Extracting exemplars and prototypes

Antti Arppe

December 20, 2013

1 Fitting a polytomous logistic regression model

Before carrying out the statistical analyses, we need to invoke the `polytomous` package to make it available within R, having installed the package earlier. As subsequent preliminary steps, we load in the `think` data frame.

```
> library(polytomous)
> data(think)
```

Next, we transform using the `multinomial2logical` function the multinomial predictors into dummy predictor variables , which are simply TRUE or FALSE for each categorical value of the selected multinomial variables. The `think` data frame has three extralinguistic variables that we have decided to exclude from the following statistical modeling, namely `Register`, `Section`, and `Author` (which are columns 25, 26, and 27 in the original data frame). The results of this conversion are stored in the data frame `think.logical`.

```
> names(think)
 [1] "Lexeme"           "Polarity"         "Voice"            "Mood"
 [5] "Person"           "Number"           "Covert"           "ClauseEquivalent"
 [9] "Agent"            "Patient"          "Manner"           "Time"
[13] "Modality1"        "Modality2"        "Source"           "Goal"
[17] "Quantity"         "Location"         "Duration"         "Frequency"
[21] "MetaComment"      "ReasonPurpose"    "Condition"        "CoordinatedVerb"
[25] "Register"         "Section"          "Author"

> think.logical <- multinomial2logical(data=think, outcome="Lexeme", variables=names(think)[2:24])
```

With logical predictor variables we need to exclude per each of the original multinomial predictors one variable value so that we will not end up with exact multicollinearity. Often, one ends up excluding those variable values (individual categories/classes) that can be considered the least surprising, or default values. In the case of the `think` data frame, such default variable values are often either designated as `None` (i.e. entirely absent) or `Other` (typically a lump category of the most infrequent variable values). For two multinomial variables, `ClauseEquivalent` and `Overt`, with only two fully complementary values in each case, one has to base the choice on one's domain knowledge, thus deciding to exclude `ClauseEquivalent.FiniteVerbChain` and `Overt.Overt` (explicit syntactic *Subject* argument which is not obligatory for FIRST and SECOND person verb forms in Finnish). We also need to leave out the first column with the

dependent outcome variable `Lexeme` (which has been retained in multinomial form. After the preceding exclusions, we are left with the following 45 logical predictor variables, which we want to incorporate into a formula for the prediction of the outcome`Lexeme` through a few intermediate stages. Finally, the resultant multivariate formula is stored in `think.formula`:

The names of all the dummy variables `think.logical` are the following (excluding the first column with the dependent outcome variable):

```
> names(think.logical)[-1]
 [1] "PolarityOther"                    "PolarityNegation"
 [3] "VoiceOther"                       "VoicePassive"
 [5] "MoodOther"                        "MoodConditional"
 [7] "MoodIndicative"                   "PersonNone"
 [9] "PersonFirst"                      "PersonSecond"
[11] "PersonThird"                      "NumberOther"
[13] "NumberPlural"                     "CovertOvert"
[15] "CovertCovert"                     "ClauseEquivalentFiniteVerbChain"
[17] "ClauseEquivalentClauseEquivalent" "AgentNone"
[19] "AgentGroup"                       "AgentIndividual"
[21] "PatientNone"                      "PatientAbstraction"
[23] "PatientActivity"                  "PatientCommunication"
[25] "PatientDirectQuote"               "Patientetta.CLAUSE"
[27] "PatientIndirectQuestion"          "PatientIndividualGroup"
[29] "PatientInfinitive"                "PatientParticiple"
[31] "MannerNone"                       "MannerAgreement"
[33] "MannerFrame"                      "MannerGeneric"
[35] "MannerJoint"                      "MannerNegative"
[37] "MannerPositive"                   "TimeNone"
[39] "TimeDefinite"                     "TimeIndefinite"
[41] "Modality1None"                    "Modality1Necessity"
[43] "Modality1Possibility"             "Modality2None"
[45] "Modality2Accidental"              "Modality2External"
[47] "Modality2Temporal"                "Modality2Volition"
[49] "SourceNone"                       "SourceSource"
[51] "GoalNone"                         "GoalGoal"
[53] "QuantityNone"                     "QuantityQuantity"
[55] "LocationNone"                     "LocationLocation"
[57] "DurationNone"                     "DurationDuration"
[59] "FrequencyNone"                    "FrequencyFrequency"
[61] "MetaCommentNone"                  "MetaCommentMetaComment"
[63] "ReasonPurposeNone"                "ReasonPurposeReasonPurpose"
[65] "ConditionNone"                    "ConditionCondition"
[67] "CoordinatedVerbNone"              "CoordinatedVerbCoordinatedVerb"
```

