

## УСВОЕНИЕ ИМЕННЫХ КЛАССОВ В ЦЕЗСКОМ ЯЗЫКЕ\*

© 2015 г.

Магомед Ибрагимович Магомедов<sup>а</sup>,  
Арсен Курбанович Абдулаев<sup>а</sup>,  
Анни Гальярди<sup>б</sup>,  
Мария Семеновна Полинская<sup>б, @</sup>

<sup>а</sup> Институт языка, литературы и искусства им. Г. Цадасы Дагестанского научного центра РАН, Махачкала, 367025, Россия; <sup>б</sup> Гарвардский университет, Кембридж, 02138, США; <sup>@</sup> polinsky@fas.harvard.edu

Данная статья преследует две взаимосвязанные цели. Первая, методологическая, — представление и обсуждение доступной методики по изучению детской речи; такая методика может быть применена к самым разным языкам, в том числе языкам малых народов или бесписьменным. Методика включает в себя создание корпуса детской речи и речи, обращенной к детям (в нашем исследовании около 10 часов живой речи), анализ этого корпуса и дальнейшее экспериментальное исследование моделей и закономерностей, которые попадают в поле зрения исследователя корпусного материала. Наша вторая цель состоит в том, чтобы, пользуясь этой методикой, рассмотреть систему именных классов (родов) в цезском языке нахско-дагестанской семьи и показать, как дети усваивают подобную систему при освоении родного языка. Главным вопросом является набор и соотношение предсказательных признаков, которые позволяют детям и взрослым приписывать конкретные слова к тому или иному именному классу. Данные нашего экспериментального исследования показывают, что дети и взрослые опираются на разные предсказательные признаки. В частности, для детей — носителей языка — формальные признаки существительного (например, начальный или конечный звук слова) играют более важную роль, чем семантика, в то время как для взрослых семантика имеет большее значение.

**Ключевые слова:** детская речь, именной класс (род), корпусная лингвистика, нахско-дагестанские языки, семантические признаки, фонологические признаки, цезский язык, экспериментальная методология

## THE ACQUISITION OF NOUN CLASSES IN TSEZ

Magomed I. Magomedov<sup>а</sup>, Arsen K. Abdulaev<sup>а</sup>,  
Annie Gagliardi<sup>б</sup>, Maria Polinsky<sup>б, @</sup>

<sup>а</sup> The Gamzat Tsadasa Institute of Language, Literature and Arts, Daghestan Academic Centre, Russian Academy of Sciences, Makhachkala, 367025, Russia; <sup>б</sup> Harvard University, Cambridge, 02138, USA; <sup>@</sup> polinsky@fas.harvard.edu

This paper follows two related goals. First, it introduces a simple methodology that can be used to investigate first language acquisition in lesser-studied languages, including those that may be endangered. The proposed methodology consists of three steps, namely, (i) corpus building, (ii) data mining based on that corpus, and (iii) subsequent experimental studies motivated by patterns observed in the corpus. The second goal of this paper is to apply the proposed methodology to the investigation of noun-class acquisition in the Nakh-Dagestanian language Tsez. Based on the proposed methodology, we have constructed a new corpus of child-directed speech and child language (about ten hours of speech), which we analyze to determine the role of predictive cues that

\* Работа выполнена при поддержке Научного фонда США (NSF Graduate Research Fellowship Program; NSF IGERT 0801465; NSF 1144223) и стипендии по документации исчезающих языков (DEL).

Авторы выражают искреннюю благодарность Патимат Абдулаевой и Хайрат Гамзатовой за помощь в экспериментальном исследовании, а также следующим коллегам за обсуждение этой работы на разных ее этапах: М. Е. Алексееву, А. Е. Кибрику, Б. Комри, Дж. Лидзу, Я. Г. Тестельцу и Д. А. Эршлеру.

help the learner assign a given noun to a noun class. Since Tsez does not overtly mark agreement on all its verbs or adjectives, we also use the corpus data to assess the extent of overt noun-class agreement. We conclude that a variety of semantic and phonological (formal) cues allow Tsez speakers to determine the noun class of a given nominal. We then present and analyze an elicited-production experiment that uncovers asymmetries in the classification of nouns with predictive features in the corpus and by children and adults. We show that children are biased to use phonological information over semantic information, despite a statistical asymmetry in the other direction.

**Keywords:** corpus building, experimental methodologies, language acquisition, Nakh-Dagestanian languages, noun class (gender), phonological properties, semantic properties, Tsez

## 1. Введение

Данная статья преследует две цели: общелингвистическую и методологическую. С общелингвистической точки зрения наша задача состоит в том, чтобы рассмотреть систему именных классов (родов) в цезском языке и показать, как дети усваивают подобную систему при освоении родного языка. Усвоение рода детьми изучалось на материале индоевропейских языков (ср. [Гвоздев 1961; Rodina 1999] для русского, [Karmiloff-Smith 1979] для французского, [MacWhinney 1978; Mills 1985; 1986; Szagun et al. 2007] для немецкого), однако в дагестанских языках такое исследование проводится впервые. Новизна данного исследования определяет вторую цель нашей работы: продемонстрировать, что изучение детской речи в малых языках в принципе возможно, и предложить конкретные методы подобного изучения. Отчасти данная статья перекликается с англоязычной работой [Gagliardi, Lidz 2014], в которой изложен материал того же эксперимента, проведенного в Дагестане с детьми, изучающими цезский язык как родной. Однако статья [Gagliardi, Lidz 2014] не преследует общеметодологических целей; между тем мы надеемся, что данная публикация послужит стимулом для дальнейшей работы по изучению того, как дети усваивают бесписьменные языки. Если нам удастся представить читателю то, что по-английски называется *proof of concept* (проверка концепции) относительно экспериментальной работы с малыми языками, это будет шагом вперед в изучении и сохранении этих языков.

