

УДК 004.896

MSC: 68T10; 68T50; 68T42

DOI: 10.35330/1991-6639-2020-3-95-13-22

МУЛЬТИАГЕНТНАЯ МОДЕЛЬ ФОРМИРОВАНИЯ ПЕРЦЕПТИВНОГО ПРОСТРАНСТВА В ПРОЦЕССЕ ОВЛАДЕНИЯ ЯЗЫКОВОЙ КОМПЕТЕНЦИЕЙ*

З.В. НАГОЕВ¹, И.А. ГУРТУЕВА², К.Ч. БЖИХАТЛОВ¹

¹ФГБНУ «Федеральный научный центр
«Кабардино-Балкарский научный центр Российской академии наук»
360002, КБР, г. Нальчик, ул. Балкарова, 2
E-mail: kbncran@mail.ru

²Институт информатики и проблем регионального управления –
филиал ФГБНУ «Федеральный научный центр
«Кабардино-Балкарский научный центр Российской академии наук»
360000, КБР, г. Нальчик, ул. И. Арманд, 37-а
E-mail: iipru@rambler.ru

Предлагаемая в настоящей работе модель раннего развития языковой компетенции с учетом влияния социальных факторов представляет собой имитационную модель фонематического импринтинга, описывающую процесс отображения аудиостимулов в классы элементарных единиц языка. Алгоритм машинного обучения разработан с использованием результатов исследования особенностей речи, обращенной к детям. Данная модель позволит исследовать особенности фонетического восприятия, когнитивные механизмы, лежащие в основе языкового развития, выделить основные факторы, влияющие на продолжительность периода пластичности. Предлагаемая модель дает возможность строить перцептивные карты, проектировать диагностический инструментарий для описания и исследования сенситивного периода. Модель также может быть использована для создания речевых систем, устойчивых к многообразию акцентов и эффективных при использовании в условиях высокой зашумленности.

Ключевые слова: мультиагентные системы, искусственный интеллект, искусственные нейронные сети, распознавание речи, период сенситивности, речь, обращенная к детям.

1. ВВЕДЕНИЕ

Системы автоматического распознавания речи представляют собой программные комплексы, осуществляющие транскрипцию звуковой волны речевого сообщения [1]. Поскольку речевые системы часто являются первой ступенью автоматических систем понимания, высокая точность при решении данной задачи приобретает принципиальное значение. Безусловно, применение в течение последнего десятилетия методов машинного обучения, в том числе глубокого обучения, вывело разработки в области речевых технологий на качественно новый уровень [1, 2]. Так, в [3] представлены результаты экспериментальных исследований, в которых процент ошибок, допускаемых человеком при транскрибировании телефонных разговоров, незначительно выше эффективности автоматической системы (11% против 11,3% на материалах CallHome English). Более того, отдельные эксперименты [4] на тех же тестовых данных показали, что процент ошибок речевых систем значительно ниже человеческого уровня (6,8%). Оставив в стороне формализацию вопроса о том, что считать уровнем человеческой точности, а также дискуссионные вопросы о выборе тестовых данных и условий проведения экспериментов, хотелось бы отметить различия в качестве ошибок,

* Работа выполнена при поддержке грантов РФФИ №№18-01-00658, 19-01-00648

допускаемых человеком и автоматическими системами, которые все же сводятся к различению прагматики – доля семантических ошибок у автоматических систем значительно выше.

На наш взгляд, успешное решение данной задачи возможно при условии выявления стратегий, которыми пользуются дети при овладении языковыми навыками. Необходимо исследовать критический период для фонетического обучения, чтобы создать систему, построенную на основе когнитивных функций, которыми пользуется человек при декодировании звуковых сообщений [5].

Разрабатываемая концепция опирается на результаты исследований экспериментальной психолингвистики и нейролингвистики, фундаментальные представления возрастной психологии, когнитивных наук, исследования материнской речи [5-11].

2. МУЛЬТИАГЕНТНАЯ МОДЕЛЬ ФОНЕМАТИЧЕСКОГО ИМПРИНТИНГА

Предлагаемые в настоящей работе модельные представления о приобретении фонематического слуха основаны на важнейшем открытии П. Эймаса, доказывающем, что младенцы обладают уникальной способностью различать фонетические контрасты любых языков [12]. Этот результат был подтвержден в ходе многочисленных экспериментов [5].