The names of the non-default dummy variables are the following:

```
> grep("(Other)|(None)|(FiniteVerbChain)|(Overt)",
+     names(think.logical)[-1],value=T,invert=T)
 [1] "PolarityNegation"                 "VoicePassive"
 [3] "MoodConditional"                  "MoodIndicative"
 [5] "PersonFirst"                      "PersonSecond"
 [7] "PersonThird"                      "NumberPlural"
 [9] "CovertCovert"                     "ClauseEquivalentClauseEquivalent"
[11] "AgentGroup"                       "AgentIndividual"
[13] "PatientAbstraction"               "PatientActivity"
[15] "PatientCommunication"             "PatientDirectQuote"
[17] "Patientetta.CLAUSE"               "PatientIndirectQuestion"
[19] "PatientIndividualGroup"           "PatientInfinitive"
[21] "PatientParticiple"                "MannerAgreement"
[23] "MannerFrame"                      "MannerGeneric"
[25] "MannerJoint"                      "MannerNegative"
[27] "MannerPositive"                   "TimeDefinite"
[29] "TimeIndefinite"                   "Modality1Necessity"
[31] "Modality1Possibility"             "Modality2Accidental"
[33] "Modality2External"                "Modality2Temporal"
```

```
[35] "Modality2Volition"                "SourceSource"
[37] "GoalGoal"                         "QuantityQuantity"
[39] "LocationLocation"                 "DurationDuration"
[41] "FrequencyFrequency"               "MetaCommentMetaComment"
[43] "ReasonPurposeReasonPurpose"       "ConditionCondition"
[45] "CoordinatedVerbCoordinatedVerb"
```

This character vector can then be transformed into a formula as follows:

```
> paste(grep("(Other)|(None)|(FiniteVerbChain)|(Overt)",
+     names(think.logical)[-1],value=T,invert=T),collapse="")
[1] "PolarityNegationVoicePassiveMoodConditionalMoodIndicativePersonFirstPersonSecondPersonThirdNumberPluralCovert
> paste("Lexeme",paste(grep("(Other)|(None)|(FiniteVerbChain)|(Overt)",
+     names(think.logical)[-1],value=T,invert=T),collapse=" + "),sep=" ~ ")
[1] "Lexeme ~ PolarityNegation + VoicePassive + MoodConditional + MoodIndicative + PersonFirst + PersonSecond + Pe
> as.formula(paste("Lexeme",
+     paste(grep("(Other)|(None)|(FiniteVerbChain)|(Overt)",
+     names(think.logical)[-1],value=T,invert=T),collapse=" + "),sep=" ~ "))
Lexeme ~ PolarityNegation + VoicePassive + MoodConditional +
    MoodIndicative + PersonFirst + PersonSecond + PersonThird +
    NumberPlural + CovertCovert + ClauseEquivalentClauseEquivalent +
    AgentGroup + AgentIndividual + PatientAbstraction + PatientActivity +
    PatientCommunication + PatientDirectQuote + Patientetta.CLAUSE +
    PatientIndirectQuestion + PatientIndividualGroup + PatientInfinitive +
    PatientParticiple + MannerAgreement + MannerFrame + MannerGeneric +
    MannerJoint + MannerNegative + MannerPositive + TimeDefinite +
    TimeIndefinite + Modality1Necessity + Modality1Possibility +
    Modality2Accidental + Modality2External + Modality2Temporal +
    Modality2Volition + SourceSource + GoalGoal + QuantityQuantity +
    LocationLocation + DurationDuration + FrequencyFrequency +
    MetaCommentMetaComment + ReasonPurposeReasonPurpose + ConditionCondition +
    CoordinatedVerbCoordinatedVerb
> think.formula <- as.formula(paste("Lexeme",
+     paste(grep("(Other)|(None)|(FiniteVerbChain)|(Overt)",
+     names(think.logical)[-1],value=T,invert=T),collapse=" + "),sep=" ~ "))
```