Информацию о принадлежности существительного к тому или иному согласовательному классу можно почерпнуть из двух источников: во-первых, из согласовательных показателей и, во-вторых, из информации, содержащейся непосредственно в существительном — в его семантике и фонологическом облике. Наше внимание направлено именно на информацию, заложенную в существительном, и в связи с этим нас интересует вопрос о том, на что обращают внимание дети и взрослые при классификации новых для них существительных. Забегая вперед, заметим: несмотря на то, что со статистической точки зрения семантическая информация является более надежным признаком класса, чем фонологическая, дети в отличие от взрослых, по-видимому, более предрасположены к тому, чтобы ориентироваться в первую очередь на фонологические признаки. Этот результат перекликается с выводами исследователей, которые изучали освоение грамматического рода в таких языках, как французский или немецкий.

## 2. Система именных классов в цезском языке

Цезский язык принадлежит к цезской подгруппе аваро-андо-цезской группы нахско-дагестанской семьи. По данным Всероссийской переписи населения 2012 г., на нем говорят около 12 тысяч человек, живущих в Цунтинском районе Республики Дагестан, а также переселенцы в равнинные районы этой республики.

В цезском языке в единственном числе существительные делятся на четыре класса, которые во множественном числе сливаются в два, см. [Бокарёв 1959; Имнайшвили 1963; Comrie, Polinsky 1999].

Класс, к которому принадлежит существительное, можно определить по морфологическим показателям, выражающим согласование в начинающихся на гласный глаголах, а также некоторых прилагательных и наречиях.

Согласовательный префикс в прилагательных фонологически пуст для существительных I класса, согласование с существительными II класса выражается с помощью префикса *y-*, с существительными III класса — с помощью префикса *b-*, а IV класса — *r-*. Согласовательные префиксы во множественном числе, а также некоторые формы личных и указательных местоимений зависят от именного класса, но из-за заметного синкретизма эти показатели служат куда менее надежными признаками именного класса, к которому принадлежит существительное — контролер согласования. Кроме того, класс существительного можно до некоторой степени предсказать по его семантике и фонологическому облику. Мы, соответственно, будем впредь говорить о внешних и внутренних признаках, позволяющих носителю языка определять класс данного существительного. Под внешними признаками мы подразумеваем согласовательные показатели, а под внутренними — информацию, содержащуюся в самом существительном.

Ниже приводятся иллюстративные примеры, показывающие классное согласование в цезском языке (мы приводим только примеры в единственном числе):

- (1) Ø-igu      uži  
I-хороший    мальчик.АБС
- (2) y-igu      kid  
II-хорошая    девочка.АБС
- (3) b-igu      k'et'u  
III-хорошая    кошка.АБС
- (4) r-igu      y<sup>h</sup>utku  
IV-хороший    дом.АБС

Глагол цезского языка всегда согласуется с абсолютивом, т. е. с подлежащим в непереходном предложении и с дополнением в переходном. Следующие примеры показывают, как работает глагольное согласование (примеры серии «а» содержат непереходный глагол, серии «б» — переходный):

- (5) а. Uži            Ø-ay-s.  
мальчик.АБС I-приходить-ПРОШ  
'Мальчик пришел'.
- б. Eñiy-ä      uži            Ø-ayir-si.  
мать-ЭРГ мальчик.АБС I-приводить-ПРОШ  
'Мать привела мальчика'.
- (6) а. Kid            y-ay-s.  
Девочка.АБС II-приходить-ПРОШ  
'Девочка пришла'.
- б. Eñiy-ä      kid            y-ayir-si.  
мать-ЭРГ девочка.АБС II-приводить-ПРОШ  
'Мать привела девочку'.
- (7) а. K'et'u        b-ay-s.  
кошка.АБС III-приходить-ПРОШ  
'Кошка пришла'.
- б. Eñiy-ä      k'et'u        b-ayir-si.  
мать-ЭРГ кошка.АБС III-приводить-ПРОШ  
'Мать привела кошку'.
- (8) а. Łaci        r-ay-s.  
ветер.АБС IV-пришел  
'Ветер налетел (пришел)'.
- б. Allah-ä      łaci        r-ayir-si.  
Бог-ЭРГ ветер.АБС IV-приводить-ПРОШ  
'Бог прислал ветер'.

В таблице 1 представлен состав именных классов, основанный на традиционных описаниях цезского языка. Класс I наиболее четко выделяется из остальных — к нему относятся все мужчины (а также некоторые сверхъестественные существа) и только они, и, таким образом, на отнесение существительных к этому классу имеются наибольшие семантические ограничения. Класс II включает имена, обозначающие лиц женского пола, но в отличие от I класса он состоит главным образом из неодушевленных и абстрактных существительных. Класс III самый обширный — к нему относятся названия всех живых существ, кроме людей, а также большое число неодушевленных, в том числе абстрактных существительных. Класс IV включает многие неодушевленные существительные, среди них абстрактные, в том числе образованные при помощи абстрактного суффикса *-i*.

Таблица 1

Именные классы в цезском языке

	Класс I	Класс II	Класс III	Класс IV
Доля существительных, входящих в класс, от общего числа существительных	13 %	13 %	40 %	34 %
Состав класса	Наименования мужчин и некоторых сверхъестественных существ	Наименования женщин; неодушевленные существительные	Все прочие одушевленные существительные; неодушевленные существительные	Неодушевленные существительные

Проценты подсчитаны по существительным, представленным в словаре [Халилов 1999]. При подсчете мы не учитывали слова, которые могут относиться к нескольким классам в зависимости от пола денотата, например ‘учитель’.

[Plaster et al. 2013] подробно рассматривают вопрос о том, какого рода информация позволяет соотносить цезские существительные с именными классами. Авторы этой работы разместили все существительные из словаря [Халилов 1999] в соответствии с потенциальными признаками — индикаторами класса и установили, что можно сформулировать правила, апеллирующие ко внутренним признакам существительных. Эти правила позволяют классифицировать около 70 % цезских существительных.

В дополнение к материалам, использованным в работе [Plaster et al. 2013], мы собрали и обработали корпус детской речи и речи взрослых, обращенной к детям [Gagliardi 2012]. На материале нашего корпуса мы уточнили состав признаков — индикаторов класса, предложенных в [Plaster et al. 2013]. Они включают в себя простые семантические характеристики (мужчина, женщина, животное, все прочие сущности) и столь же простые, легко доступные формальные (фонологические) признаки (исходный звук слова, конечный звук слова). Мы сознательно использовали именно эти признаки, так как они позволяют увидеть, на какие семантические или фонологические свойства ориентируются носители языка при отнесении существительного к определенному классу.

