Age of acquisition and borrowing

This analysis treats POS as a random effect.

TODO: why is freq and borrowing apparently positively correlated in the last graph? Something's wrong with the predicted values from the model, or the linear model is not capturing the data well. Need to add quadratic effect of frequency

TODO: scaling phon length should center on whole number.

Load libraries

```
library(lme4)
```

```
## Warning: package 'lme4' was built under R version 3.3.2
```

```
## Loading required package: Matrix
```

```
## Warning: package 'Matrix' was built under R version 3.3.2
```

```
library(sjPlot)
```

```
## Warning: package 'sjPlot' was built under R version 3.3.2
```

```
## Warning in checkMatrixPackageVersion(): Package version inconsistency detected.
## TMB was built with Matrix version 1.2.10
## Current Matrix version is 1.2.11
## Please re-install 'TMB' from source or restore original 'Matrix' package
```

```
## Learn more about sjPlot with 'browseVignettes("sjPlot")'.
```

```
library(lattice)
```

```
## Warning: package 'lattice' was built under R version 3.3.2
```

```
library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(party)
```

```
## Warning: package 'party' was built under R version 3.3.2
```

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

```
## Loading required package: mvtnorm
```

```
## Warning: package 'mvtnorm' was built under R version 3.3.2
```

```
## Loading required package: modeltools

## Loading required package: stats4

##
## Attaching package: 'modeltools'

## The following object is masked from 'package:lme4':
##
##     refit

## Loading required package: strucchange

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 3.3.2

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Loading required package: sandwich

## Warning: package 'sandwich' was built under R version 3.3.2
```

```r
logit2per = function(X){
  return(exp(X)/(1+exp(X)))
}
```

Load data

1 = definately borrowed 5 = no evidence of borrowing

```r
dataloan <- read.csv("../analysis/loanword8.csv",stringsAsFactors = F)
dataloan$bor15 <- ifelse(dataloan$borrowing==1,1, ifelse(dataloan$borrowing==5,0,NA))
dataloan$bor15.cat <- factor(dataloan$bor15)
```

Convert to numbers and scale.

```r
dataloan$subtlexzipf = as.numeric(dataloan$subtlexzipf)
```

```
## Warning: NAs introduced by coercion
```

```r
dataloan$AoA <- as.numeric(dataloan$AoA)
```

```
## Warning: NAs introduced by coercion
```

```r
aoaSD = sd(dataloan$AoA,na.rm = T)
aoaMean = mean(dataloan$AoA/aoaSD,na.rm=T)

dataloan$AoAscale <- scale(as.numeric(dataloan$AoA))
dataloan$subtlexzipfscale <- scale(as.numeric(dataloan$subtlexzipf))
dataloan$phonlengthscale <-
  dataloan$phonlength - median(dataloan$phonlength)
dataloan$concscale <- scale(as.numeric(dataloan$conc))
```

```
## Warning in scale(as.numeric(dataloan$conc)): NAs introduced by coercion
```
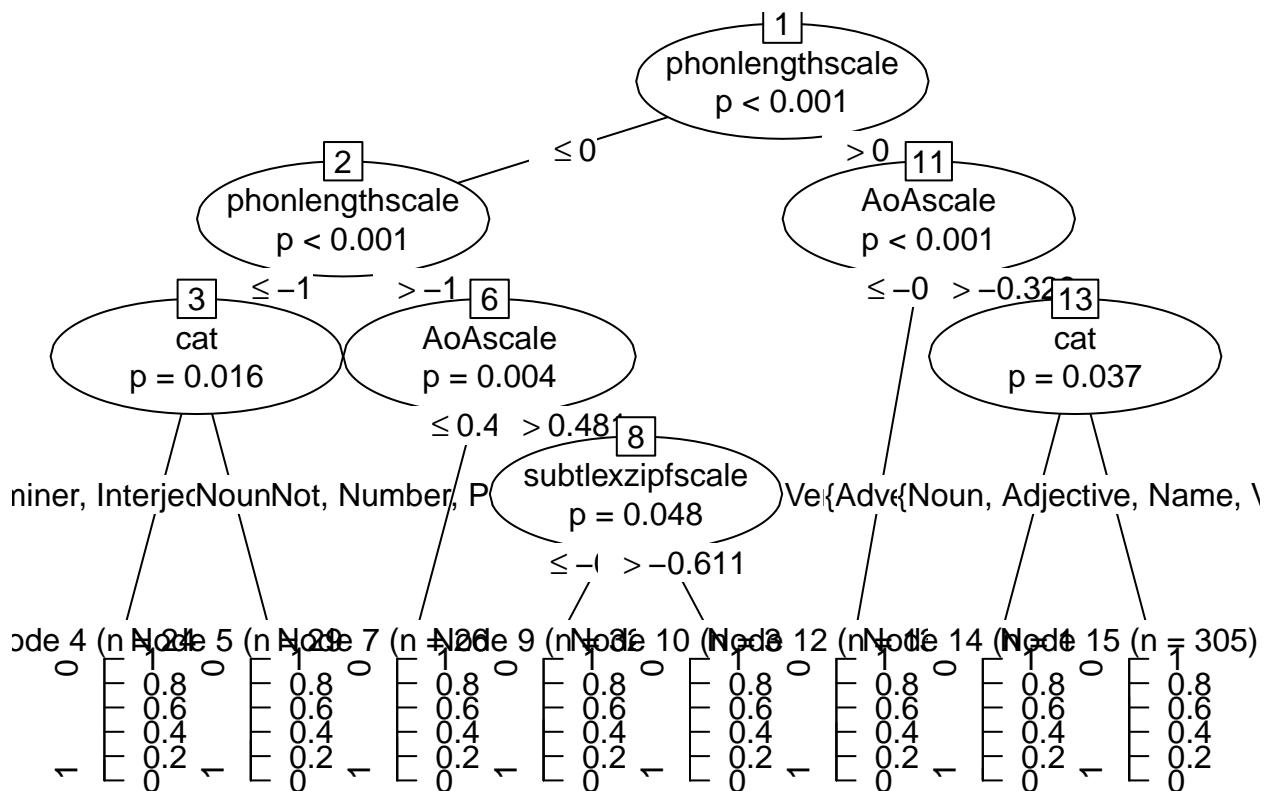
```
dataloan$cat = factor(dataloan$cat)
dataloan$cat = relevel(dataloan$cat,"Noun")
```