Now we can fit a polytomous logistic regression model, the result of which we assign to the data frame `think.polytomous`:

```
> think.polytomous <- polytomous(think.formula, data=think.logical)
> print(summary(think.polytomous),max.print=NA)
Formula:
Lexeme ~ PolarityNegation + VoicePassive + MoodConditional +
    MoodIndicative + PersonFirst + PersonSecond + PersonThird +
    NumberPlural + CovertCovert + ClauseEquivalentClauseEquivalent +
    AgentGroup + AgentIndividual + PatientAbstraction + PatientActivity +
    PatientCommunication + PatientDirectQuote + Patientetta.CLAUSE +
    PatientIndirectQuestion + PatientIndividualGroup + PatientInfinitive +
    PatientParticiple + MannerAgreement + MannerFrame + MannerGeneric +
    MannerJoint + MannerNegative + MannerPositive + TimeDefinite +
    TimeIndefinite + Modality1Necessity + Modality1Possibility +
    Modality2Accidental + Modality2External + Modality2Temporal +
    Modality2Volition + SourceSource + GoalGoal + QuantityQuantity +
    LocationLocation + DurationDuration + FrequencyFrequency +
    MetaCommentMetaComment + ReasonPurposeReasonPurpose + ConditionCondition +
    CoordinatedVerbCoordinatedVerb
Heuristic:
one.vs.rest

Odds:
                                  ajatella  harkita  miettia   pohtia
(Intercept)                          2.177   0.0619   0.2165    0.124
AgentGroupTRUE                      0.1952    (1.15)   0.5151    4.132
AgentIndividualTRUE                (0.8071) (0.6853)  (1.009)    1.665
ClauseEquivalentClauseEquivalentTRUE  (1.181)  (1.984)   0.5822 (0.8531)
```

```
ConditionConditionTRUE                  0.457    2.885  (1.228)  (0.564)
CoordinatedVerbCoordinatedVerbTRUE     0.4856 (0.8283)    2.314 (0.8375)
CovertCovertTRUE                      (1.098) (0.8045)    (1.2)  (0.746)
DurationDurationTRUE                    0.119  (1.038)    3.414   (1.26)
FrequencyFrequencyTRUE                  0.384  (1.749)    1.764 (0.7861)
GoalGoalTRUE                            3.611   0.2143 (0.5823) (0.5989)
LocationLocationTRUE                   0.2586    0.458 (0.9318)    3.675
MannerAgreementTRUE                     16.61  (1/Inf)   0.0707    0.225
MannerFrameTRUE                         2.577    0.277   0.2786  (1.166)
MannerGenericTRUE                        22.4  (1/Inf)    0.149  (1/Inf)
MannerJointTRUE                        0.3716  (1.512)    2.091 (0.7307)
MannerNegativeTRUE                      3.984 (0.5963) (0.5536)   0.2184
MannerPositiveTRUE                    (0.7149)    1.846  (0.974) (0.8232)
MetaCommentMetaCommentTRUE            (0.8361)    1.624  (1.053) (0.7936)
Modality1NecessityTRUE                   0.345  (1.449)    2.025 (0.9376)
Modality1PossibilityTRUE               (1.182)  (1.227)  (1.055) (0.8118)
Modality2AccidentalTRUE                   5.84  (1/Inf) (0.4212) (0.4715)
Modality2ExternalTRUE                   2.584 (0.8987) (0.7864) (0.7142)
Modality2TemporalTRUE                  0.2612   0.1477    1.766    2.353
Modality2VolitionTRUE                 (0.5982) (0.6531)  (1.467)  (1.229)
MoodConditionalTRUE                    (1.273)    2.323   0.5446 (0.6908)
MoodIndicativeTRUE                       1.993 (0.8237)  (0.672) (0.8055)
NumberPluralTRUE                       (1.154)  (1.211)    0.595    1.548
PatientAbstractionTRUE                  0.2369  (1.077)    1.511    4.383
PatientActivityTRUE                     0.1374    9.293 (0.7723)    1.744
PatientCommunicationTRUE                0.1081  (1.878)     2.68     3.07
PatientDirectQuoteTRUE                 0.00921  (1/Inf)    3.172    8.789
PatientIndirectQuestionTRUE            0.07091 (0.8492)    4.145    2.847
PatientIndividualGroupTRUE              2.747  (1.008)   0.3704   0.3345
PatientInfinitiveTRUE                   5.171  (1.486)  (1/Inf) (0.2253)
PatientParticipleTRUE                   5.738  (1.024)  (1/Inf) (0.3336)
Patientetta.CLAUSETRUE                  2.583   0.2616    0.509   0.5163
PersonFirstTRUE                       (0.8882)  (1.814)  (1.731)   0.2754
PersonSecondTRUE                      (0.7091) (0.6757)    2.368   0.4054
PersonThirdTRUE                        (0.654)  (1.504)  (1.262) (0.9667)
PolarityNegationTRUE                    2.136   (1.14)    0.705   0.4665
QuantityQuantityTRUE                  (0.6843)   0.3387    2.607 (0.7461)
ReasonPurposeReasonPurposeTRUE          0.4337  (1.617)  (1.066)  (1.239)
SourceSourceTRUE                        3.129 (0.1365)  (0.755)   0.2869
TimeDefiniteTRUE                        0.4039  (0.774) (0.9586)    2.254
TimeIndefiniteTRUE                      0.5772  (1.241)    1.508 (0.9791)
VoicePassiveTRUE                      (0.6373)  (1.101) (0.8802)     1.84

Null deviance:              8701  on  13616  degrees of freedom
Residual (model) deviance:  5981  on  13432  degrees of freedom

R2.likelihood:  0.3125
AIC:            6349
BIC:            7478
```

