Associative plural suffix -txe- in Adyghe

Evgeniya Murzinova, Maria Yaroslavtseva

About Adyghe

Adyghe, also known as West Circassian, is one of the two official languages of the Republic of Adygea in the Russian Federation. It is spoken by various tribes of the Adyghe people, each with its own dialect. Outside Circassia Adyghe is spoken by Circassians in Iraq, by Circassians in Israel, and by Circassians in Syria.

There are around 128,000 speakers of Adyghe in Russia, almost all of them native speakers. In total, some 300,000 speak it worldwide. The largest Adyghe-speaking community is in Turkey, spoken by the post Russian–Circassian War (circa 1763–1864) diaspora; in addition to that, the Adyghe language is spoken by the Cherkesogai in Krasnodar Krai. Adyghe belongs to the family of Northwest Caucasian languages. Kabardian (also known as East Circassian) is a very close relative, treated by some as a dialect of Adyghe or of an overarching Circassian language. Ubykh, Abkhaz and Abaza are somewhat more distantly related to Adyghe.mnb The language was standardised after the October Revolution in 1917. Since 1936, the Cyrillic script has been used to write Adyghe. Before that, an Arabic-based alphabet was used together with the Latin.

Data collection and creation of database

The research is based upon primary data that was collected in Krasnodar region during several interview sessions with native speakers. All the interviewees speak Bzhedug dialect (which is not formal but the most widely distributed among native speakers).

Present research is built upon Animacy Hierarchy [Silverstein 1976]:

1. 1,2 > 3
2. ProN > ProperN > Kin > CommonN
3. Hum > NHum > Inanim
4. Def > Specific > Non-specific
5. Sg > Pl > Mass

Line 4 is excluded from the examination due to impossibility to check appearance of the suffix out of context. In our study we have used constructed examples not controlled for the referential status of noun phrases and so we excluded line 4 (Def > Specific > Non-specific) from the examination. Further investigation is needed.

Lines of this hierarchy are not mutually exclusive so we have combined them into one table, grouping intercross attributes into levels. That work can be described by following list of actions:

- creating a list of words from the questionnaires
- dividing the list into classe: Pronouns (pron), Proper Nouns (propern), Kinship Terminology (kin),

Common Nouns (commonn);

- assign each word with attribute of animacy: Human (hum), Non Human (nhum), Inanimated (inanim) and number: Singular (sg), Plural (pl);
- information about ability to add APL suffix -тхэ- was added (yes, no)

The Associative Plural

Associative plural constructions consist of a noun **X** (typically of human reference, usually a person's name or a kin term) and some other material, most often an affix, a clitic, or a word. The meaning of the construction is **'X and other people associated with X'**. Associative plural constructions are identified by two semantic properties: referential heterogeneity and reference to groups. The first property, referential heterogeneity, distinguishes associative plurals from additive plurals which happen (the most common type in most languages). The second semantic feature of the associative plural is reference to groups: it refers to a close-knit group of individuals rather than to sets without any internal cohesion. This characteristic ties associative plurals to collective plurals such as Russian 'listva' - 'a mass of leaves'; the difference is that collective plurals are referentially homogeneous. [Michael Daniel, Edith Moravcsik. 2013]

Here is the definition of Associative Plural:

'X and X's associate(s)', where all members are individuals, X is the focal referent, and the associate(s) form a group centering around X

In Adyghe The Associative Plural concept is realized by suffix -тхэ- (-txe-) The associative plural form of the name N (1) is opposed to additive plural form (2). Second one consists of several objects, each of which is a referent of N:

> 1. шыпхъутхэр
>   sister-APL
>   'sister and her associates'
> 2. шыпхъухэр
>   sister-PL
>   'sisters'

Hypothesis

We assume that in Adyghe the possibility of attaching suffix -txe- is higher when the focal referent is singular and animated.

