

Локальные дейктики и жестикуляция

2023-06-06

Локальные дейктики и жестикуляция

О проекте

Проект посвящен изучению связи вербального и кинетического каналов коммуникации и их участия в указании.

Цель проекта - исследовать данную связь статистически.

Для проекта был выбран датасет SaGA (подробнее в Lücking et al., 2010; Kucherenko, 2021). Он мультимодален, состоит из 25 видео, в которых производится указание пути одним участником другому. Язык датасета - немецкий.

Датасет доступен по ссылке.

Источники:

1. Lücking, A., Bergmann, K., Hahn, F., Kopp, S., & Rieser, H. (2010). *The Bielefeld Speech and Gesture Alignment Corpus (SaGA)*. In M. Kipp, J. - P. Martin, P. Paggio, & D. Heylen (Eds.), LREC 2010 Workshop: Multimodal Corpora—Advances in Capturing, Coding and Analyzing Multimodality (pp. 92-98).
2. Kucherenko, T., Nagy, R., Neff, M., Kjellström, H., & Henter, G. E. (2021). *Multimodal analysis of the predictability of hand-gesture properties*. arXiv preprint arXiv:2108.05762.

Данные

Исходный датасет включает в себя 1) файлы с аннотациями к видео (.eaf), 2) файлы с аудиальными признаками (.пру), 3) файлы с 3д-моделированием кинетического поведения говорящих (.bvh). Для исследования были взяты только первые файлы, из которых с помощью программы ELAN были получены аннотации нужных слоёв в текстовом виде. Полученный файл далее обрабатывался с помощью языка программирования Python.

Аннотация датасета включает в себя транскрипт речи говорящих (дающего указания и им следующего), лемматизированный и размеченный частеречно, информацию о различных характеристиках жестов отдельно для каждой из рук и также их вместе.

Процесс обработки данных описан в скрипте `preprocessing_data.py`. Он заключался в сопоставлении аннотаций с разных слоёв по времени для дальнейшего создания датасета, содержащего произносимые слова (а также их леммы и часть речи) и сопоставленные им жестовые характеристики (тип жеста, фаза жеста, семантика жеста и др. характеристики). Далее для каждой леммы была проведена дополнительная разметка, является ли лексема дейтиком. Для этого был составлен список немецких дейктических единиц, исходя из работы Blühdm (1994), а также ресурсов 2 и 3. Заметим, что отбирались только единицы, которые относятся к категориям *Zeichendeiktika* и *Lokaldeiktika* у Blühdm'a, поскольку связь между жестами и местными дейктиками наиболее сильна, а также из-за специфики материала, фокусирующегося на указании на местности.

Источники:

1. Blühdorn, H. (1994). *Deixis und Deiktika in der deutschen Gegenwartssprache*. Deutsche Sprache, 21(1), 44-62.
2. *Präpositionaladverb*.
3. *Deiktisches Adverb*.

Подгружаем данные...

```
data <- read.csv("https://raw.githubusercontent.com/anna-a-m/andan-project/main/data.csv")
```

Анализ статистики

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.2      v readr      2.1.4
## v forcats    1.0.0      v stringr   1.5.0
## v ggplot2    3.4.2      v tibble    3.2.1
## v lubridate  1.9.2      v tidyr     1.3.0
## v purrr      1.0.1
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
str(data)
```

```
## 'data.frame':    90776 obs. of  11 variables:
## $ Form          : chr  "ja" "also" "ich" "bin" ...
## $ Lemma         : chr  "ja" "also" "ich" "sein" ...
## $ Pos           : chr  "PTKANT" "ADV" "PPER" "VAFIN" ...
## $ Phrase        : chr  "" "" "" "" ...
## $ Hand          : chr  "" "" "" "" ...
## $ Phase         : chr  "" "" "" "" ...
## $ Practice      : chr  "" "" "" "" ...
## $ Semantic      : chr  "" "" "" "" ...
## $ HandShapeShape: chr  "" "" "" "" ...
## $ PalmDirection : chr  "" "" "" "" ...
## $ IsDeictic     : chr  "non-deictic" "non-deictic" "non-deictic" "non-deictic" ...
```

В полученном датасете **90,776** наблюдений и 11 признаков, первые три из которых описывают речевую составляющую, следующие 7 жестовую. Последний же признак - это наша целевая переменная. Далее посмотрим на её распределение в датасете.