Мультиагентная [13-17] модель фонематического импринтинга в процессе приобретения языковых навыков, как показывает рисунок 1, состоит из трех этапов: регистрация, оценка, размещение.

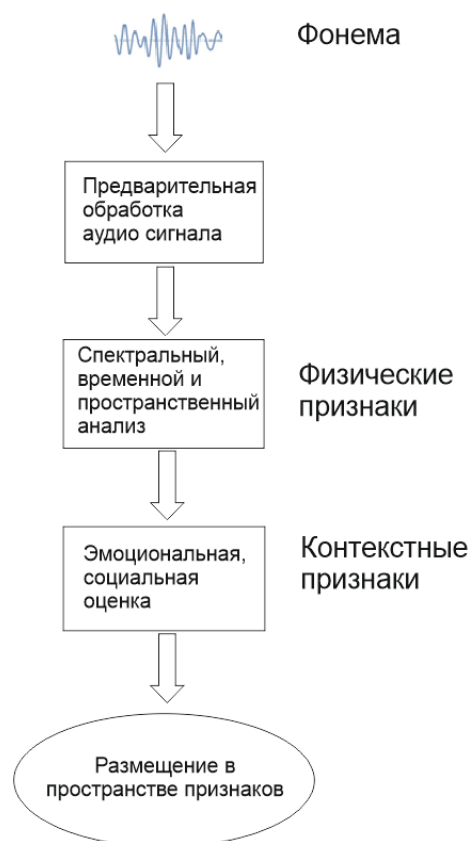


Рис. 1. Этапы обработки звукового сообщения мультиагентной моделью фонематического импринтинга

На этапе предварительной обработки звуковой сигнал преобразуется в набор сигнатур, на основе которого создается матрица, характеризующая акустические особенности сиг-

нала. Поскольку звуковую волну с достаточной полнотой можно охарактеризовать четырьмя физическими параметрами, внутримодальная дифференциация сводится к выделению в структуре звукового потока четырех слоев: амплитуды, спектрального состава, продолжительности звучания и местоположения источника сигнала [18]. Данная информация подается на вход мультиагентной системы [15], где создаются агенты-акторы, функционал которых соответствует слуховым сенсорам человека [19, 20].

Первый этап – оценка – имитирует активацию нейронов речевых зон мозга звуковыми стимулами.

В формировании представлений о принципах функционирования данного блока мы опирались на исследования материнской речи [7, 21-24].

Использование материнской речи, за очень редкими исключениями [25], наблюдается во всех языковых сообществах. Материнская речь традиционно исследуется на четырех уровнях – фонологическом, лексическом, синтаксическом и коммуникативном. Исследователи выделяют от четырех и более признаков, отличающих речь, обращенную к детям, от речи, используемой в коммуникации сверстников [6].

В рамках настоящей работы наибольший интерес вызывают исследования фонологических и просодических признаков, и, главное, принципы целесообразности, лежащие в основе коммуникации с детьми. Во-первых, преувеличение акустических диапазонов, направленное, предположительно, на формирование представлений о допустимых пределах вариативности речи. Прежде всего к данному принципу можно отнести модуляцию уровня тона, причем характерно расширение верхней границы диапазона частоты основного тона [22]. Преувеличиваются также и интонационные контуры высказываний [22, 23]. Во-вторых, упрощения и повторы, способствующие на начальном этапе освоения облегчению сегментации непрерывной речи, а затем ускорению понимания и обучения. А именно: лабиализация, исключение сложных в артикуляции согласных, звукоподражание, снижение темпа речи, наблюдаемое не только в регистре обращения к детям, но также и в речи опытных педагогов. Кроме того, отмечается тенденция к увеличению продолжительности пауз. Частотным признаком речи, обращенной к детям, также является пролонгация гласных фонем. По результатам исследований [21], средняя продолжительность произнесения слогового ядра значительно выше во время общения с ребенком. Данные исследований [5] документируют превышение пролонгации звучания гласных, реализуемых в речи матерей, по сравнению с их звучанием в коммуникации с равными по возрасту адресантами в три раза. Используются также повторы просодических паттернов. Третий принцип, устойчиво сохраняющийся в материнской речи, сводится к стимуляции социальной вовлеченности, направленной на повышение эффективности обучения и продолжительности запоминания. Речь, обращенная к детям, привлекает внимание ребенка, подает четкие эмоциональные сигналы, вовлекает в коммуникацию. При этом материнская речь не обязательно звучит громче, чем обычная. Интересно, что в процессе коммуникации взрослого и ребенка часто используется шепот. Наиболее распространенная из объясняющих это гипотез заключается в том, что матери инстинктивно используют шепот для привлечения внимания. В большинстве случаев так звучат заключительные слоги в последней синтагме, иногда все высказывание произносится шепотом. В разговоре с двухлетками шепотом звучит каждая шестая фраза. В работе [24] показано, что при взаимодействии с детьми используются повторяющиеся ритмические структуры, рифмы, удвоение звуков, провоцируют младенцев на синхронную двигательную активность, что также способствует более эффективному овладению языковыми навыками. Несомненно, наибольшее влияние оказывает предпочтение определенных мелодических контуров, придающих позитивную окраску коммуникативной ситуации.

