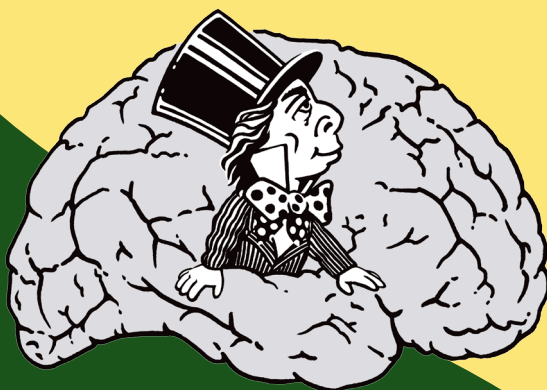


КОГНИТИВНАЯ НАУКА

В МОСКВЕ



НОВЫЕ ИССЛЕДОВАНИЯ

МАТЕРИАЛЫ
КОНФЕРЕНЦИИ
2023

Под ред. Е.В. Печенковой, М.В. Фаликман, А.Я. Койфман

УДК 159.9
ББК 88.25
К57

Когнитивная наука в Москве: новые исследования. Материалы конференции 21 – 22 июня 2023 г. Под ред. Е.В. Печенковой, М.В. Фаликман, А.Я. Койфман. – М.: ООО «Буки Веди», Московский институт психоанализа. 2023 г. – 604 стр.

© Авторы статей, 2023

ISBN 978-5-4465-3880-5

УДК 159.9
ББК 88.25

ISBN 978-5-4465-3880-5

© Авторы статей, 2023

РАЗРАБОТКА И АПРОБАЦИЯ ИНДЕКСА ПРОДУКТИВНОСТИ СИНТАКСИСА ПРИ ОСВОЕНИИ ИМЕННЫХ ГРУПП РУССКОГОВОРЯЩИМИ ДЕТЬМИ

М. Д. Дьячкова* (1), И. А. Секерина (2), С. В. Дорофеева (1)
masha.dyachcova@yandex.ru

1 – Центр языка и мозга, Национальный исследовательский университет
«Высшая школа экономики», Москва; 2 – Городской университет
Нью-Йорка, США

Аннотация. В рамках настоящего проекта было положено начало разработке Индекса продуктивности синтаксиса для именных групп – нового метода для оценки грамматической сложности образцов спонтанной речи русскоязычных детей. Такой инструмент был предложен в 1990 году для англоязычных детей, однако на базе русского языка подобные методы отсутствуют, речевой профиль маленьких детей чаще оценивается субъективно. Опираясь на оригинальный Индекс, а также учитывая синтаксические и морфологические особенности русского языка, мы разработали список из 11 параметров, оценивающих грамматическую сложность детской речи, связанную с именными группами. Параметры были апробированы на лонгитюдных записях трех русскоязычных детей, собранных в рамках проекта CHILDES. Были выбраны записи в возрасте 20, 22, 24, 26, 28, 30 и 32 месяцев. Первое применение инструмента показало значимое влияние возраста на значение показателей Индекса.

Ключевые слова: детская речь, грамматическая сложность, CHILDES, спонтанная речь, развитие речи

Введение

Индекс продуктивности синтаксиса (ИПС) – это метод количественной оценки грамматической сложности образцов спонтанной речи маленьких детей. Метод был предложен в 1990 году Холлис Скарборо (Scarborough, 1990) для оценки речевых навыков англоязычных детей. Особенность и преимущество этого метода состоит в том, что ИПС позволяет оценивать спонтанную детскую речь с самого раннего возраста, когда применение стандартизированных тестов на порождение и восприятие речи может быть проблематичным. В настоящее время для русского языка подобные инструменты для оценки речевого профиля детей отсутствуют. На данный момент для этих целей используются субъективный метод наблюдения, речевые карты, опросники. Также существует тест для оценки речевых навыков русскоязычных детей «КОРАБЛИК» (Клиническая оценка развития базовых лингвистических компетенций, Гомозова и др., 2021), однако пока он используется крайне ограниченным кругом исследователей.

При этом создание ИПС позволило бы определять речевой профиль ребенка без личной встречи исследователя (педагога, логопеда, нейропсихолога) и ребенка, только по записям речи. Кроме того, предполагается, что ИПС даст возможность оценивать речевые навыки детей до трех лет.

В нашем исследовании мы предпринимаем попытку создать аналогичный Индексу продуктивности синтаксиса инструмент для оценки речевых навыков русскоязычных детей, начав с разработки Индекса для именных групп. Исследование состоит из двух основных частей: собственно разработки параметров Индекса и проверки и апробации Индекса на данных корпусах детской речи.