Полный набор признаков приведен в таблице 2, где также показана вероятность, с которой данное существительное может быть членом того или иного класса при наличии у этого существительного соответствующего признака — индикатора класса<sup>1</sup>. Например, если слово начинается со звука /b/, то вероятность его принадлежности к классу III составляет 51 %, что обозначено в таблице как 0,51. Вероятность этого типа вводится с помощью записи  $p(C_i \dots | \text{признак } X)$ . Второй тип условной вероятности, показанный в таблице 2, — это

<sup>1</sup> Имеется в виду условная вероятность, т. е. возможность принадлежности к определенному классу при наличии у формы некоторого признака.

степень вероятности, с которой можно ожидать, что имя существительное будет обладать данным признаком (например, начинаться с некоторого звука или обозначать животное и т. д.), являясь членом данного именного класса. Эта вероятность определяется тем, сколько всего имен входит в конкретный класс. Она обозначена в таблице как  $p(\text{признак } X | \text{СI } \dots)$ . Например, для существительных III класса вероятность того, что слово будет начинаться со звука /b/, составляет 0,10.

Таблица 2

**Признаки — индикаторы класса в речи, обращенной к детям**

Класс	Семантический признак класса	Прочие семантические признаки	Фонологические признаки
I	‘Мужчина’ $p(\text{СI I}   \text{‘мужчина’}) = 0,99$ $p(\text{‘мужчина’}   \text{СI I}) = 0,99$	—	—
II	‘Женщина’ $p(\text{СI II}   \text{‘женщина’}) = 0,99$ $p(\text{‘женщина’}   \text{СI II}) = 0,22$	‘Из бумаги’, ‘одежда’ $p(\text{СI II}   \text{‘из бумаги’ или ‘одежда’}) = 0,52$ $p(\text{‘из бумаги’ или ‘одежда’}   \text{СI II}) = 0,04$	—
III	‘Одушевленный’* $p(\text{СI III}   \text{‘одушевленный’}) = 0,98$ $p(\text{‘одушевленный’}   \text{СI III}) = 0,13$	—	‘Начинается с b-’ $p(\text{СI III}   \text{‘b-’}) = 0,51$ $p(\text{‘b-’}   \text{СI III}) = 0,10$
IV	‘Неодушевленный’	—	‘Начинается с r-’ $p(\text{СI IV}   \text{‘r-’}) = 0,61$ $p(\text{‘r-’}   \text{СI IV}) = 0,09$ ‘заканчивается на -i’ $p(\text{СI IV}   \text{‘-i’}) = 0,54$ $p(\text{‘-i’}   \text{СI IV}) = 0,41$

\*Одушевленными здесь условно названы все живые существа за исключением мужчин и женщин, т. е. животные, а также дүхи.

Мы исходим из предположения, что для взрослых именной класс является частью лексической информации, что означает, что он приписан к слову, а не вычисляется заново при каждом употреблении существительного. В связи же с детской речью возникает вопрос о том, каким путем дети приходят к той системе лексической информации, которой располагают взрослые.

В литературе по усвоению именных классов показано, что дети при усвоении языка обращают внимание и на согласование, и на (внутренние) признаки — индикаторы класса существительных, подобные тем, что приведены в таблице 2. Более того, дети используют внутренние признаки при отнесении новых для них существительных к классу /роду. Это наводит на мысль о том, что оба типа информации имеют вес при усвоении языка. Кроме того, дети более младшего возраста предпочитают использовать морфофонологическую, а не семантическую информацию, несмотря на то что в части случаев именно последняя наиболее надежно предсказывает именной класс.

### 3. Экспериментальное исследование именных классов в речи цезских детей

Нами были записаны 10 часов речи, обращенной к детям во время обычного, повседневного общения на цезском языке (материалы собраны в 2008 и 2009 гг.). Материал нашего корпуса содержит речь матери, тетки и старшей сестры двух двадцатимесячных детей,

живущих в поселке Шамхал (Дагестан). Сбор данных проводился в течение месяца. Около шести часов из этих записей были затранскрибированы с помощью двух членов этой же семьи, носителей цезского языка, в результате чего было получено около 3 000 строк текста. Затем текст был размечен вручную; были обозначены части речи, согласовательная морфология и именные классы. Хотя по меркам корпусной лингвистики наш корпус и мал, в нем все же содержится достаточно данных для того, чтобы было можно оценить распределение признаков в наиболее частотных цезских существительных.

Как уже упоминалось, согласование в цезском языке маркируется только в некоторых глаголах и прилагательных, начинающихся с гласного звука. Ни одна форма, начинающаяся на согласный, не показывает согласования, однако даже не все формы, начинающиеся на гласный, согласуются. Согласующиеся глаголы и прилагательные составляют лишь небольшую часть от общего числа лексем (4 % прилагательных и 27 % глаголов) и, в принципе, могло бы оказаться, что информация о согласовании очень слабо представлена в обращенной к ребенку речи. Тем не менее в нашем корпусе глаголы и прилагательные такого рода достаточно частотны: частота типов (*type frequency*) для согласующихся глаголов и прилагательных составляет соответственно 60 % и 35 %, а частота экземпляров (*token frequency*) еще выше: 84 % всех вхождений глаголов и 77 % вхождений прилагательных. Эти данные усреднены по всем четырем согласовательным показателям (см. примеры (1)–(4)), но примерно такая же картина наблюдается, когда подсчет производится по отдельности для каждого из классов в единственном числе. Таким образом, ребенок, усваивающий цезский язык, может уверенно извлечь информацию о принадлежности имен к классам по данным о согласовании.

Подсчет показывает, что из 114 лексем существительных, представленных в нашем корпусе, у 24 % были признаки — индикаторы класса; если вести подсчет по вхождениям словоформ существительных, то из 1 189 вхождений признаки — индикаторы класса были у 39 %<sup>2</sup>. Подсчет также показывает, что внутренние признаки — индикаторы класса (см. таблицу 2) присутствуют у 100 % участвующих в согласовании существительных I класса, 52 % таких существительных II класса, 51 % существительных III класса и 45 % существительных IV класса. Таким образом, информация о классах имен, необходимая ребенку, усваивающему язык, вполне доступна. То, что кажется малочастотным и редким при рассмотрении слов в словаре, оказывается широко распространено в речи, обращенной к детям, что, в свою очередь, позволяет им эффективно усвоить информацию о классах.