Select only complete cases.

```
dataloan2 = dataloan[complete.cases(dataloan[,
              c("phonlengthscale","AoAscale",
              "subtlexzipfscale", "cat",
              'concscale','bor15')]),]
```

Overall patterns

```
ctx = ctree(bor15.cat~cat+phonlengthscale+
          AoAscale + subtlexzipfscale+ concscale,
       dataloan2)
plot(ctx)
```


Random effect structure

Baseline model.

```
m0 = glmer(bor15.cat ~ 1 + (1 | cat),
           data=dataloan2, family = binomial)
```

Random slope for AoA

```
m0.1 = glmer(bor15.cat ~ 1 + (1 | cat) + (0 + AoAscale|cat),
            data=dataloan2,
            family = binomial)
anova(m0,m0.1)
```

```
## Data: dataloan2
## Models:
## m0: bor15.cat ~ 1 + (1 | cat)
## m0.1: bor15.cat ~ 1 + (1 | cat) + (0 + AoAscale | cat)
##      Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## m0    2 1726.1 1736.5 -861.05   1722.1
## m0.1  3 1682.5 1698.1 -838.27   1676.5 45.56      1   1.48e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Random slope for freq.

```
m0.2 = glmer(bor15.cat ~ 1 +
               (1 | cat) +
               (0 + AoAscale|cat) +
               (0 + subtlexzipfscale | cat),
            data=dataloan2, family = binomial)
anova(m0.1,m0.2)
```

```
## Data: dataloan2
## Models:
## m0.1: bor15.cat ~ 1 + (1 | cat) + (0 + AoAscale | cat)
## m0.2: bor15.cat ~ 1 + (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m0.2:     cat)
##      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## m0.1  3 1682.5 1698.1 -838.27   1676.5
## m0.2  4 1679.9 1700.6 -835.95   1671.9 4.6422      1     0.0312 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Random slope for length.

```
m0.3 = glmer(bor15.cat ~ 1 +
               (1 | cat) +
               (0 + AoAscale|cat) +
               (0 + subtlexzipfscale | cat) +
               (0 + phonlengthscale| cat),
            data=dataloan2, family = binomial)
anova(m0.2,m0.3)
```

```
## Data: dataloan2
## Models:
## m0.2: bor15.cat ~ 1 + (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m0.2:     cat)
## m0.3: bor15.cat ~ 1 + (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m0.3:     cat) + (0 + phonlengthscale | cat)
##      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## m0.2  4 1679.9 1700.6 -835.95   1671.9
## m0.3  5 1598.5 1624.4 -794.24   1588.5 83.415      1  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Random slope for concreteness?