Разработка параметров. Был разработан список параметров Индекса для русского языка. При его разработке мы отталкивались от оригинального Индекса, предложенного Х. Скарборо и содержащего 12 параметров для именных групп. Кроме того, мы постарались учесть синтаксические и морфологические особенности русского языка с опорой на «Русскую грамматику» (Шведова, 1980). Список параметров и пояснения к ним приведены в табл. 1.

Метод

Материалы. Проверка и апробация Индекса были проведены на базе лонгитюдных корпусов детской речи, собранных в рамках проекта CHILDES (MacWhinney, 2000). На них запечатлено ежедневное взаимодействие детей со взрослыми. Исследуемые нами корпуса размечены в латинице в программе CLAN по определенным правилам, для удобства дальнейшей работы с данными все реплики были транслитерированы в кириллицу. Индекс был апробирован на материалах лонгитюдных записей трех русскоязычных детей-монолингвов (две девочки и мальчик) без неврологических нарушений. Были использованы записи в возрасте 20, 22, 24, 26, 28, 30, 32 месяцев. Такой возраст был выбран вслед за (Xathos et al., 2011).

Процедура. Для отбора детских реплик из корпуса, дальнейшего их морфологического анализа и оценки по параметрам был написан скрипт на языке Python. Для оценивания образца речи каждого ребенка в каждом из выбранных возрастов мы отобрали 100 первых его реплик в корпусе в этом возрасте. Для морфологического разбора и оценивания выбранных реплик мы использовали морфологический анализатор Mystem, определялись части речи и морфологические характеристики. Реплики детей и их морфологический анализ записывались в таблицу. Ошибки анализатора и случаи омонимии исправлялись и разрешались вручную. По каждому параметру каждая подвыборка из 100 реплик (подкорпус) ребенка могла получить от 0 до 2 баллов (вслед за процедурой в (Scarborough, 1990)). Автоматически оценивались все реплики, выполнялся поиск слов и сочетаний, удовлетворяющих условию параметра. За подходящий пример начислялся 1 балл. После того как внутри подкорпуса по конкретному параметру набиралось 2 балла, поиск подходящих примеров по этому параметру завершался, и можно было переходить к следующему. В результате была получена таблица, в которой для каждого параметра по каждому подкорпусу был указан балл от 0 до 2, а также конкретные примеры из подкорпуса (слова и сочетания слов), за которые эти баллы были начислены.

Результаты и обсуждение

Для оценки показателей Индекса был проведен однофакторный дисперсионный анализ с помощью языка программирования R и пакета car. Было вы-

Таблица 1. Параметры ИПС-ИГ для русского языка

	Название параметра	Описание
Параметры, унаследованные из (Scarborough, 1990)		
1	Существительное	Ребенок употребляет в своей речи существительные
2	Местоимение	Ребенок употребляет в речи местоимения-существительные: личные местоимения, возвратные местоимения, вопросительные, неопределенные или отрицательные местоимения
3	Модификаторы	Ребенок употребляет в речи полные прилагательные, а также местоимения-прилагательные или счетно-порядковые прилагательные
4	Двусловные ИГ с модификатором	Ребенок использует двусловные словосочетания, слова в которых связаны согласованием. Зависимыми в таких словосочетаниях выступают слова, изменяющиеся по родам, числам и падежам
5	Трехсловные ИГ	Ребенок использует сочетания, в которых с существительным согласуются два или более модификатора
6	Число	Ребенок использует существительные или прилагательные во множественном числе
Новые параметры, отсутствующие в оригинальном Индексе		
7	Двусловные ИГ с управлением	Ребенок использует двусловные словосочетания, слова в которых связаны управлением. Главным словом в таких словосочетаниях может выступать существительное, прилагательное или количественное числительное
8	Двусловные ИГ с примыканием	Ребенок использует трехсловные словосочетания, слова в которых связаны примыканием. Главным словом в таких словосочетаниях может выступать существительное или прилагательное
9	Предложная группа	Ребенок использует сочетания предлога и существительного или местоимения
10	Падеж	Ребенок использует существительные, местоимения-существительные или прилагательные в косвенных падежах
11	Редкие формы	Ребенок использует существительные в форме вторичных падежей (второй родительный, второй предложный), образует косвенные формы разносклоняемых существительных

явлено значимое влияние *возраста* на значения по Индексу, $F(6,20) = 3.976$, $p = .0156$. Кроме того, было обнаружено, что не все параметры демонстрируют постепенное увеличение значения от возраста к возрасту. Средние баллы по каждому параметру представлены в табл. 2.