```
data %>%
  count(IsDeictic, sort = TRUE, name = "count") %>%
  mutate(part = round(count/nrow(data), digits=2))
```

```
##      IsDeictic count part
## 1 non-deictic 82465 0.91
## 2      deictic  8311 0.09
```

Целевая переменная, дейктик ли лемма, распределена крайне неравномерно. Она составляет только 10-ю часть всего датасета.

Посмотрим, сколько уникальных значений имеют жестовые признаки и каковы эти значения.

```
data[, 5:length(data)-1] %>%
  summarise_each(list(n_distinct)) %>%
  pivot_longer(cols=Phrase:Palmdirection, names_to="Variable", values_to="Distinct_count") %>%
  arrange(Distinct_count)
```

```
## Warning: `summarise_each()` was deprecated in dplyr 0.7.0.
## i Please use `across()` instead.
## i The deprecated feature was likely used in the dplyr package.
## Please report the issue at <https://github.com/tidyverse/dplyr/issues>.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

```
## # A tibble: 7 x 2
##   Variable      Distinct_count
##   <chr>          <int>
## 1 Hand              3
## 2 Phase              6
## 3 Semantic          13
## 4 Phrase            14
## 5 Practice          25
## 6 Palmdirection     56
## 7 HandShapeShape    141
```

```
data %>%
  filter(Phase != "") %>%
  distinct(Phase, IsDeictic) %>%
  arrange(IsDeictic)
```

```
##      Phase  IsDeictic
## 1    prep    deictic
## 2   stroke    deictic
## 3 post.hold    deictic
## 4     retr    deictic
## 5 pre.hold    deictic
## 6    prep non-deictic
## 7   stroke non-deictic
## 8 post.hold non-deictic
## 9     retr non-deictic
## 10 pre.hold non-deictic
```

Как видно, по признаку фазы жеста, что дейктики и не-дейктики не различаются. Это логично, поскольку само протекание жеста не должно зависеть от производимой вербальной составляющей. Так, у жеста есть три основных фазы: экскурсия, выдержка и ретракция (примерно так же, как и у артикуляции звука в акустике). Также жест может быть прерван или задержан в фазе экскурсии или ретракции.

```
data %>%
  filter(Phrase != "") %>%
  distinct(Phrase, IsDeictic) %>%
  arrange(IsDeictic)
```

```
##      Phrase  IsDeictic
## 1    iconic    deictic
## 2      move    deictic
## 3    deictic    deictic
## 4 beat-iconic    deictic
```

```

## 5      discourse      deictic
## 6      deictic-iconic  deictic
## 7      beat-deictic    deictic
## 8      beat            deictic
## 9      beat-deictic-iconic deictic
## 10     beat-discourse   deictic
## 11     iconic non-deictic
## 12     beat-iconic non-deictic
## 13     discourse non-deictic
## 14     move non-deictic
## 15     beat non-deictic
## 16     deictic non-deictic
## 17     beat-deictic non-deictic
## 18     deictic-iconic non-deictic
## 19     beat-deictic-iconic non-deictic
## 20     beat-discourse non-deictic
## 21     deictic-discourse non-deictic
## 22     discourse-iconic non-deictic
## 23     discourse-indexing non-deictic

```

В датасете присутствуют достаточно много разных типов жестов как для дейктиков, так и для не-дейктиков. Интересно, что и не-дейктики часто сопровождаются указательными жестами. Посмотрим, какие именно это слова.

```

data %>%
  select(Lemma, Phrase, IsDeictic) %>%
  filter(IsDeictic == 'non-deictic') %>%
  filter(stringr::str_detect(Phrase, 'deictic')) %>%
  distinct(Lemma) %>%
  sample_n(30)