Значительное влияние социальных факторов в приобретении языковых навыков традиционно отмечалось в теории социального обучения [26]. Однако новые исследования показывают, что социальное взаимодействие способствует не только лексическому развитию, но и овладению элементарными единицами языка. Исследования [27] продемонстрировали сильную корреляцию между различиями в социальном поведении детей во время занятий и степенью овладения фонетическими контрастами и даже словами.

Таким образом, способность приобретать языковые знания на раннем этапе развития представляет собой набор общих вычислительных навыков и экстраординарные способности к обучению. Поэтому на этапе оценки система дополняет зафиксированные периферической слуховой системой человека акустические (физические) признаки сигнала параметрами, соответствующими психическим представлениям о них. А именно: присваивает бинарный признак прототип/непрототип и эмоциональную окраску на основе сравнительной оценки продолжительности звучания материнских гласных и средней продолжительности звучания гласных, извлеченных из речи взрослых, а также используя сведения о том, что сигналы невысокой громкости привлекают внимание и оцениваются позитивно.

В финале, на третьем этапе, система размещает агента в признаковом пространстве. Размещение осуществляется с учетом концептуальных представлений магнитной теории родного языка [5]. Фонемы, размеченные как прототипические, конвергируют в признаковом пространстве, стимулы, соответствующие изученным категориям, обрабатываются быстрее. На наш взгляд, это поможет отразить тот факт, что раннее обучение может ограничить последующее [28], а также известное свидетельство о том, что, завершив обучение родному языку, взрослые, изучающие новый язык, не способны неявно поглощать статистические свойства фонетических единиц нового языка – внимание и когнитивные усилия определяются структурой освоенной категории. Повышенное внимание и ментальное усилие необходимы для обработки стимулов, резко несоответствующих существующим фонетическим категориям. θ -ритмы головного мозга указывают на устойчивое изменение в восприятии речи, обусловленное воздействием окружающей языковой среды [28]. Нейронная сеть мозга ребенка концентрируется на регистрации высокочастотных речевых событий, представленных фонетическими категориями, используемыми в данной среде.

3. АЛГОРИТМ ФОРМИРОВАНИЯ ПЕРЦЕПТИВНОГО ПРОСТРАНСТВА

Речевой поток фиксируется системой микрофонов. Затем выявляется его спектральный состав с помощью кратковременного преобразования Фурье [1] с использованием алгоритма Кули и Тьюки [1]. Далее методом YIN [29], аналогичным функции автокорреляции [1], оценивается частота основного тона, поскольку данный метод наиболее эффективен для выделения частоты основного тона монофонных музыкальных инструментов и речи. Состав гармоник определяется методом двухстороннего несоответствия [30]. Таким образом, на этапе предварительной обработки аудиосигнал преобразуется в следующий набор сигнатур:

$$\langle F_0, F_1, F_2, \Delta t \rangle,$$

где F_0 – частота основного тона,

F_1, F_2 – первая и вторая форманты,

Δt – продолжительность звучания исследуемой фонемы.