Чтобы подробнее исследовать, насколько дети и взрослые чувствительны к свойствам, характеризующим существительные каждого из классов, мы провели эксперимент, определяющий, какими свойствами пользуются носители, чтобы классифицировать как знакомые, так и новые существительные, в которых представлены признаки, позволяющие отнести их к тому или иному классу.

В эксперименте были использованы как реально существующие в языке, так и специально составленные искусственные слова (*nonce words*)<sup>3</sup>. Часть слов обладала каким-либо одним из выделенных нами признаков — маркеров класса, у некоторой группы слов присутствовало сочетание двух признаков и, наконец, в эксперимент были включены в качестве контрольной группы слова, не обладающие ни одним из явных признаков, которые мы изучали.

В таблице 3 исчислены сочетания признаков, присутствующих у слов каждого из усваиваемых классов. У каждого из слов, использованных в эксперименте, имеется основной семантический признак, позволяющий очертить группы слов, соответствующие мужскому

<sup>2</sup> Все наши подсчеты исключают имена собственные. Это, возможно, уменьшает как долю существительных с признаками — индикаторами класса, так и долю согласующихся существительных, если рассматривать пол обладателя имени собственного как признак-индикатор для соответствующего существительного.

<sup>3</sup> Полный список использованных слов содержится в Приложении.

полу, женскому полу, животным; в группе неодушевленных существительных были выделены дополнительные, более тонкие семантические признаки — принадлежность к предметам одежды или принадлежность к бумажным изделиям; наконец, были выделены формальные фонологические признаки (например, исходный звук слова).

Реальные цезские слова брались из нашего корпуса. Если слова с нужным нам сочетанием семантических и формальных признаков там не оказывалось, то вместо него подбирались семантически близкий эквивалент английского слова, часто встречающегося в английских корпусах обращенной к детям речи. Искусственные слова были составлены в соответствии с цезской фонотактикой и затем проверены носителями языка, чтобы удостовериться в том, что никакое из них случайно не совпало с реально существующим словом.

Таблица 3

**Состав лексики, обладающей выделенными признаками  
и использованной в эксперименте**

Класс	Основной семантический признак	Прочие семантические признаки	Фонологический признак	Согласованные признаки	Несогласованные признаки
I	‘Мужчина’ (3/3)	—	—	—	‘Мужчина’ и ‘начинается с y-’ (0/3) ‘Мужчина’ и ‘начинается с b-’ (0/3)
II	‘Женщина’ (3/3)	‘Из бумаги’ (3/3); ‘одежда’ (3/3)	‘Начинается с y-’ (3/3)	‘Женщина’ и ‘начинается с y-’ (0/3)	‘Женщина’ и ‘начинается с r-’ (0/3)
III	‘Одушевленный’ (3/3)	—	‘Начинается с b-’ (3/3)	‘Животное’ и ‘начинается с b-’ (3/3)	‘Животное’ и ‘начинается с r-’ (2/3)
IV	‘Неодушевленный’	—	‘Начинается с r-’ (3/3) ‘кончается на -i’ (3/3)	‘Начинается с r-’ и ‘кончается на -i’ (2/3)	‘Неодушевленный’ и ‘начинается с b-’ (3/0)

Числа в скобках показывают количество в выборке реальных / искусственных слов с данным признаком.

Чтобы придать искусственным словам нужные семантические признаки, испытуемым в эксперименте сообщалось, что это названия неизвестных ранее животных или редкие имена людей. Для того чтобы сравнить, как носители языка классифицируют существительные, обладающие признаками-индикаторами, и существительные, лишенные таких признаков, в выборку были также включены и искусственные существительные без каких-либо семантических или фонологических признаков-индикаторов.

Задание в эксперименте было очень простым и легко доступным для детей. В цезском языке есть два глагола со значением ‘есть, кушать’: непереходный *-iš* и переходный *-ac*’. Оба глагола высоко частотны и согласуются с абсолютивом. Все экспериментальные предложения содержали эти глаголы.

Сначала испытуемого обучали выполнению задания. Дети считали, что их задача — разделить предметы на те, что можно есть, и те, что нельзя.

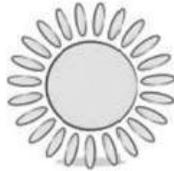
Ребенку показывали игрушку-фигурку, и он должен был скомандовать фигурке начать есть, используя непереходный глагол *-iš*. При этом в глаголе появлялся согласовательный префикс, указывавший на то, к какому классу ребенок относит фигурку. Например, если ребенок относил фигурку к классу III, он говорил ей *b-iš!*, а если к классу I, то *Ø-iš!*

Затем фигурка двигалась по набору картинок, а экспериментатор вслух называл нарисованные там предметы. Ребенок должен был говорить фигурке, есть ей или не есть каждой из них. При этом ребенок употреблял переходный глагол *-ac'o* (повелительная форма от *-ac'*), и согласовательный префикс на глаголе показывал, к какому классу ребенок относит изображенный предмет — прямое дополнение глагола. Экспериментатор записывал, какой согласовательный показатель был использован с каждым из предметов. Эксперимент проводился с большим числом картинок, и в итоге были задействованы все существительные из нашей выборки (общим числом 107). Пример страницы с картинками показан на рисунке, а идеализированная расшифровка диалога во время эксперимента — в таблице 4.