Our data

As a first step we load the data using the read.csv() function. Then we list the first 6 lines and summary of our data with head() and summary():

```
ourdata <- read.csv("F:/Project/txe.pro-utf8.csv", header=TRUE)
head (ourdata)
```

```
##   isSuffix Class Number Animacy word
## 1       no  pron     sg     hum   сэ
## 2       no  pron     sg     hum    о
## 3       no  pron     sg     hum   ар
## 4       no  pron     sg    nhum   ар
## 5       no  pron     sg  inanim   ар
## 6       no  pron     pl     hum   тэ
```

```
summary(ourdata)
```

```
##  isSuffix       Class      Number      Animacy          word
##  no :63   commonn:24   pl:36   hum    :46   ар     : 3
##  yes:13   kin    :16   sg:40   inanim:16   ахэр   : 3
##           pron   :22           nhum   :14   мо     : 3
##           propern:14                        мохэр  : 3
##                                             мы     : 3
##                                             мыхэр  : 3
##                                             (Other):58
```

Visualization

In order to visualize the data, we need to run library(tidyverse)

```
library(tidyverse)
```

```
## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr
```

```
## Conflicts with tidy packages ----------------------------------------------
```

```
## filter(): dplyr, stats
## lag():    dplyr, stats
```

We may see the correlation between existence of Associative Plural forms and the class of the focal referent: common nouns do not attach the suffix -txe- at all.

Possibility to attach suffix -txe- by Class
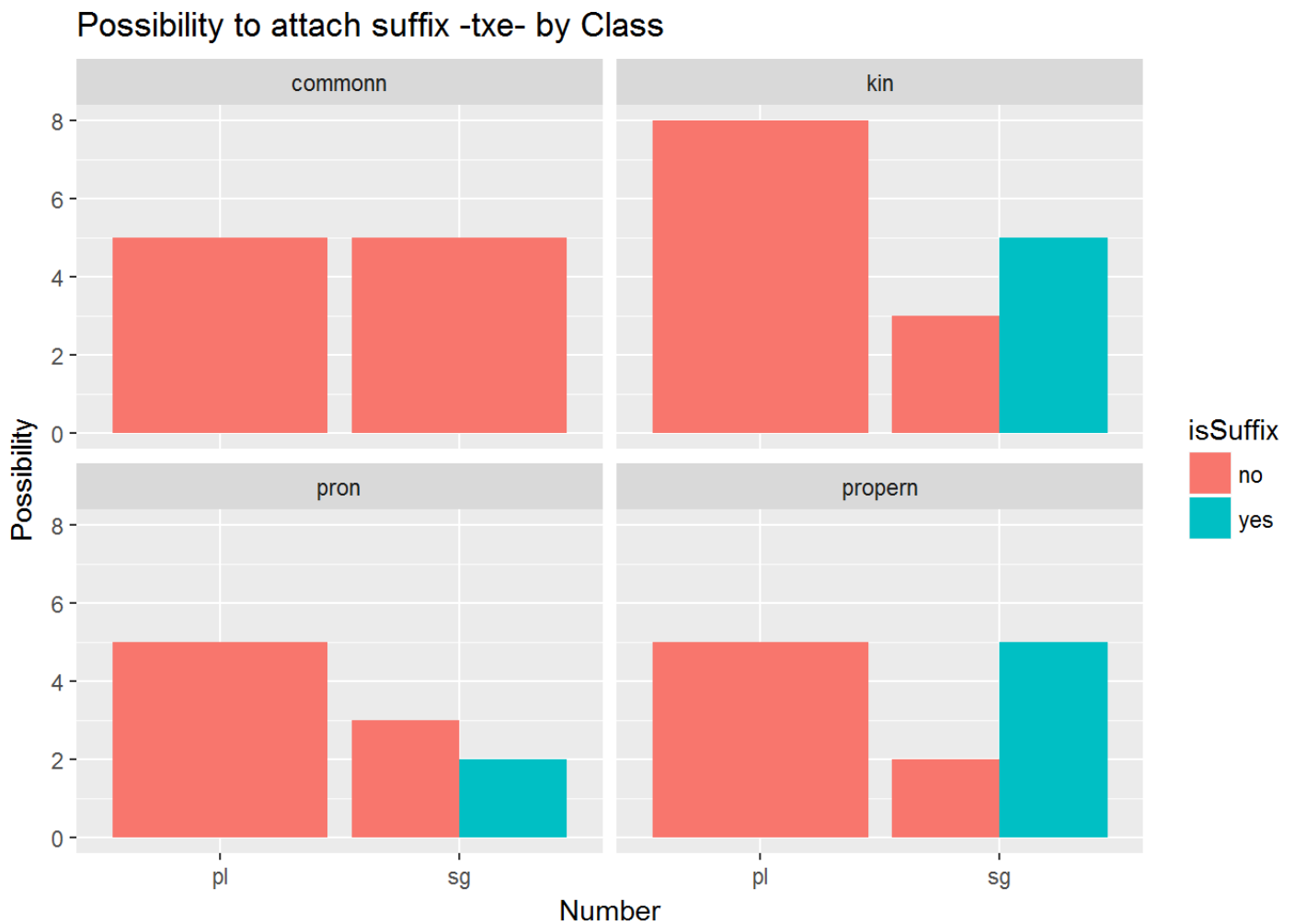
Next bar demonstrates the correlation between Associative Plural forms and the animacy of the focal referent: animated and non-human (with a proper noun as a FR) can attach the suffix -txe-.