Полученный вектор признаков подается на вход мультиагентной рекурсивной когнитивной архитектуры [15], в которой разработчиками предварительно сформирован набор так называемых нейрофабрик – агентов специального типа, которые динамиче-

ски, по требованию создают агентов, определяют их тип и размещают в соответствующей области пространства мультиагентной системы. Итак, нейрофабрика создает агента, отвечающего за некоторую фонему. Для создания такого агента используется специальная программа, которая считывает геном агента – стартовый набор продукционных правил в базе знаний агента [15].

Процесс обучения предполагает подачу на вход системы полного набора обучающего материала, который подбирается с учетом особенностей речи, используемой при обращении к детям [5]. Как было подробно описано выше, одним из универсальных признаков речи, обращенной к детям, является продолжительность звучания гласных фонем, превышающая по продолжительности звучания гласные в речи, обращенной к взрослым, в три раза [5]. Мультиагентная система, исходя из сравнительной оценки продолжительности звучания анализируемой фонемы с существующими статистическими оценками долготы звучания «материнской» [5] фонемы, присваивает агенту дополнительный признак – эмоциональную окраску, а также идентифицирует его как прототип/непрототип. Поскольку вне данного контекста продолжительность звучания неинформативна, указанный параметр представляет собой бинарный признак. Эмоциональная окраска определяется в диапазоне от 0 до 1 с учетом экспериментальных данных о позитивном восприятии пролонгированных гласных и концентрации внимания на звуках с невысоким уровнем громкости.

Таким образом, на данном этапе базу знаний агента, характеризующего фонему, составляет следующий набор признаков:

$$\Phi_{vowel} = (\langle F_0, F_1, F_2, \Delta t \rangle, K_{letter}, \varepsilon, P),$$

где F_0 – частота основного тона,

$F_1 F_2$ – первая и вторая гармоники,

Δt – продолжительность звучания фонемы,

K_{letter} – классифицирующая контрактная связь с агентом-литерой,

$\varepsilon \in [0, 1]$ – эмоциональная окраска,

$P(0, 1)$ – прототип/непрототип.

Затем агент, реализуя поведение, детерминированное правилами, записанными в его собственном геноме, с целью отыскания класса, к которому принадлежит, обращается с вопросом к эксперту. На основе ответа эксперта заключается контракт между агентом первого уровня и агентом-литерой, характеризующей данный класс, то есть реализуется контролируемое машинное обучение [31].

Важно отметить, что признак прототип/непрототип и оценка эмоциональной окраски определяют не только пространственное положение агента в мультиагентном перцептивном пространстве, но и начальное время жизни агента в системе, которое затем управляется ступенчатой функцией памяти. В случае долгого периода бездействия агента происходит его исключение из системы. Настройки параметра продолжительности жизни агента позволяют исследовать проблемы эффективности овладения детьми новыми знаниями, а также проблемы запоминания [6].

Как показывает рисунок 2, результатом функционирования первого уровня разрабатываемой архитектуры являются создание агентов-актеров, регистрирующих акустические характеристики сигнала подобно слуховым рецепторам человека, и формирование множеств агентов, соответствующих каждой минимальной речевой единице языка, а также агентов, соответствующих фонеме. Предлагаемый алгоритм позволяет отследить механизм формирования слуховых паттернов человека.