*kid*  
'девочка'

класс II  
семантический  
признак:  
'женщина'



*buq*  
'солнце'

класс III  
фонологический  
признак:  
'начинается с b-'



*k'uray*  
'лук'

класс IV  
без признаков



*zamil*  
искусственное слово

класс III  
семантический  
признак:  
'животное'

Рис. Образцы изображений, использованных в эксперименте

Таблица 4

Фрагмент эксперимента, расшифровка записи

Говорящий	Реплика	Действие
АЭ (Ассистент экспериментатора, носитель языка)	<i>Kid.</i> 'Девочка'. (Класс II)	Объясняет задание, показывает и называет предмет
Р(ребенок)	<i>Sis q'ano ʔono</i> один два три <i>y-iš!</i> II-есть.НЕПЕР.ИМПЕР 'Раз, два, три, ешь!'	Командует фигурке; глагол согласуется с абсолютивом подлежащего
АЭ	<i>Buq.</i> 'Солнце'. (Класс III)	Показывает на солнце и называет его
Р	<i>Buq</i> солнце.АБС <i>b-ac'-xosi ānu!</i> III-есть-ПРИЧ ОTR 'Солнце несъедобное'.	Командует фигурке / описывает ситуацию
АЭ	<i>K'uray.</i> 'Лук'. (Класс IV)	Показывает на лук и называет его

Таблица 4 (окончание)

Говорящий	Реплика	Действие
Р	K'uraŋ g-ac'-o! лук IV-есть.ПЕРЕХ-ИМПЕР 'Ешь лук!'	Командует фигурке / описывает ситуацию; глагол согласуется с абсолютным дополнением
АЭ	Zamił. Искусственное слово, должно относиться к III классу	Показывает на вымышленное животное и называет его
Р	Zamił замил.АБС b-ac'-xosi ānu! III-есть.ПЕРЕХ-ПРИЧ ОTR 'Замил несъедобный'.	Командует фигурке / описывает ситуацию; причастие согласуется с искусственным словом

Испытуемыми были носители цезского языка, живущие в поселке Шамхал и городе Кызыльюрт (Республика Дагестан), которые были привлечены к участию в эксперименте при помощи ассистента — местного жителя, знакомого с цезскими семьями, живущими там. Эксперимент проходил в 2009 г. Он проводился в отдельной комнате, где находились экспериментатор, ассистент — носитель цезского языка, ребенок, а также иногда родители ребенка, другие родственники или еще какие-нибудь сопровождающие взрослые, которым было заранее объяснено, что во время эксперимента они должны молчать. Разрешалось только подбодрять ребенка, если тот очень стеснялся и не хотел выполнять задание.

#### 4. Основные результаты

Результаты, представленные ниже, отражают материал, собранный в группе младших детей (от четырех до семи лет, всего 10 испытуемых), в группе более старших детей (от восьми до двенадцати лет, всего 12 испытуемых), а также в контрольной группе из 10 взрослых<sup>4</sup>.

Экспериментальные данные анализировались следующим образом: для каждого типа примеров (скажем, для искусственных слов с семантическим признаком 'женский род' или для существующих слов с фонологическим признаком 'начинается с b-') была вычислена доля ответов, в которых экспериментальные слова были отнесены к каждому из классов, среди общего числа ответов. Эти ответы были просчитаны по отдельности для каждой из возрастных групп. Например, у младших детей классификация искусственных слов с семантическим признаком «женский род» оказалась такой: в 4% ответов они были отнесены к I классу, в 52% ко II классу, в 22% к III классу и в 22% к IV классу.

Для каждого набора слов с заданным признаком или комбинацией признаков (см. таблицу 3) была вычислена доля ответов, в которых они были отнесены к каждому из классов, среди общего числа ответов про эти слова. Таким образом, для каждого из наборов слов и для каждой возрастной группы мы получили распределение ответов о принадлежности этих слов к именованным классам.

Мы также провели сравнение классификации реально существующих слов с их классификацией в словаре. Для классификации искусственных слов такого сравнения быть не может, но здесь мы провели сравнение между классификацией искусственных слов, обладающих признаками-индикаторами, и классификацией искусственных слов, таких признаков лишенных.

Говоря о классификации искусственных слов, мы подразумеваем долю ответов, в которых эти слова относятся к ожидаемому классу — тому, на который сильнее всего (в смысле приведенных в таблице 2 условных вероятностей) указывают признаки-индикаторы данного

<sup>4</sup> Мы разделяем детей только на две группы по возрасту, поскольку общее число детей, участвовавших в эксперименте, относительно невелико.

слова, по сравнению с долей таких же ответов для слов, лишенных признаков-индикаторов. Например, ожидаемый класс искусственного слова, обозначающего женщину, — второй. Мы сравнивали долю ответов про искусственные слова с этим семантическим признаком, который позволяет их относить ко II классу, с долей ответов про слова, не имеющие этого признака, но тем не менее отнесенные к этому классу.

Наш анализ позволил обнаружить следующие общие тенденции. Во-первых, классификация искусственных слов с фонологическими или семантическими индикаторами принадлежности к I, II и III классам значительно отличалась от классификации искусственных слов, лишенных признаков-индикаторов, причем различия между возрастными группами не наблюдалось. Это показывает, что все носители цезского языка подсознательно обращают внимание на семантические и фонологические классные признаки.

Во-вторых, в обеих группах детей классификация искусственных слов с противоречивыми индикаторами отличалась от классификации слов, у которых были либо только семантические, либо только фонологические индикаторы; тем самым, дети были чувствительны к взаимодействию признаков. У взрослых такого различия между подгруппами искусственных слов не наблюдалось.

В-третьих, дети младшей группы классифицировали реально существующие слова с противоречивыми индикаторами не так, как они классифицировались в словаре. Это наводит на мысль о том, что младшие дети еще находятся в процессе построения системы признаков, по которым они могут относить слова к классам. Чем старше дети, тем более привычны для них формальные признаки, тогда как семантические признаки, особенно то, что выходит за пределы обычной одушевленности, по-прежнему вызывают затруднения. Этот результат хорошо согласуется с наблюдениями об усвоении рода французскими детьми, которые в старших группах тоже отдают приоритет формальным признакам (см. [Karmiloff-Smith 1979]).