Так, можно заметить, что параметры «Существительное» (1 в таблице), «Двусловные ИГ с примыканием» (8), «Редкие формы» (11) не демонстрируют постепенного увеличения значения внутри выбранного временного промежутка. Что касается параметра «Существительное», отсутствие различий по возрастам в этом параметре может объясняться тем, что номинации – черта, появляющаяся в речи на самых ранних этапах (Воейкова, 2015). Отсутствие

различий в «Редких формах» может быть связано с тем, что формы второго родительного и второго предложного вообще довольно редки в русском языке (Копотев, 2008).

Таблица 2. Средние баллы по параметрам Индекса продуктивности синтаксиса для именных групп

Возраст (мес.)	Средний балл по каждому параметру											Итого
	1	2	3	4	5	6	7	8	9	10	11	
20	2.00	0.33	0.33	0.00	0.00	1.00	0.00	0.00	0.33	0.00	0.00	0.36
22	2.00	1.00	0.67	0.00	0.00	1.00	0.00	0.00	0.67	0.00	0.00	0.48
24	2.00	0.67	1.33	0.00	0.00	0.67	0.00	0.00	0.67	0.33	0.33	0.55
26	2.00	1.67	1.33	1.00	0.33	1.33	0.33	0.00	1.33	0.67	0.33	0.94
28	2.00	2.00	1.67	1.00	0.00	1.33	0.33	0.00	1.33	1.33	0.00	1.00
30	2.00	2.00	2.00	1.33	0.67	2.00	0.33	0.00	2.00	2.00	0.00	1.30
32	2.00	2.00	2.00	1.67	1.33	2.00	0.67	0.00	2.00	2.00	0.33	1.45

Первая апробация разработанного нами инструмента на небольшом объеме данных может говорить об Индексе как о надежном способе оценки грамматической сложности детской речи. Однако для более аккуратных выводов необходима проверка Индекса на большем количестве данных, а также более подробный анализ различий показателей соседних возрастов.

Литература

Войкова М.Д. Становление имени: Ранние этапы усвоения детьми именной морфологии русского языка. М.: Фонд «Развития фундаментальных лингвистических исследований», 2015.

Гомозова М.А., Арутюнян В.Г., Лопухина А.А., Драгой О.В. Инструмент для комплексного обследования речевых навыков КОРАБЛИК и опыт его применения в группе младших школьников с РАС // Аутизм и нарушения развития. 2021. Т. 19. № 4. С. 24 – 31. <https://doi.org/10.17759/autidd.2021190403>

Копотев М.В. К построению частотной грамматики русского языка: падежная система по корпусным данным // Инструментарий русистики: корпусные подходы / Под ред. А. Мустайоки, М.В. Копотева, Л.А. Бирюлина, Е.Ю. Протасовой. Helsinki: 2008. Т. Slavica Helsingiensia. 34. С. 136 – 151.

Шведова Н.Ю. Русская грамматика. Т. 1. М.: Наука, 1980.

MacWhinney B. The CHILDES project: The database. London: Psychology Press, 2000.

Scarborough H.S. Index of productive syntax // Applied Psycholinguistics. 1990. Vol. 11. No. 1. P. 1 – 22. <https://doi.org/10.1017/s0142716400008262>

Xanthos A., Laaha S., Gillis S., Stephany U., Aksu-Koç A., Christofidou A., Gagarina N., Hrzica G., Ketz F.N., Kilani-Schoch M., Korecky-Kröll K., Kovačević M., Laalo K., Palmović M., Pfeiler B., Voeikova M.D., Dressler W.U. On the role of morphological richness in the early development of noun and verb inflection // First Language. 2011. Vol. 31. No. 4. P. 461 – 479. <https://doi.org/10.1177/0142723711409976>

DEVELOPMENT AND APPROVAL OF THE INDEX OF PRODUCTIVE SYNTAX IN NOUN PHRASE ACQUISITION BY RUSSIAN-SPEAKING CHILDREN

M. D. Dyachkova* (1), I. A. Sekerina (2), S. V. Dorofeeva (1)

masha.dyachkova@yandex.ru

1 – Center of Language and Brain, HSE University, Moscow;

2 – The City University of New York, USA

Abstract. A new method has been developed for assessing the grammatical complexity of samples of spontaneous speech of Russian-speaking children: the Index of Productive Syntax. Such a tool was proposed in 1990 for English-speaking children, but there have been no such methods for the Russian language; the language development of young children is more often evaluated subjectively. Based on the original Index, and also taking into account the syntactic and morphological features of the Russian language, we suggested a list of 11 parameters that assess the grammatical complexity of children's speech associated with noun phrases. The parameters were tested on samples from longitudinal recordings collected from three Russian-speaking children as part of the CHILDES project. Entries were selected at 20, 22, 24, 26, 28, 30 and 32 months of age. The first use of the tool showed a significant influence of age on the value of the Index.

Keywords: child language, grammatical complexity, CHILDES, spontaneous speech, language development