```

```

##      Lemma
## 1      Gras
## 2     Hufeisen
## 3     gerade
## 4     derselbe
## 5     jeweils
## 6     einmal
## 7     Pforte
## 8     irgend
## 9     Anfang
## 10     sonst
## 11     spitze
## 12 eigentlich
## 13     Haken
## 14     Eingang
## 15     Podest
## 16     Treppe
## 17     nachdem
## 18     immerhin
## 19     halb
## 20 Startpunkt
## 21 einfassen
## 22 irgendwie
## 23     erstmal

```

```
## 24      weit
## 25      zwei
## 26  sozusagen
## 27      Tür
## 28 rausgucken
## 29      paar
## 30      lieb
```

Указательные жесты сопровождаются

- неразмеченными в датасете дейктиками (поскольку либо не являются локальными, либо не входят в выбранную классификацию),
dies, jen, dann
- местные предлоги,
hinterm, über
- наименования мест,
Kirche, Park, See
- глаголы действия,
stehen, biegen
- междометия и дискурсивные слова
quasi, aeh

Сбалансируем наш датасет, убрав из него те слова, которые сопровождаются указательными жестами, чтобы сделать более явной разницу между предметом исследования, локальными дейктиками, и другими вербальными единицами. При этом в таком случае будем считать признак `Phrase` менее значимым для наблюдений.

```
non_deictics <- data %>%
  filter(IsDeictic == 'non-deictic') %>%
  filter(!stringr::str_detect(Phrase, 'deictic'))

non_deictics <- non_deictics[sample(1:nrow(non_deictics)), ]
non_deictics <- non_deictics %>%
  drop_na()
non_deictics <- non_deictics[1:8311,]
```

```
deictics <- data %>%
  filter(IsDeictic == 'deictic')

final_df <- bind_rows(deictics, non_deictics)
final_df <- final_df %>%
  mutate(across(where(is.character), as.factor))

final_df %>%
  count(IsDeictic, name="count")
```

```
##      IsDeictic count
## 1      deictic  8311
## 2 non-deictic  8311
```

Визуализации

График-мозаика распределения признака Phrase

Посмотрим, как теперь распределён наш признак Phrase. Он представляет собой категоризацию жеста по типу. Жесты бывают иллюстративные (когда мы словно изображаем вербальную составляющую), ритмические (они поддерживают ритмическую структуру речи, помогают сохранить темп речи), дискурсивные, указательные, а также их сочетания.

```
library(vcd)
```

```
## Loading required package: grid
```

```
mosaic(IsDeictic ~ Phrase, data=final_df,  
       main = "Phrase feature distribution",  
       direction=c("h","v"), spacing=spacing_increase(),  
       gp = gpar(fill=c('lightblue', 'lightgreen')))
```

The treemap visualizes the distribution of phrases across deictic and non-deictic categories. The y-axis lists the phrases, and the x-axis shows the 'IsDeictic' status. The size of each rectangle represents the frequency of that phrase in that category.

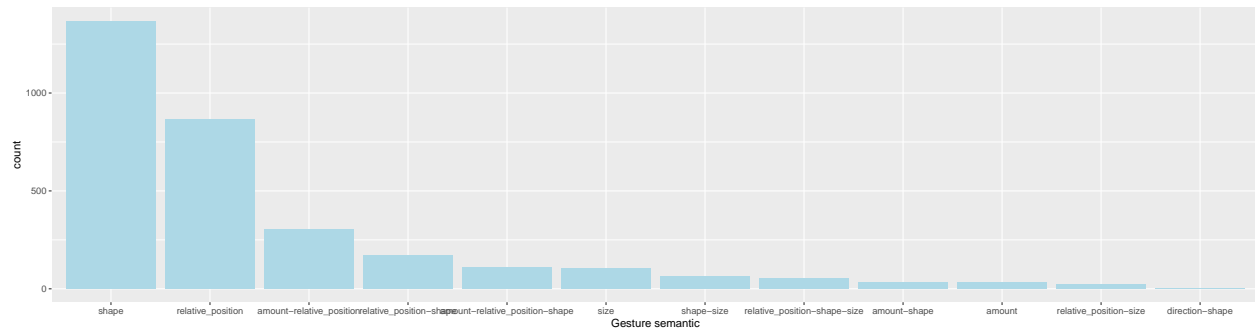
Phrase	IsDeictic	Frequency (approx.)
move	deictic	15
move	non-deictic	15
iconic	deictic	45
iconic	non-deictic	0
disjunctive	deictic	45
disjunctive	non-deictic	0
disjunctive-deictic	deictic	15
disjunctive-deictic	non-deictic	35
deictic	deictic	45
deictic	non-deictic	0
beat	deictic	15
beat	non-deictic	35
beat-deictic	deictic	15
beat-deictic	non-deictic	0
beat-non-deictic	deictic	0
beat-non-deictic	non-deictic	15