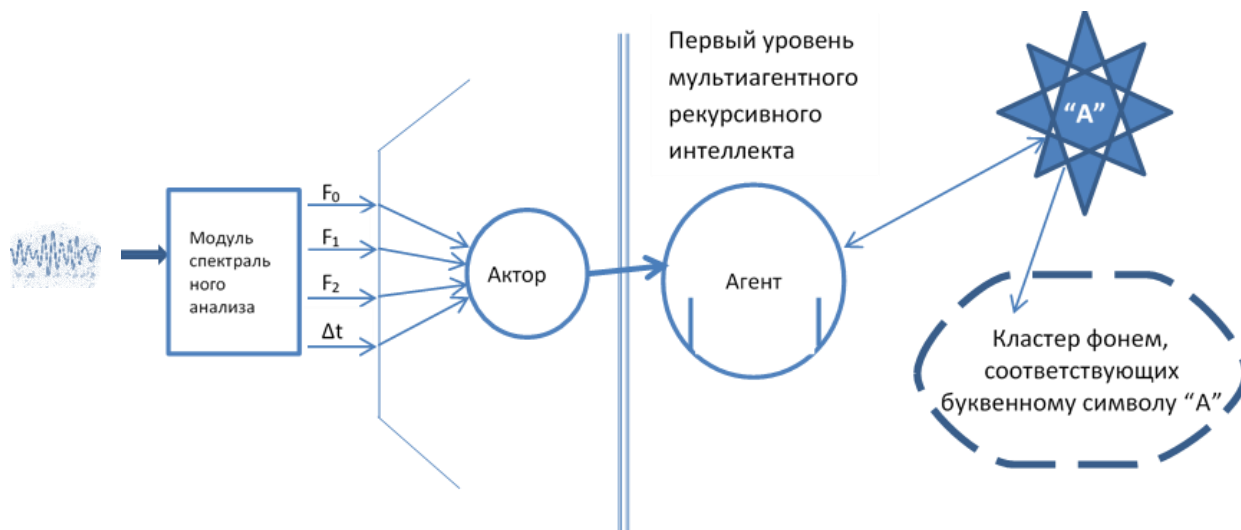


Рис. 2. Мультиагентная модель перцептивного пространства гласных фонем родного языка

На последующих уровнях для ускорения процесса распознавания планируется применять фонологические, грамматические ограничения [32]. Планируется также введение обратных связей для коррекции и уточнения результатов декодирования.

Таким образом, на основе анализа экспериментальных данных поведенческих исследований и модельных представлений о механизмах распознавания речи с точки зрения психолингвистического знания был разработан метод машинного обучения с имитацией формирования фонематического слуха человека.

4. ЗАКЛЮЧЕНИЕ

В данной работе предлагается модель раннего развития фонематической компетенции и алгоритм машинного обучения с имитацией механизма формирования слуховых нейронных групп человека. Предлагаемая модель позволит исследовать факторы, влияющие на продолжительность периода сенситивности, его причины и механизмы, создать диагностический инструментарий для описания периода сенситивности, изучить содержание этапов его открытия и закрытия, позволит создать речевые системы, устойчивые к многообразию акцентов.

ЛИТЕРАТУРА

1. *Jurafsky D., Martin J.* Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition. Boston, Prentice Hall, 2008. P. 1032
2. *Waibel A., Lee K.-F.* Readings in Speech Recognition. Berlington, Morgan Kaufman, 1990. P. 680.
3. *Stolcke Andreas & Droppo Jasha.* Comparing Human and Machine Errors in Conversational Speech Transcription. 137-141. 10.21437/Interspeech. 2017-1544.
4. *Saon George & Kurata Gakuto & Sercu Tom & Audhkhasi Kartik & Thomas Samuel & Dimitriadis Dimitrios & Cui Xiaodong & Ramabhadran Bhuvana & Picheny Michael & Lim Lynn-Li & Roomi Bergul & Hall Phil.* English Conversational Telephone Speech Recognition by Humans and Machines. 132-136. 10.21437/Interspeech.2017-405.
5. *Strange W.* Speech Perception and Linguistic Experience: Issues in Cross-Language Research. Baltimore, York Press, Baltimore, 1995. P. 492.