Таблица 5

**Процент реально существующих слов, правильно отнесенных к своему классу**

	<b>Семантический признак пола / одушевленности</b>	<b>Прочие семантические признаки</b>	<b>Фонологические признаки</b>	<b>Отсутствие признаков-индикаторов</b>	<b>Несогласованные признаки</b>
Младшие дети	79	71	84	77	42
Старшие дети	86	58	94	78	47
Взрослые	87	75	92	86	71

Из таблицы 5 видно, что ни в одной из групп и ни для одного из типов слов их отнесение к грамматическому классу не оказалось стопроцентным. Однако высокие доли правильного отнесения имен к классам заставляют нас предположить, что по крайней мере часть слов была известна всем трем возрастным группам. Если это так, то часть отклонений от 100 % может объясняться незнанием слов или отсутствием быстрого лексического доступа, а другая часть — шумом в эксперименте. В свою очередь, соотношение лексических лагун и шума требует дальнейших исследований, которые не ограничиваются одним языком. Наличие шума в эксперименте подтверждается и тем, что отклонения от правильной классификации оказались равномерно распределены по лексемам; если бы испытуемые не знали каких-либо отдельных слов, мы могли бы ожидать ошибок только в этих словах, но этого не произошло.

Следует также иметь в виду, что многие дети и даже взрослые не были привычны к структурированным тестам, что могло увеличить число неправильных ответов. Кроме

того, условия, в которых проводились эксперименты, были далеки от идеальных, так как найти тихую комнату без отвлекающих факторов возможности не было. Это тоже приводит к увеличению числа ошибок. Тем не менее можно констатировать, что представители всех трех возрастных групп очень хорошо справились с классификацией слов, обладающих только семантическим или фонологическим признаком или лишенных каких бы то ни было признаков. В тех же случаях, когда признаки противоречили фактическому классу слова, дети из обеих групп справлялись с заданием хуже. Во всех этих случаях информация, не соответствующая настоящему классу слова, была фонологической. Так, например, *recenyo* ‘муравей’ принадлежит III классу, но начинается с *r-*, что служит индикатором IV класса. Дети часто делали ошибки в выборе класса этого слова, относя его не к III, а к IV классу. Это значит, что для детей фонологический индикатор определенного класса перевешивает возможный опыт обращения с данным словом.

В отличие от классификации реально существующих слов, в отношении которых мы ожидали, что большинство из них будет отнесено к их реальному грамматическому классу, для искусственных слов можно предположить, что они будут классифицироваться в соответствии с признаками-индикаторами, если таковые имеются. При наличии каких-либо опознавательных признаков ожидаемым классом искусственного существительного мы называем тот, который наиболее значимо предсказывается внутренними признаками данного слова. В таблице 6 приведены проценты слов, правильно отнесенных к ожидаемому классу.

Таблица 6

**Процент искусственных слов, отнесенных к ожидаемому классу**

	<b>Семантический признак пола / одушевленности</b>	<b>Прочие семантические признаки</b>	<b>Фонологические признаки</b>	<b>Несогласованные признаки</b>
Младшие дети	54	8	61	38
Старшие дети	65	9	63	53
Взрослые	53	23	61	55

Первое, что бросается в глаза, — это степень совпадения классификаций во всех трех группах испытуемых. Это означает, что все группы чувствительны к признакам идентификации слов, которые предсказывают грамматический класс. Можно видеть, что при классификации искусственных слов задействованы и семантические, и фонологические признаки; существенная часть слов, обладающих такими признаками, оказывается отнесена к ожидаемому классу представителями всех возрастных групп. Для классов I и II получающееся распределение существенно отличается от дефолтного, показанного в таблице 1, в то время как для классов III и IV наблюдаемое и дефолтное распределения оказываются очень близки в тех случаях, когда подходящие признаки-индикаторы присутствуют. В тех же случаях, когда признаки-индикаторы отсутствуют, отклонения от дефолтного распределения значительно сильнее. Полная классификация для каждого из типов признаков дается в статье [Gagliardi, Lidz 2014].

Перейдем к более специфическим семантическим признакам. Так, [Plaster et al. 2013] показывают, что более специфические семантические признаки (например, в случае нашей выборки — ‘быть из бумаги’) также играют роль в отнесении имен к классам; в частности, это признаки, которые были использованы в нашей работе для отнесения имен к классу II. Наш эксперимент показал, что дети менее активно пользуются этими семантическими признаками, чем взрослые (различие между 8 % и 9 % слов с этими признаками, отнесенных к классу II, и 1 % слов, лишенных каких бы то ни было индикаторов и отнесенных, тем не менее, к тому же классу, оказалось статистически не значимо). Между тем взрослые активно опираются на эти признаки. Так, они отнесли слова, обозначающие бумажные

предметы и одежду, к классу II в 23 % случаев, что статистически значимо отличается от процента слов, не содержащих признака-индикатора, которые были отнесены к этому классу. В целом эти количественные данные показывают, что значимость индикаторов с семантикой, не связанной с полом и одушевленностью, при отнесении цезских имен к классам существенно ниже.

Семантическая информация обладает наибольшей предсказательной силой для искусственных слов с противоречивыми признаками-индикаторами. У таких искусственных слов, если семантика предлагает принадлежность к одному классу, а фонологические признаки указывают на другой класс, выбор делается в пользу семантики. Однако такая стратегия — признак носителей языка, которые уже в полной мере им владеют. В нашем случае это взрослые и дети старшей группы. Младшие дети даже в такой ситуации ориентируются в большей степени на фонологическую информацию.

Подведем итоги. Мы продемонстрировали, что, используя экспериментальные материалы и прибегая к помощи информированных носителей языка, в принципе возможно изучать усвоение детьми малых бесписьменных языков. Мы установили, что усвоение именных классов в цезском — языке с большим числом именных классов (родов), чем в языках, которые обычно подлежат изучению в работах по детской речи (например, если сравнивать с индоевропейским родом), следует тем же принципам, которые были ранее обнаружены в исследованиях детской речи. Дети определяют именные классы, опираясь на сочетание базовых семантических признаков (пол, одушевленность) и очевидных формальных признаков; в цезском языке к последним относятся исходный и конечный звук существительного в абсолютиве — падеже, с которым всегда согласуется глагол. Однако дети более младшего возраста при отнесении имени к грамматическому классу ориентируются в первую очередь на фонологический облик слова, а не на его семантические признаки.