Отметим, что в целом иллюстративные жесты очень распространены. Также часто используются указательные жесты. Где-то в пятой части случаев жесты не используются (первая “строка” мозаики), причём при не-дейктиках чаще, чем при них.

Под семантикой жеста в данном датасете понимаются абстрактные категории, которые уточняются жестом - размер объекта в пространстве, его относительная позиция, форма или количество, а также совмещение данных признаков.

Поскольку для многих жестов не определен данный признак, посмотрим только на ту часть, для которой он размечен.

```
semantic <- final_df %>%
  filter(Semantic != "")
ggplot(semantic, aes(x=reorder(Semantic, Semantic, function(x)-length(x)))) +
  geom_bar(fill='lightblue') + labs(x='Gesture semantic')
```



Наиболее часто жестами поясняется форма предмета, что коррелирует с данными о том, что иллюстративные жесты распространены в датасете. Примерно также часто (будем считать жесты с семантикой *относительная позиция* и их комбинации с другими вместе) жесты поясняют относительную позицию объекта относительно говорящего, что коррелирует с распространенностью указательных жестов. Значение *количество* и *направление* встречаются редко.

То есть получается, что при объяснении пути говорящему важнее объяснить ориентиры и их положение относительно себя, нежели направление движения.

Множественный анализ соответствий для признака Practice

Признак Practice представляет собой некие способы представления объекта с помощью жеста. Всего таких способов в датасете выделено 8 таких способов (по оригинальной статье Lücking et al., 2010):

1. *Indexing* - указание на точку пространства
2. *Placing* - объект оказывается фигурально помещен или поставлен на пол
3. *Shaping* - форма объекта словно лепится говорящим в воздухе
4. *Drawing* - в воздухе рисуется форма объекта, в частности границы
5. *Posturing (modelling)* - руки образуют статичную конфигурацию в форме объекта
6. *Sizing* - указание на расстояние до объекта или его размер
7. *Counting* - счет на пальцах
8. *Hedging* - фигуральный жест для чего-то неопределенного