The polytomous logistic regression model is the basis of both extracting a set of exemplary exemplars as well as sets of properties (variable values) which are prototypical for each outcome (i.e. `Lexeme`).

2 Extracting exemplars and prototypes

Having fit the polytomous logistic regression model, we can use it as a basis for extracting individual contexts which are most exemplary of the various outcomes, using the function `extract.exemplars`. This function takes the dataset incorporated in the fitted result and applies *Hierarchical Cluster Analysis* on it with the function `hclust`, using `method="binary"` for distance measure and `method="ward"` for the distance-based clustering:

```
> extract.exemplars(think.polytomous, n.clusters=100)
     indices outcomes max.probs
899      899   pohtia 0.7522760
286      286   pohtia 0.7197809
514      514   pohtia 0.6736778
1648    1648  miettia 0.7680341
3233    3233 ajatella 0.9905042
3035    3035 ajatella 0.8284270
2637    2637 ajatella 0.8995969
1171    1171 ajatella 0.9813306
3162    3162   pohtia 0.6292819
1991    1991   pohtia 0.6205248
2422    2422 ajatella 0.9384550
121      121 ajatella 0.6648627
255      255  harkita 0.5131767
2015    2015 ajatella 0.9814404
2482    2482 ajatella 0.8895892
19        19 ajatella 0.9893241
1387    1387   pohtia 0.7458646
76        76 ajatella 0.3894004
3112    3112 ajatella 0.7909073
1856    1856 ajatella 0.9842436
1684    1684 ajatella 0.9238340
574      574 ajatella 0.8194565
1521    1521 ajatella 0.9874202
151      151 ajatella 0.9407626
558      558  miettia 0.8239422
1891    1891 ajatella 0.9737094
267      267 ajatella 0.9851833
521      521 ajatella 0.8863296
37        37 ajatella 0.9731117
3013    3013  miettia 0.8588000
1704    1704 ajatella 0.9279253
474      474   pohtia 0.7424814
2711    2711 ajatella 0.9806257
323      323 ajatella 0.9683607
1300    1300 ajatella 0.9257161
2081    2081 ajatella 0.9534887
1765    1765 ajatella 0.9705500
2922    2922 ajatella 0.9906011
230      230 ajatella 0.8357151
52        52   pohtia 0.7062646
2611    2611 ajatella 0.9803866
2328    2328 ajatella 0.9295380
2163    2163 ajatella 1.0000000
763      763  harkita 0.5832702
2458    2458 ajatella 0.9517910
2362    2362  miettia 0.8384125
207      207 ajatella 0.8357440
1015    1015 ajatella 0.9886027
1759    1759 ajatella 0.8985800
88        88 ajatella 0.9688293
2369    2369 ajatella 0.9375479
479      479 ajatella 0.9368280
166      166 ajatella 0.6846686
2254    2254 ajatella 0.9328456
540      540 ajatella 0.8310588
894      894 ajatella 0.8172398
310      310   pohtia 0.4478400
3278    3278 ajatella 0.9671088
3002    3002   pohtia 0.6181336
1943    1943  miettia 0.7404593
1754    1754  miettia 0.7539501
2168    2168 ajatella 0.9558096
133      133 ajatella 0.8449074
2680    2680 ajatella 0.9098211
2811    2811 ajatella 0.9837570
148      148 ajatella 0.6943053
```

```
2582    2582   pohtia 0.8507527
2211    2211 ajatella 0.8993147
482      482 ajatella 0.6688125
1753    1753 ajatella 1.0000000
655      655   pohtia 0.8493617
1579    1579   pohtia 0.6620441
1874    1874 ajatella 0.9719171
1613    1613 ajatella 0.5357194
2828    2828  harkita 0.5458486
180      180 ajatella 0.7531416
2585    2585 ajatella 0.6915454
695      695   pohtia 0.7891253
194      194 ajatella 0.9866129
204      204 ajatella 0.8225811
836      836   pohtia 0.6927355
2069    2069 ajatella 0.9731780
2122    2122 ajatella 0.9290423
1619    1619  miettia 0.8239539
907      907 ajatella 0.9815937
1687    1687 ajatella 0.8896726
313      313 ajatella 0.9198923
560      560 ajatella 0.8655428
1444    1444  miettia 0.3592241
818      818 ajatella 0.8619285
353      353 ajatella 0.5544356
1009    1009 ajatella 0.8867894
802      802  miettia 0.6701509
362      362  harkita 0.5212785
1161    1161 ajatella 0.8555719
397      397 ajatella 0.6337628
1997    1997 ajatella 0.9591542
2895    2895   pohtia 0.3726344
1261    1261 ajatella 0.8920517
2722    2722 ajatella 0.9472946

