed and the model of the text extraction from long-term memory is described. The paper also hypothesizes that the associative-semantic network principle plays a key role in the formation of human mental space. Based upon the data available, the author develops a methodology of modeling structures of knowledge representation in the systems of artificial intelligence.

Key words: mental lexicon; long-term memory; associative-semantic networks; mental representations; human brain; artificial intelligence; connectionism.