## ПРИЛОЖЕНИЕ

### Слова, использованные в эксперименте

Тип слова	Русский перевод	Цезское слово
искусственное, I, несогласованные признаки	новое наименование мужчины	<i>yasi</i>
искусственное, I, несогласованные признаки	новое наименование мужчины	<i>yeža</i>
искусственное, I, несогласованные признаки	новое наименование мужчины	<i>banu</i>
искусственное, I, несогласованные признаки	новое наименование мужчины	<i>yušon</i>
искусственное, I, несогласованные признаки	новое наименование мужчины	<i>bino</i>
искусственное, I, несогласованные признаки	новое наименование мужчины	<i>buma</i>
искусственное, I, семантический признак	новое наименование мужчины	<i>cina</i>
искусственное, I, семантический признак	новое наименование мужчины	<i>kirop</i>
искусственное, I, семантический признак	новое наименование мужчины	<i>melu</i>
искусственное, II, согласованные признаки	новое наименование женщины	<i>yeħu</i>
искусственное, II, согласованные признаки	новое наименование женщины	<i>yunik</i>
искусственное, II, согласованные признаки	новое наименование женщины	<i>yina</i>
искусственное, II, несогласованные признаки	новое наименование женщины	<i>riħu</i>
искусственное, II, несогласованные признаки	новое наименование женщины	<i>rak'o</i>
искусственное, II, несогласованные признаки	новое наименование женщины	<i>ruya</i>
искусственное, II, фонологический признак	новое наименование пищи	<i>yobar</i>

Продолжение таблицы приложения

Тип слова	Русский перевод	Цезское слово
искусственное, II, фонологический признак	новое наименование предмета	<i>yuto</i>
искусственное, II, фонологический признак	новое наименование пищи	<i>yala</i>
искусственное, II, универсальный семантический признак	новое наименование женщины	<i>kuna</i>
искусственное, II, универсальный семантический признак	новое наименование женщины	<i>haba</i>
искусственное, II, универсальный семантический признак	новое наименование женщины	<i>sohaq</i>
искусственное, II, специфический семантический признак	новое наименование бумаги	<i>molo</i>
искусственное, II, специфический семантический признак	новое наименование одежды	<i>lemin</i>
искусственное, II, специфический семантический признак	новое наименование бумаги	<i>mačum</i>
искусственное, II, специфический семантический признак	новое наименование одежды	<i>kenu</i>
искусственное, II, специфический семантический признак	новое наименование бумаги	<i>ħidar</i>
искусственное, II, специфический семантический признак	новое наименование одежды	<i>zubu</i>
искусственное, III, согласованные признаки	новое наименование животного	<i>bazu</i>
искусственное, III, согласованные признаки	новое наименование животного	<i>budu</i>
искусственное, III, согласованные признаки	новое наименование животного	<i>bičan</i>
искусственное, III, несогласованные признаки	новое наименование животного	<i>yugi</i>
искусственное, III, несогласованные признаки	новое наименование животного	<i>resu</i>
искусственное, III, несогласованные признаки	новое наименование животного	<i>riga</i>
искусственное, III, несогласованные признаки	новое наименование животного	<i>čohi</i>
искусственное, III, несогласованные признаки	новое наименование животного	<i>rola</i>
искусственное, III, несогласованные признаки	новое наименование животного	<i>t'awi</i>
искусственное, III, фонологический признак	новое наименование блюда	<i>betu</i>
искусственное, III, фонологический признак	новое наименование блюда	<i>baka</i>
искусственное, III, фонологический признак	новое наименование блюда	<i>bidan</i>
искусственное, III, семантический признак	новое наименование животного	<i>zamil</i>
искусственное, III, семантический признак	новое наименование животного	<i>seno</i>
искусственное, III, семантический признак	новое наименование животного	<i>kiru</i>
искусственное, IV, согласованные признаки	новое наименование блюда	<i>rubi</i>
искусственное, IV, согласованные признаки	новое наименование блюда	<i>rehi</i>
искусственное, IV, согласованные признаки	новое наименование блюда	<i>rabi</i>
искусственное, IV, фонологический признак	новое наименование блюда	<i>tali</i>
искусственное, IV, фонологический признак	новое наименование предмета	<i>joni</i>
искусственное, IV, фонологический признак	новое наименование предмета	<i>q'omi</i>

Продолжение таблицы приложения

Тип слова	Русский перевод	Цезское слово
искусственное, IV, фонологический признак	новое наименование предмета	<i>rega</i>
искусственное, IV, фонологический признак	новое наименование блюда	<i>rulo</i>
искусственное, IV, фонологический признак	новое наименование блюда	<i>rinay</i>
искусственное, без признаков	новое наименование блюда	<i>miraj</i>
искусственное, без признаков	новое наименование блюда	<i>lesi</i>
искусственное, без признаков	новое наименование блюда	<i>kola</i>
искусственное, без признаков	новое наименование блюда	<i>nola</i>
искусственное, без признаков	новое наименование блюда	<i>kela</i>
искусственное, без признаков	новое наименование блюда	<i>šiwa</i>
искусственное, без признаков	новое наименование блюда	<i>dero</i>
искусственное, без признаков	новое наименование предмета	<i>norib</i>
искусственное, без признаков	новое наименование блюда	<i>žewu</i>
искусственное, без признаков	новое наименование блюда	<i>nawe</i>
реально существующее, I, семантический признак	младенец	<i>k'ak'a</i>
реально существующее, I, семантический признак	мальчик	<i>uži</i>
реально существующее, I, семантический признак	отец	<i>baba</i>
реально существующее, II, без признаков	соль	<i>cijo</i>
реально существующее, II, без признаков	дверь	<i>ac</i>
реально существующее, II, без признаков	сыр	<i>izu</i>
реально существующее, II, фонологический признак	камень	<i>yuł</i>
реально существующее, II, фонологический признак	молоко	<i>ɣaj</i>
реально существующее, II, фонологический признак	штаны	<i>ɣel'o</i>
реально существующее, II, универсальный семантический признак	женщина	<i>ɣana</i>
реально существующее, II, универсальный семантический признак	девочка	<i>kid</i>
реально существующее, II, универсальный семантический признак	мать	<i>eni</i>
реально существующее, II, специфический семантический признак	письмо	<i>kayat</i>
реально существующее, II, специфический семантический признак	рубашка / платье	<i>ged</i>
реально существующее, II, специфический семантический признак	трусы	<i>turusik</i>
реально существующее, II, специфический семантический признак	шапка	<i>šapka</i>
реально существующее, II, специфический семантический признак	книга	<i>t'ek</i>
реально существующее, II, специфический семантический признак	газета	<i>gazit</i>
реально существующее, III, согласованные признаки	рыба	<i>besuro</i>