Попробуем провести MAC для данного категориального признака и посмотреть как он распределен относительно нашей целевой переменной.

```
library(factoextra)

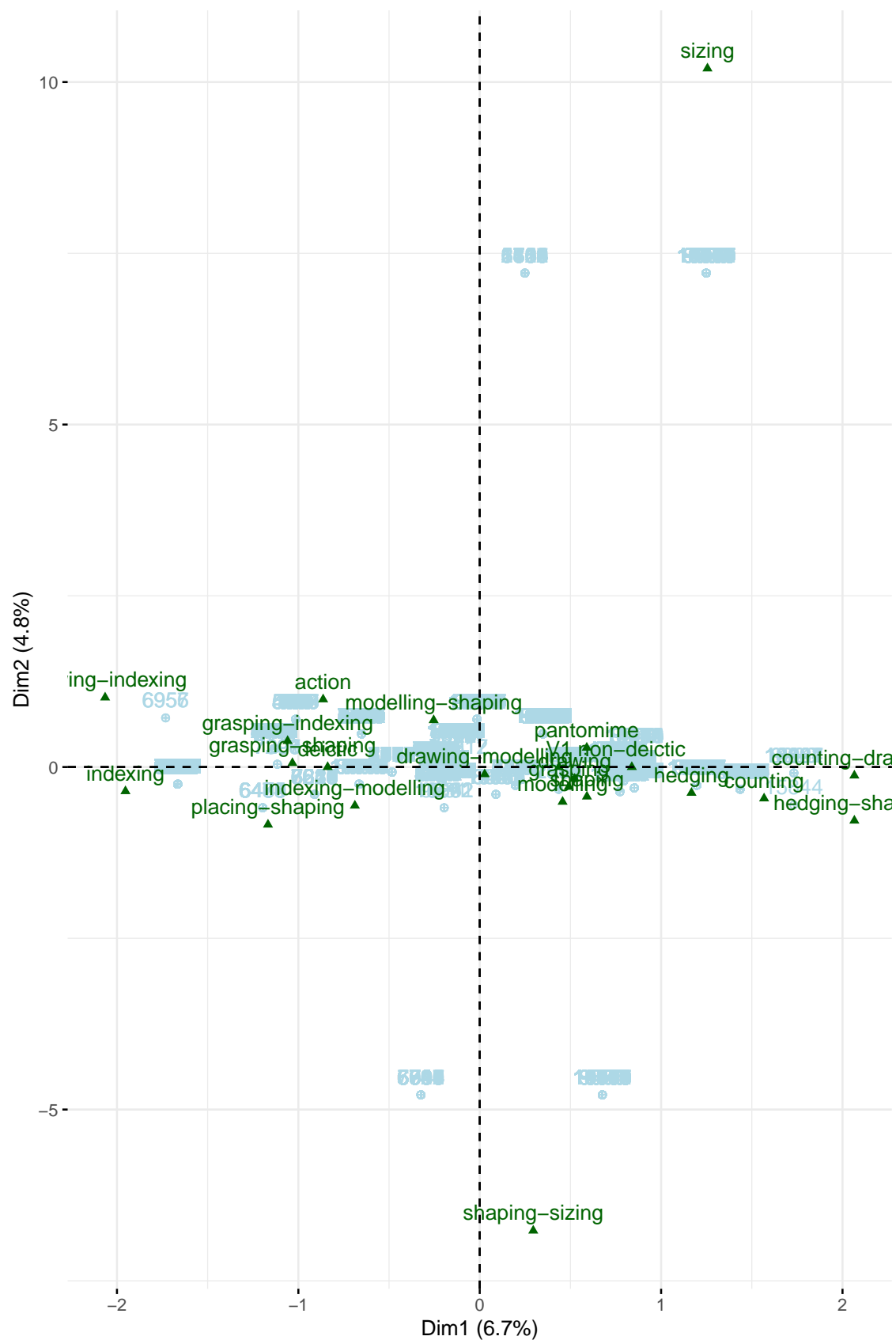
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
library(FactoMineR)

gesture_features <- final_df %>%
  select(Practice, IsDeictic)
```



```
gesture_features_mca <- MCA(gesture_features,  
                             ncp = 2,  
                             graph = FALSE)  
  
fviz_mca_biplot(gesture_features_mca,  
                 col.ind = "lightblue",  
                 col.var = "darkgreen",  
                 shape.ind = 10,  
                 ggtheme = theme_minimal())
```

MCA – Biplot



Как видно, в датасете в отличие от оригинальной статьи названия данной категории несколько отличаются (*posturing=modelling*), также добавлены некоторые другие категории типа *action*, *grasping*, *pantomime*. По-видимому, разметчики решили расширить данную категорию.

На картинке дейктики и не-дейктики оказываются разделены осью Оу. Соответственно будем считать, что те способы представления объекта, что характерны для обоих типов вербальных единиц, находятся в окрестности нуля. Здесь можно сделать вывод, что для жестов в целом характерна описательность (вспомним, что большая часть наших жестов иллюстративная), которая включает в себя задание формы предмета (*drawing | shaping | modelling*) и определение его положения (*indexing*).