Possibility to attach suffix -txe- by Animacy
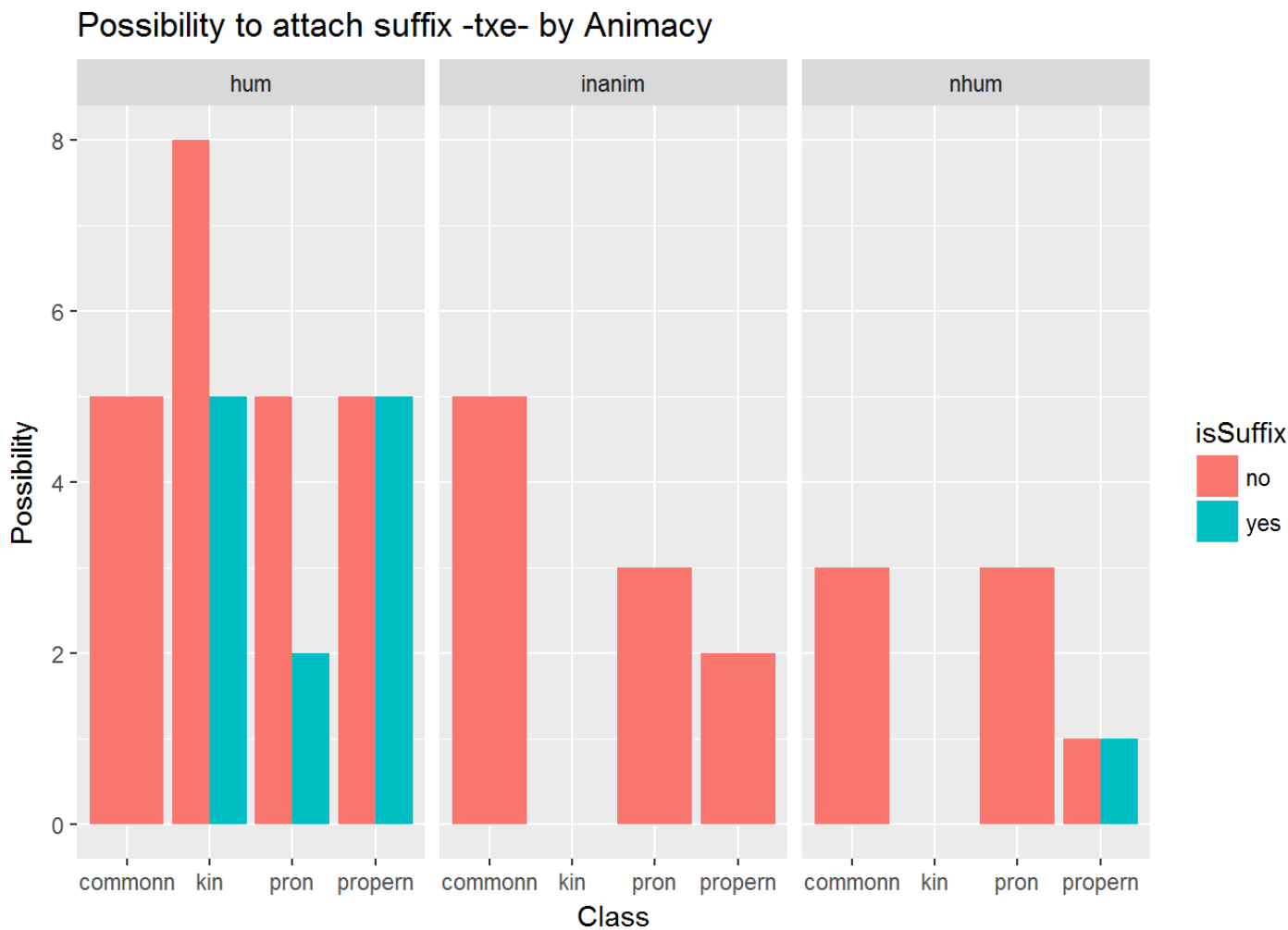
Logistic regression