Окончание таблицы приложения

Тип слова	Русский перевод	Цезское слово
реально существующее, III, согласованные признаки	змея	<i>bikori</i>
реально существующее, III, согласованные признаки	овца	<i>be'ʔyu</i>
реально существующее, III, несогласованные признаки	море	<i>raʔad</i>
реально существующее, III, несогласованные признаки	муравей	<i>recenyo</i>
реально существующее, III, без признаков	яблоко	<i>heneš</i>
реально существующее, III, без признаков	картофель	<i>hek'u</i>
реально существующее, III, без признаков	хлеб	<i>magalu</i>
реально существующее, III, фонологический признак	солнце	<i>buq</i>
реально существующее, III, фонологический признак	вишня	<i>ba'li</i>
реально существующее, III, фонологический признак	палец	<i>baša</i>
реально существующее, III, семантический признак	цыпленок	<i>onoču</i>
реально существующее, III, семантический признак	корова	<i>ziya</i>
реально существующее, III, семантический признак	кошка	<i>k'et'u</i>
реально существующее, IV, несогласованные признаки	будка (или туалет?)	<i>butka</i>
реально существующее, IV, несогласованные признаки	флаг	<i>bairaq</i>
реально существующее, IV, несогласованные признаки	кольцо	<i>basčiqow</i>
реально существующее, IV, без признаков	лук	<i>k'uray</i>
реально существующее, IV, без признаков	суп	<i>čorpa</i>
реально существующее, IV, без признаков	глаз	<i>ozura</i>
реально существующее, IV, фонологический признак	вода	<i>li</i>
реально существующее, IV, фонологический признак	каша	<i>qiqi</i>
реально существующее, IV, фонологический признак	окно	<i>aki</i>
реально существующее, IV, фонологический признак	рука	<i>reʔ'a</i>
реально существующее, IV, фонологический признак	сливочное масло	<i>riʔ</i>
реально существующее, IV, фонологический признак	ключ	<i>reka</i>
реально существующее, IV, согласованные признаки	мусор	<i>rešoni</i>
реально существующее, IV, согласованные признаки	колыбель	<i>rikini</i>

## СПИСОК СОКРАЩЕНИЙ

- I, II, III, IV — именной класс в единственном числе  
 АБС — абсолютив  
 ИМПЕР — императив  
 НЕПЕР — непереходный глагол  
 ОTR — отрицание  
 ПЕРЕХ — переходный глагол  
 ПРИЧ — причастие  
 ПРОШ — прошедшее время  
 ЭРГ — эргатив

## СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Бокаре́в 1959 — Бокаре́в А. Н. Цезские (дидойские) языки Дагестана. М.: Издательство АН СССР, 1959. [Bokarev A. N. *Tsezskie (didoiskie) yazyki Dagestana* [Tsez (Dido) languages of Dagestan]. Moscow: Academy of Sciences of the USSR, 1959.]
- Гвозде́в 1961 — Гвозде́в А. Н. Вопросы изучения детской речи. М.: Наука, 1961. [Gvozdev A. N. *Voprosy izucheniya detskoj rechi* [Problems of child speech study]. Moscow: Nauka, 1961.]
- Имнайшвили 1963 — Имнайшвили Д. С. Дидойский язык в сравнении с гинухским и хваршийским языками. Тбилиси: Академия наук Грузинской ССР, 1963. [Imnaishvili D. S. *Didoiskii yazyk v sravnenii s ginukhskim i khvarshinskimi yazykami* [The Dido language in comparison with the Hinukh and Khvarshin languages]. Tbilisi: Academy of Sciences of the Georgian SSR, 1963.]
- Халилов 1999 — Халилов М. Ш. Цезско-русский словарь. М.: Academia, 1999. [Khalilov M. Sh. *Tsezskorusskii slovar'* [Tsez-Russian dictionary]. Moscow: Academia, 1999.]
- Comrie, Polinsky 1999 — Comrie B., Polinsky M. Some observations on class categorization in Tsez. *Studies in Caucasian linguistics: Selected papers of the Eighth Caucasian Colloquium*. Van den Berg H. (ed.). Leiden: Universiteit Leiden, 1999. Pp. 125—139.
- Gagliardi 2012 — Gagliardi A. *Input and intake in language acquisition*. Doct. dissertation. University of Maryland, 2012.
- Gagliardi, Lidz 2014 — Gagliardi A., Lidz J. Statistical insensitivity in the acquisition of Tsez noun classes. *Language*. Vol. 90. No. 1. Pp. 58—89.
- Karmiloff-Smith 1979 — Karmiloff-Smith A. *A functional approach to child language*. Cambridge: Cambridge University Press, 1979.
- MacWhinney 1978 — MacWhinney B. The acquisition of morphophonology. *Monographs of the Society for research in child development*. 1978. Vol. 43. No. 1/2. Pp. 1—123.
- Mills 1985 — Mills A. E. The acquisition of German. *The crosslinguistic study of language acquisition*. Slobin D. I. (ed.). Vol. 1. Hillsdale (NJ): Lawrence Erlbaum, 1985. Pp. 141—254.
- Mills 1986 — Mills A. E. *The acquisition of gender: A study of English and German*. Berlin: Springer, 1986.
- Plaster et al. 2013 — Plaster K., Polinsky M., Harizanov B. Noun classes grow on trees: Noun classification in the North-East Caucasus. *Language typology and historical contingency: In honor of Johanna Nichols*. Bickel B., Grenoble L. A., Peterson D. A., Timberlake A. (eds). Amsterdam: John Benjamins, 2013. Pp. 153—170.
- Rodina 1999 — Rodina Y. *Semantics and morphology: The acquisition of grammatical gender in Russian*. Doct. dissertation. Tromsø: University of Tromsø, 2009.
- Szagun et al. 2007 — Szagun G., Stumper B., Sondag N., Franik M. The acquisition of gender marking by young German-speaking children: Evidence for learning guided by phonological regularities. *Child Language*. 2007. Vol. 34. No. 3. Pp. 445—471.

Статья поступила в редакцию 07.09.2014.