Различия между дейктиками и не-дейктиками может заключаться в том¹, что с не-дейктиками чаще используются жесты “размера”, “неопределенности” и “счета”. Для дейктиков более характерны жесты с указанием точки пространства (*indexing*), а также комбинация указания формы объекта и помещения его в некоторое пространство.

¹ - вспомним, что используемый датасет был специально уменьшен, поэтому тут возможен перекося

Статистика и модель

Проверим, связаны ли такие жестовые признаки как *HandShapeShape* (форма жеста) и *PalmDirection* (направление ладони) с нашей целевой переменной с помощью статистического теста Фишера.

Для начала отметим, что форма жеста была размечена согласно стандартам американского жестого языка. Для разметки были использованы буквы, цифры и слова типа *маленький*, *ребёнок*. Направление же жеста размечалось аббревиатурами, включающими в себя а) движется ли рука ладонью вверх или вниз, б) направление движения вперед-назад, вправо-влево или вверх-вниз.

Сначала составим гипотезы:

H₀: жесты, сопровождающие дейктики, не отличаются по своей форме / направлению движения ладони.

H₁: жесты, сопровождающие дейктики, отличаются по своей форме / направлению движения ладони.

```
handshape_corr <- fisher.test(final_df$HandShapeShape, final_df$IsDeictic, workspace=2e9, simulate.p.val  
handshape_corr
```

```
##  
## Fisher's Exact Test for Count Data with simulated p-value (based on  
## 2000 replicates)  
##  
## data: final_df$HandShapeShape and final_df$IsDeictic  
## p-value = 0.0004998  
## alternative hypothesis: two.sided
```

```
palmdirection_corr <- fisher.test(final_df$PalmDirection, final_df$IsDeictic, workspace=2e9, simulate.p.val  
palmdirection_corr
```

```
##  
## Fisher's Exact Test for Count Data with simulated p-value (based on  
## 2000 replicates)  
##  
## data: final_df$PalmDirection and final_df$IsDeictic  
## p-value = 0.0004998  
## alternative hypothesis: two.sided
```

Поскольку значение p-value меньше, чем **0.05** в обоих статистических тестах, то будем считать, что данные признаки статистически связаны с нашей целевой переменной.

Учитывая выводы предыдущих разделов, можно предположить, что поскольку для дейктиков и не-дейктиков свойственны разные типы жестов, то и форма жеста, и направление ладони будут различаться. Так, для

типичного указательного жеста будет свойственна форма 5 (см. картинку выше) и направление ладони (вперед).

В качестве статистической модели была выбрана логистическая регрессия, поскольку целевая переменная бинарна. Для предсказания будем использовать только жестовые признаки *Practice* и *Semantic*, поскольку по предыдущим наблюдениям они являются важными в определении разницы между дейктиками и недейктиками, а также имеют приемлемое количество значений внутри категории, в отличие от *HandShapeShape* и *PalmDirection*.

```
library(stats)

gest.glm <- glm(IsDeictic ~ Semantic, final_df, family = binomial)
summary(gest.glm)

##
## Call:
## glm(formula = IsDeictic ~ Semantic, family = binomial, data = final_df)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.08000    0.01724  -4.641 3.47e-06 ***
## Semanticamount    1.54633    0.45324   3.412 0.000645 ***
## Semanticamount-relative_position    0.30314    0.11633   2.606 0.009164 **
## Semanticamount-relative_position-shape    0.36768    0.19172   1.918 0.055135 .
## Semanticamount-shape    1.21943    0.40657   2.999 0.002706 **
## Semanticdirection-shape   -0.32547    0.91303  -0.356 0.721488
## Semanticrelative_position   -0.77006    0.07621 -10.104 < 2e-16 ***
## Semanticrelative_position-shape    0.81759    0.16485   4.960 7.06e-07 ***
## Semanticrelative_position-shape-size    0.63961    0.28084   2.278 0.022755 *
## Semanticrelative_position-size    0.16004    0.40069   0.399 0.689595
## Semanticshape    1.03915    0.06282  16.541 < 2e-16 ***
## Semanticshape-size    1.42707    0.31180   4.577 4.72e-06 ***
## Semanticsize    1.95180    0.28760   6.786 1.15e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 23043  on 16621  degrees of freedom
## Residual deviance: 22455  on 16609  degrees of freedom
## AIC: 22481
##
## Number of Fisher Scoring iterations: 4
```

Отметим, что значение признака *Semantic direction-shape* сильно путает модель и с ним достаточно высокое значение ошибки. Возможно, это потому что он крайне редко встречается в датасете.