Logistic regression is a method for fitting a regression curve when **y** is a categorical variable. The typical use of this model is predicting **y** given a set of predictors **x**. The predictors can be continuous, categorical or a mix of both.

The categorical variable **y**, in general, can assume different values. In the simplest case scenario **y** is binary meaning that it can assume either the value 1 or 0, in our case, **yes** or **no** (about possibility to form associative plural)

Let us build several models with different predictors. We suppose that the best predictor for our model should be **Animacy** (according to the hypothesis).

1. We are going to check the influence of **Class** to the possibility of attaching suffix -txe-. We use standard glm() function with parameter "family" specified to "binomial".By using function summary() we obtain the results of our model:

```
model.cl <- glm(isSuffix~Class, data = ourdata, family = "binomial")
summary(model.cl)
```

```
##
## Call:
## glm(formula = isSuffix ~ Class, family = "binomial", data = ourdata)
##
## Deviance Residuals:
##       Min        1Q    Median        3Q       Max
## -1.05794  -0.54387  -0.43660  -0.00008   2.18993
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -19.57    2195.15  -0.009    0.993
## Classkin         18.78    2195.15   0.009    0.993
## Classpron        17.26    2195.15   0.008    0.994
## Classpropern     19.28    2195.15   0.009    0.993
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 69.548  on 75  degrees of freedom
## Residual deviance: 52.400  on 72  degrees of freedom
## AIC: 60.4
##
## Number of Fisher Scoring iterations: 18
```

2. Now we are going to use **Number** as a predictor.

```
model.nu <- glm(isSuffix~Number, data = ourdata, family = "binomial")
summary(model.nu)
```

```
##
## Call:
## glm(formula = isSuffix ~ Number, family = "binomial", data = ourdata)
##
## Deviance Residuals:
##       Min        1Q    Median        3Q       Max
## -0.88661  -0.88661  -0.00008  -0.00008   1.49929
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -19.57    1792.34  -0.011    0.991
## Numbersg         18.84    1792.34   0.011    0.992
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 69.548  on 75  degrees of freedom
## Residual deviance: 50.446  on 74  degrees of freedom
## AIC: 54.446
##
## Number of Fisher Scoring iterations: 18
```