6. *Цейтлин С.Н.* Язык и ребенок: Лингвистика детской речи. М.: Гуманитарный издательский центр, 2000. С. 240.
7. *Chomsky N.A.* A Review of Skinner's Verbal Behavior. In: Jakobovits, L.A., Miron, M.S. (eds.) Readings in the Psychology of Language. Boston, Prentice-Hall, 1967. P. 636.
8. *Морозов В.П., Вартамян И.А., Галунов В.И.* Восприятие речи: вопросы функциональной асимметрии мозга. Ленинград: Наука, 1988. С. 135.
9. *Newell A.* Unified Theories of Cognition. Cambridge, Massachusetts: Harvard University Press, 1990. P. 576.
10. *Haikonen P.* The Cognitive Approach to Conscious Machines, Exeter, UK, imprint Academic, 2003. P. 300.
11. *Schunk D.H.* Learning Theories: An Educational Perspective, Boston, Pearson Merrill Prentice Hall, 2011. P. 576.
12. *Pinker S.* The Language Instinct: How the Mind Creates Language. New-York, Harper Perennial, 2007. P. 494.
13. *Kotseruba Iu, Tsotsos J.K.* A Review of 40 Years of Cognitive Architecture Research: Core Cognitive Abilities and Practical Applications. arxiv.org/abs/1610.08602
14. *Wooldridge M.* An Introduction to Multi-Agent Systems. Hoboken, New-Jersey, Wiley, 2009. P. 366.
15. *Нагоев З.В.* Интеллектика, или Мышление в живых и искусственных системах. Нальчик: Изд-во КБНЦ РАН, 2013. С. 232.
16. *De Mulder W., Bethard S., Moens M.-F.* A Survey on the Application of Recurrent Neural Networks to Statistical Language Modeling. Computer Speech and Language, 2015. № 30(1). P. 61-98.
17. *Deng L., Li X.* Machine Learning Paradigms for Speech Recognition: An Overview. IEEE Transactions on Audio, Speech, and Language Processing. 2013. №21(5). Pp. 1060-1089.
18. *Nagoev Z., Lyutikova L., Gurtueva I.* Model for Automatic Speech Recognition Using Multi-Agent Recursive Cognitive Architecture, Annual International Conference on Biologically Inspired Cognitive Architectures BICA, Prague, Czech Republic <http://doi.org/10.1016/j.procs.2018.11.089>
19. *Nagoev Z., Gurtueva I., Malyshev D., Sundukov Z.* Multi-agent Algorithm Imitating Formation of Phonemic Awareness. In: Samsonovich A. (eds) Biologically Inspired Cognitive Architectures 2019. BICA 2019. Advances in Intelligent Systems and Computing. Vol. 948. Springer, Cham https://doi.org/10.1007/978-3-030-25719-4_47
20. *Нагоев З.В., Гуртуева И.А.* Базовые элементы когнитивной модели механизма восприятия речи на основе мультиагентного рекурсивного интеллекта // Известия КБНЦ РАН. 2019. № 3(89). С. 3-14.
21. *Garnica O.* Some prosodic and paralinguistic features of speech to young children. In: Snow, C., Ferguson, Ch. (eds.): Talking to Children. Cambridge, Cambridge University Press, 1977. P. 63.
22. *Fernald A.* Four-month-old infants prefer to listen to motherese. Infant Behavior and Development, 1985. № 8. Pp. 181-95.
23. *Fernald A., Kuhl P.* Acoustic determinants of infant preference for Motherese Speech. Infant Behavior and Development, 1987. № 10. Pp. 279-93.
24. *Moerk E.L.* Principles of interaction in language learning. Merrill-Palmer Quaterly, 1972. № 18. Pp. 229-257.
25. *Pye C.* Quiché Mayan speech to children // Journal of Child Language. 1986. № 13(1). Pp. 85-100.

26. *Выготский Л.С.* Мышление и речь. Санкт Петербург: Питер, 2019. С. 432.
27. *Conboy B.T., Kuhl P.K.* Impact of second-language experience in infancy: Brain measures of first- and second-language speech perception. *Developmental Science*, 2011. № 14. Pp. 242-248. <https://doi.org/10.1111/j.1467-7687.2010.00973.x>
28. *Doupe A.J., Kuhl P.K.* Birdsong and human speech: Common themes and mechanisms. In: *Zeigler, H.P., Marler, P. (eds.): Neuroscience of birdsong.* Cambridge University Press, 2008. Pp. 5-31.
29. *De Cheveigne A., Kawahara H.* YIN, a fundamental frequency estimator for speech and music // *The Journal of the Acoustical Society of America*, 2002. № 111(4). Pp. 1917-1930.
30. *Maher R.C., Beauchamp J.W.* Fundamental frequency estimation of musical signals using a two-way mismatch procedure // *The Journal of the Acoustical Society of America*. 1994. №95, 2254 <https://doi.org/10.1121/1.408685>
31. *Coates A., Ng A.Y.* Learning Feature Representations with K-Means. In: *Montavon G., Orr G.B., Müller KR. (eds) Neural Networks: Tricks of the Trade. Lecture Notes in Computer Science*, 2012. Vol 7700. Springer, Berlin, Heidelberg.
32. *Зиндер Л.П.* Общая фонетика. М.: Высшая школа, 1979. С. 312.