Модель, обученная только на признаке *Semantic*, имеет достаточно высокое значение AIC. Попробуем сравнить его с моделью, в которой будет учтён также признак *Practice*.

```
gest.glm2 <- glm(IsDeictic ~ Practice + Semantic, family=binomial, data=final_df)
gest.glm2.bw <- step(gest.glm2, direction = "backward")
```

```
## Start:  AIC=19709.47
## IsDeictic ~ Practice + Semantic
##
##              Df Deviance   AIC
```

```
## <none>          19644 19710
## - Semantic 12    19812 19854
## - Practice 20    22455 22481
```

Для модели с двумя признаками видно, что признак *Semantic* чуть более информативен, нежели *Practice*, поскольку его AIC сравнительно ниже. Но возможно, что дело лишь в количестве значений признака. Теперь посмотрим, какие параметры оказались самыми влиятельными.

```
library(caret)
```

```
## Loading required package: lattice
```

```
##
```

```
## Attaching package: 'caret'
```

```
## The following object is masked from 'package:purrr':
```

```
##
```

```
## lift
```

```
varImp(gest.glm2) %>%
  arrange(desc(Overall))
```

```
##                                Overall
## Practiceindexing               27.54755596
## Practicegrasping-indexing      17.21898317
## Practicemodelling-shaping      13.54156055
## Semanticshape                  9.80726685
## Practiceaction                 9.21447878
## Semanticamount-relative_position 5.44850700
## Semanticsize                   4.94750762
## Practicehedging                4.20199761
## Practiceindexing-modelling      3.72786200
## Practicecounting               3.71038691
## Practiceplacing-shaping         3.37049924
## Practicedrawing-modelling       3.34519393
## Semanticshape-size             3.22375960
## Semanticrelative_position-shape 2.40496231
## Practicesizing                 2.30771823
## Practiceshaping-sizing          2.20402999
## Practicegrasping-shaping        1.86666202
## Practicegrasping               1.85100607
## Semanticamount-shape            1.83614700
## Practicedrawing                1.48098822
## Practicemodelling              1.44333514
## Semanticamount                 1.27538639
## Semanticdirection-shape         0.88179846
## Semanticrelative_position-shape-size 0.81801319
## Semanticrelative_position       0.72370870
## Semanticrelative_position-size   0.70596144
## Practiceshaping                 0.51112690
## Practicepantomime              0.49692992
## Practicecounting-drawing        0.10525764
## Practicedrawing-indexing        0.06967281
## Semanticamount-relative_position-shape 0.06014482
## Practicehedging-shaping         0.03738089
```

Среди параметров с наибольшим весом большая часть относится к признаку *Practice* - indexing, grasping-indexing, modelling-shaping, action. Если обратиться вновь к графику MAC, то станет ясно, что эти параметры

наиболее характерны и отличают дейктики.

Также кажется, что признак *Practice* все же важен для модели и не стоит его исключать из неё.

Выводы

- Жесты, сопровождающие дейктики и не-дейктики действительно могут различаться между собой в *форме* и *семантике*. В частности, для дейктиков характерны жесты с указанием точки пространства (*indexing*), а также комбинация указания формы объекта и помещения его в некоторое пространство.
- Не-дейктики сравнительно чаще сопровождаются жестами, описывающими размер объекта, некую неопределенность и счет предметов. Соответственно, они являются иллюстративными.
- Фаза жеста не определяет не-дейктики и дейктики.