899                                                                    LocationLocation; Modality2Temporal; MoodIndicativ
286                                                                            AgentGroup; Modality2Temporal; MoodIndic
514                                                                  AgentIndividual; MoodIndicative; PatientDi
1648                                 AgentIndividual; CovertCovert; Modality1Necessity; MoodIndicative; PatientIndirectQues
3233                             AgentIndividual; CovertCovert; MannerAgreement; Modality1Possibility; MoodIndicative; Patientett
3035                       AgentIndividual; CovertCovert; FrequencyFrequency; Modality1Necessity; MoodIndicative; PatientInfini
2637                                          AgentIndividual; CovertCovert; GoalGoal; Modality2Accidental; MoodIndi
1171                                                                                                          MannerG
3162                                                          LocationLocation; Modality2Temporal; MoodIndicative; Pa
1991                                                       AgentGroup; MoodIndicative; NumberPlural; PatientIndirectQu
2422                                                            AgentIndividual; CovertCovert; MannerGeneric; MoodIndi
121                                                                                                    AgentIndividua
255                                                                                                    ClauseEquivale
2015                                      AgentIndividual; MannerAgreement; MetaCommentMetaComment; MoodIndicative; Numb
2482                                                                 AgentIndividual; MoodIndicative; Patientetta.Cl
19                                                                 AgentGroup; LocationLocation; MannerGeneric; MoodIn
1387                                                     AgentIndividual; LocationLocation; MannerPositive; MoodIndicativ
76                                                                                                   ClauseEquivalentC
3112 AgentIndividual; CovertCovert; DurationDuration; MannerGeneric; Modality1Necessity; MoodIndicative; PersonThi
1856                                                                                                MannerAgreement; MoodIn
1684                                                                      AgentIndividual; ClauseEquivalentClauseEquivalen
574                                                                         AgentIndividual; CovertCovert; NumberPlural; I
1521                                                                          LocationLocation; MannerGeneric; MoodIndicat
151                                                          AgentIndividual; CovertCovert; MoodIndicative; PatientPartic
558                                  AgentIndividual; CovertCovert; DurationDuration; MoodIndicative; PatientIndirectQues
1891                                                                 MetaCommentMetaComment; MoodIndicative; PatientPa
267                                                            AgentIndividual; CovertCovert; MannerAgreement; MoodCon
521                                                          AgentIndividual; CovertCovert; MoodIndicative; Patientetta.Cl
37                                     AgentIndividual; MannerGeneric; MetaCommentMetaComment; MoodIndicative; NumberPlu
3013                                 AgentIndividual; CoordinatedVerbCoordinatedVerb; CovertCovert; FrequencyFrequency; Pa
1704                               AgentIndividual; MetaCommentMetaComment; Modality2Accidental; MoodIndicative; NumberPl
474                                                          LocationLocation; MoodIndicative; PatientAbstraction; Reaso
2711                                                                   LocationLocation; MannerGeneric; MetaCommentMetaCo
```