3. Finally, we are going to check **Animacy** influence.

```
model.an <- glm(isSuffix~Animacy, data = ourdata, family = "binomial")
summary(model.an)
```

```
##
## Call:
## glm(formula = isSuffix ~ Animacy, family = "binomial", data = ourdata)
##
## Deviance Residuals:
##       Min        1Q    Median        3Q       Max
## -0.77754  -0.77754  -0.38499  -0.00013   2.29741
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -1.0415     0.3358  -3.102  0.00192 **
## Animacyinanim   -17.5246  1630.6597  -0.011  0.99143
## Animacynhum      -1.5235     1.0907  -1.397  0.16248
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 69.548  on 75  degrees of freedom
## Residual deviance: 60.010  on 73  degrees of freedom
## AIC: 66.01
##
## Number of Fisher Scoring iterations: 17
```

4. Additionally, we decided to combine **all** predictors listed above:

```
model.all <- glm(isSuffix~Number+Animacy+Class, data = ourdata, family = "binomial
")
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
summary(model.all)
```

file:///Users/olyashevskaya/Documents/!Vyshka/DataAnalysis/2018/…avtseva_Associative%20plural%20suffix%20-txe-%20in%20Adyghe.html

Page 7 of 9

```
##
## Call:
## glm(formula = isSuffix ~ Number + Animacy + Class, family = "binomial",
##     data = ourdata)
##
## Deviance Residuals:
##       Min        1Q     Median        3Q        Max
## -1.40059  -0.00003    0.00000    0.00000    1.35373
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -80.27   17973.81  -0.004    0.996
## Numbersg          58.90   13502.38   0.004    0.997
## Animacyinanim    -58.70   16660.64  -0.004    0.997
## Animacynhum      -19.24    6462.23  -0.003    0.998
## Classkin          21.88   11863.54   0.002    0.999
## Classpron         20.97   11863.54   0.002    0.999
## Classpropern      59.34   15254.00   0.004    0.997
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 69.548  on 75  degrees of freedom
## Residual deviance: 17.315  on 69  degrees of freedom
## AIC: 31.315
##
## Number of Fisher Scoring iterations: 22
```

Result Interpretation and prediction

Now we can analyze the fitting and interpret what the model is telling us. We can see that only Animacy are statistically significant. It has the lowest p-value suggesting a strong association of the animacy of the focal referent with the probability of having the form of the associative plural.

Finally, we may predict the order of the largest probability to form the associative plural.

```
ourdata %>%
count(Animacy, isSuffix) %>%
select(-n, -isSuffix) %>%
unique() ->
new.data
new.data %>%
predict(model.an, newdata = ., type = "response") ->
new.data$prediction
ura <- new.data %>%
arrange(desc(prediction))
ura
```

```
## Source: local data frame [3 x 2]
## Groups: Animacy [3]
##
##   Animacy   prediction
##    <fctr>        <dbl>
## 1     hum 2.608696e-01
## 2    nhum 7.142857e-02
## 3  inanim 8.646869e-09
```

As we see from the model results, our hypothesis was partially correct: the probability to attach the associative plural suffix -txe- in Adyghe depends on the Animacy of the focal referent, but there is no sifnificant correlation with the Number.

Sources

- Даниэль, М.А. Типология ассоциативной множественности. М., 2000.
- Аспекты полисинтетизма: очерки по грамматике адыгейского языка. М., 2009.
- Silverstein, Michael. 1976. Hierarchy of Features and Ergativity. In Dixon, R. M. W. (ed.), Grammatical Categories in Australian Languages, 112-171. Canberra: Australian National University.
- RStudio. Version 1.0.136 – © 2009-2016 RStudio, Inc.
- tidyverse-package {tidyverse}
- Data: https://vk.com/doc383823376_447062479 (https://vk.com/doc383823376_447062479)