REFERENCES

1. *Jurafsky D., Martin J.* *Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition.* Boston, Prentice Hall, 2008. P. 1032
2. *Waibel A., Lee K.-F.* *Readings in Speech Recognition.* Berlington, Morgan Kaufman, 1990. P. 680.
3. *Stolcke Andreas & Droppo Jasha.* Comparing Human and Machine Errors in Conversational Speech Transcription. 137-141. 10.21437/Interspeech.2017-1544.
4. *Saon George & Kurata Gakuto & Seru Tom & Audhkhasi Kartik & Thomas Samuel & Dimitriadis Dimitrios & Cui Xiaodong & Ramabhadran Bhuvana & Picheny Michael & Lim Lynn-Li & Roomi Bergul & Hall Phil.* English Conversational Telephone Speech Recognition by Humans and Machines. 132-136. 10.21437/Interspeech.2017-405.
5. *Strange W.* *Speech Perception and Linguistic Experience: Issues in Cross-Language Research.* Baltimore, York Press, Baltimore, 1995. P. 492.
6. *Tseitlin S.N.* *Yazyk i rebenok: Lingvistika detskoy rechi* [A Child and a Language: Child Speech Linguistics]. Humanitarian Publishing Center VLADOS. Moscow, 2000.
7. *Chomsky N.A.* A Review of Skinner's Verbal Behavior. In: *Jakobovits, L.A., Miron, M.S. (eds.) Readings in the Psychology of Language.* Boston, Prentice-Hall, 1967. P. 636.
8. *Morozov V.P., Vartanyan I.A., Galunov V.I.* *Vospriyatiye rechi: voprosy funktsional'noy asimmetrii mozga* [Speech Perception: Problems of Functional Brain Asymmetry] Science, St. Petersburg, 1988.
9. *Newell A.* *Unified Theories of Cognition.* Cambridge, Massachusetts: Harvard University Press, 1990. P. 576.
10. *Haikonen P.* *The Cognitive Approach to Conscious Machines,* Exeter, UK, imprint Academic, 2003. P. 300.
11. *Schunk D.H.* *Learning Theories: An Educational Perspective,* Boston, Pearson Merrill Prentice Hall, 2011. P. 576.
12. *Pinker S.* *The Language Instinct: How the Mind Creates Language.* New-York, Harper Perennial, 2007. P. 494.
13. *Kotseruba Iu, Tsotsos J.K.* A Review of 40 Years of Cognitive Architecture Research: Core Cognitive Abilities and Practical Applications. arxiv.org/abs/1610.08602