Even though one could select manually out of the results of the fitted polytomous logistic regression model those properties that together form the abstract prototypes per each outcome `Lexeme`, we can use the convenience function `extra.prototypes` to directly get these, sorted in terms of decreasing odds for each outcome.

```
> extract.prototypes(think.polytomous)
$ajatella
                                      Odds
MannerGenericTRUE                22.402327677
MannerAgreementTRUE              16.612737398
Modality2AccidentalTRUE           5.840152352
PatientParticipleTRUE             5.738414552
PatientInfinitiveTRUE             5.171373435
MannerNegativeTRUE                3.983843673
GoalGoalTRUE                      3.611264449
SourceSourceTRUE                  3.129202530
PatientIndividualGroupTRUE        2.747133577
Modality2ExternalTRUE             2.583958234
Patientetta.CLAUSETRUE            2.583353952
MannerFrameTRUE                   2.577198926
(Intercept)                       2.176841163
PolarityNegationTRUE              2.136006848
MoodIndicativeTRUE                1.992772273
TimeIndefiniteTRUE                0.577161840
CoordinatedVerbCoordinatedVerbTRUE 0.485562629
ConditionConditionTRUE            0.457020168
ReasonPurposeReasonPurposeTRUE    0.433650016
TimeDefiniteTRUE                  0.403936351
FrequencyFrequencyTRUE            0.384048480
MannerJointTRUE                   0.371588626
Modality1NecessityTRUE            0.345026528
Modality2TemporalTRUE             0.261189628
LocationLocationTRUE              0.258618962
PatientAbstractionTRUE            0.236903291
AgentGroupTRUE                    0.195177806
PatientActivityTRUE               0.137437012
DurationDurationTRUE              0.118993611
PatientCommunicationTRUE          0.108109179
PatientIndirectQuestionTRUE       0.070908140
PatientDirectQuoteTRUE            0.009210107
$harkita
                             Odds
PatientActivityTRUE       9.29309572
ConditionConditionTRUE    2.88457153
MoodConditionalTRUE       2.32344848
MannerPositiveTRUE        1.84576867
MetaCommentMetaCommentTRUE 1.62356351
LocationLocationTRUE      0.45796396
QuantityQuantityTRUE      0.33866227
MannerFrameTRUE           0.27702202
Patientetta.CLAUSETRUE    0.26164699
GoalGoalTRUE              0.21433596
Modality2TemporalTRUE     0.14768111
(Intercept)               0.06189904

$miettia
                                      Odds
PatientIndirectQuestionTRUE       4.14463728
DurationDurationTRUE              3.41387969
PatientDirectQuoteTRUE            3.17224511
PatientCommunicationTRUE          2.68022996
QuantityQuantityTRUE              2.60735313
PersonSecondTRUE                  2.36765400
CoordinatedVerbCoordinatedVerbTRUE 2.31432650
MannerJointTRUE                   2.09148209
Modality1NecessityTRUE            2.02479917
```

```
Modality2TemporalTRUE                1.76610909
FrequencyFrequencyTRUE               1.76351066
PatientAbstractionTRUE               1.51096183
TimeIndefiniteTRUE                   1.50815992
PolarityNegationTRUE                 0.70499520
NumberPluralTRUE                     0.59504069
ClauseEquivalentClauseEquivalentTRUE 0.58219604
MoodConditionalTRUE                  0.54458882
AgentGroupTRUE                       0.51514830
Patientetta.CLAUSETRUE               0.50895779
PatientIndividualGroupTRUE           0.37043905
MannerFrameTRUE                      0.27858938
(Intercept)                          0.21650731
MannerGenericTRUE                    0.14896994
MannerAgreementTRUE                  0.07069893

$pohtia
                                 Odds
PatientDirectQuoteTRUE      8.7891740
PatientAbstractionTRUE      4.3827275
AgentGroupTRUE              4.1324530
LocationLocationTRUE        3.6754374
PatientCommunicationTRUE    3.0701864
PatientIndirectQuestionTRUE 2.8470324
Modality2TemporalTRUE       2.3534923
TimeDefiniteTRUE            2.2542571
VoicePassiveTRUE            1.8401364
PatientActivityTRUE         1.7441448
AgentIndividualTRUE         1.6646172
NumberPluralTRUE            1.5478831
Patientetta.CLAUSETRUE      0.5163393
PolarityNegationTRUE        0.4664671
PersonSecondTRUE            0.4054136
PatientIndividualGroupTRUE  0.3344601
SourceSourceTRUE            0.2869156
PersonFirstTRUE             0.2753850
MannerAgreementTRUE         0.2250194
MannerNegativeTRUE          0.2184115
(Intercept)                 0.1240282
```