14. Wooldridge M. An Introduction to Multi-Agent Systems. Hoboken, New-Jersey, Wiley, 2009. P. 366.
15. Nagoev Z.V. *Intellektika, ili Myshleniye v zhyvykh i iskusstvennykh sistemakh* [Intellectics, or thinking in living and artificial systems]. Publishing House KBSC RAS. Nalchik, 2013.
16. De Mulder W., Bethard S., Moens M.-F. A Survey on the Application of Recurrent Neural Networks to Statistical Language Modeling. *Computer Speech and Language*, 2015. № 30(1). P. 61-98.
17. Deng L., Li X. Machine Learning Paradigms for Speech Recognition: An Overview. *IEEE Transactions on Audio, Speech, and Language Processing*. 2013. №21(5). Pp. 1060-1089.
18. Nagoev Z., Lyutikova L., Gurtueva I. Model for Automatic Speech Recognition Using Multi-Agent Recursive Cognitive Architecture, Annual International Conference on Biologically Inspired Cognitive Architectures BICA, Prague, Czech Republic <http://doi.org/10.1016/j.procs.2018.11.089>
19. Nagoev Z., Gurtueva I., Malyshev D., Sundukov Z. Multi-agent Algorithm Imitating Formation of Phonemic Awareness. In: Samsonovich A. (eds) *Biologically Inspired Cognitive Architectures 2019. BICA 2019. Advances in Intelligent Systems and Computing*. Vol. 948. Springer, Cham https://doi.org/10.1007/978-3-030-25719-4_47
20. Nagoev Z.V., Gurtueva I.A. *Bazovye elementy kognitivnoi modeli mehanizma vospriyatiya rechi na osnove multiagentnogo rekursivnogo intellekta* [Fundamental Elements for Cognitive Model of Speech Perception Mechanism Based on Multiagent Recursive Intellect] // *News of Kabardino-Balkarian Scientific Center of RAS*, 2019. № 3(89). Pp. 3-14.
21. Garnica O. Some prosodic and paralinguistic features of speech to young children. In: Snow, C., Ferguson, Ch. (eds.): *Talking to Children*. Cambridge, Cambridge University Press, 1977. P. 63.
22. Fernald A. Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 1985. №8. Pp.181-95.
23. Fernald A., Kuhl P. Acoustic determinants of infant preference for Motherese Speech. *Infant Behavior and Development*, 1987. №10. Pp. 279-93.
24. Moerk E.L. Principles of interaction in language learning. *Merril-Palmer Quaterly*, 1972. № 18. Pp. 229-257.
25. Pye C. Quiché Mayan speech to children // *Journal of Child Language*. 1986. № 13(1). Pp. 85-100.
26. Vygotsky L.S. *Myshlenie I Rech'* [Thinking and Speech]. St-Petersburg: Piter, 2019.
27. Conboy B.T., Kuhl P.K. Impact of second-language experience in infancy: Brain measures of first- and second-language speech perception. *Developmental Science*, 2011. № 14. Pp. 242-248. <https://doi.org/10.1111/j.1467-7687.2010.00973.x>
28. Doupe A.J., Kuhl P.K. Birdsong and human speech: Common themes and mechanisms. In: Zeigler, H.P., Marler, P. (eds.): *Neuroscience of birdsong*. Cambridge University Press, 2008. Pp. 5-31.
29. De Cheveigne A., Kawahara H. YIN, a fundamental frequency estimator for speech and music // *The Journal of the Acoustical Society of America*, 2002. № 111(4). Pp. 1917-1930.
30. Maher R.C., Beauchamp J.W. Fundamental frequency estimation of musical signals using a two-way mismatch procedure // *The Journal of the Acoustical Society of America*. 1994. № 95, 2254 <https://doi.org/10.1121/1.408685>
31. Coates A., Ng A.Y. Learning Feature Representations with K-Means. In: Montavon G., Orr G.B., Müller KR. (eds) *Neural Networks: Tricks of the Trade. Lecture Notes in Computer Science*, 2012. Vol. 7700. Springer, Berlin, Heidelberg.
32. Zinder L.R. *Obshaya Fonetika* [The General Phonetics]. The Higher School, Moscow, 1979.

MULTIAGENT MODEL OF PERCEPTUAL SPACE FORMATION IN THE PROCESS OF MASTERING LINGUISTIC COMPETENCE*

Z.V. NAGOEV¹, I.A. GURTUEVA², K.Ch. BZHIKHATLOV¹

¹Federal public budgetary scientific establishment «Federal scientific center
«Kabardin-Balkar Scientific Center of the Russian Academy of Sciences»
360002, KBR, Nalchik, 2, Balkarova street
E-mail: kbncran@mail.ru

² Institute of Computer Science and Problems of Regional Management –
branch of FSBSE “Federal scientific center
“Kabardin-Balkar Scientific Center of the Russian Academy of Sciences”
360000, KBR, Nalchik, I. Armand street, 37-a.
E-mail: iipru@rambler.ru

The model of the early development of language competencies proposed in this paper, takes into account the social factors effects. It is a simulation model of phonemic imprinting. The model describes the process of perceiving audio stimuli as their mapping into classes of elementary language units. The machine learning algorithm was developed using the results of the study of speech addressed to children. Our model will allow to explore the features of phonetic perception, the cognitive mechanisms that underlie language development, highlight the main factors affecting the duration of the plasticity period. The proposed model gives possibilities to build perceptual maps, design diagnostic tools to describe and study the sensitive period. The model can also be used to create speech systems that are resistant to various influences and effective when used in conditions of high noise.

Keywords: *multiagent systems, artificial intelligence, artificial neuron networks, speech recognition, plasticity period, motherese.*

Работа поступила 08.06.2020 г.

* This work was supported by grants of the Russian Foundation for Basic Research No. 18-01-00658, 19-01-00648