```
m0.35 = glmer(bor15.cat ~ 1 + (1 | cat) +
                (0 + AoAscale|cat) +
                (0 + subtlexzipfscale | cat) +
                (0 + phonlengthscale| cat) +
                (0 + concscale | cat),
            data=dataloan2, family = binomial)
anova(m0.3,m0.35)
```

```
## Data: dataloan2
## Models:
## m0.3: bor15.cat ~ 1 + (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m0.3:     cat) + (0 + phonlengthscale | cat)
## m0.35: bor15.cat ~ 1 + (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m0.35:     cat) + (0 + phonlengthscale | cat) + (0 + concscale | cat)
##       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## m0.3   5 1598.5 1624.4 -794.24   1588.5
## m0.35  6 1600.5 1631.6 -794.23   1588.5 0.0256      1     0.8728
```

Not significant.

Check correlation parameters between random slopes

```
m0.4 = glmer(bor15.cat ~ 1 +
                (1 + AoAscale+
                    subtlexzipfscale +
                    phonlengthscale| cat),
            data=dataloan2, family = binomial)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : Model failed to converge with max|grad| = 0.0266914 (tol =
## 0.001, component 1)
```

```
anova(m0.3,m0.4)
```

```
## Data: dataloan2
## Models:
## m0.3: bor15.cat ~ 1 + (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m0.3:     cat) + (0 + phonlengthscale | cat)
## m0.4: bor15.cat ~ 1 + (1 + AoAscale + subtlexzipfscale + phonlengthscale |
## m0.4:     cat)
##       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## m0.3   5 1598.5 1624.4 -794.24   1588.5
## m0.4  11 1589.6 1646.6 -783.79   1567.6 20.919      6   0.001897 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

All are significant, but correlation parameters cause convergence issue and add lots of degrees of freedom. We'll leave them out. Anyway, as we see in the later analyses, the random slopes actually account for very little variation in the final models.

Fixed effects

Main effect of length

```
m1 = glmer(bor15.cat ~
            phonlengthscale +
            (1 | cat) +
            (0 + AoAscale|cat) +
            (0 + subtlexzipfscale | cat) +
            (0 + phonlengthscale| cat),
          data=dataloan2, family = binomial)
m1.5 = glmer(bor15.cat ~
            phonlengthscale +
            I(phonlengthscale^2) +
            (1 | cat) +
            (0 + AoAscale|cat) +
            (0 + subtlexzipfscale | cat) +
            (0 + phonlengthscale| cat),
          data=dataloan2, family = binomial)
```

Main effect of aoa and quadratic term.

```
m2 = glmer(bor15.cat ~
            phonlengthscale +
            AoAscale +
            (1 | cat) +
            (0 + AoAscale|cat) +
            (0 + subtlexzipfscale | cat) +
            (0 + phonlengthscale| cat),
          data=dataloan2, family = binomial)
m2.5 = glmer(bor15.cat ~
            phonlengthscale +
            AoAscale +
            I(AoAscale^2) +
            (1 | cat) +
            (0 + AoAscale|cat) +
            (0 + subtlexzipfscale | cat) +
            (0 + phonlengthscale| cat),
          data=dataloan2, family = binomial)
```

Main effect of freq

```
m3 = glmer(bor15.cat ~ phonlengthscale +
            AoAscale +
            subtlexzipfscale  +
            (1 | cat) +
            (0 + AoAscale|cat) +
            (0 + subtlexzipfscale | cat) +
            (0 + phonlengthscale| cat),
          data=dataloan2, family = binomial)
m3.5 = glmer(bor15.cat ~ phonlengthscale +
            AoAscale +
            subtlexzipfscale  +
            I(subtlexzipfscale^2)  +
            (1 | cat) +
            (0 + AoAscale|cat) +
            (0 + subtlexzipfscale | cat) +
            (0 + phonlengthscale| cat),
          data=dataloan2, family = binomial)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : Model failed to converge with max|grad| = 0.00352504 (tol =
## 0.001, component 1)
```

Test significance by model comparison:

```r
anova(m0.3,m1,m1.5)
```

```
## Data: dataloan2
## Models:
## m0.3: bor15.cat ~ 1 + (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m0.3:     cat) + (0 + phonlengthscale | cat)
## m1: bor15.cat ~ phonlengthscale + (1 | cat) + (0 + AoAscale | cat) +
## m1:     (0 + subtlexzipfscale | cat) + (0 + phonlengthscale | cat)
## m1.5: bor15.cat ~ phonlengthscale + I(phonlengthscale^2) + (1 | cat) +
## m1.5:     (0 + AoAscale | cat) + (0 + subtlexzipfscale | cat) + (0 +
## m1.5:     phonlengthscale | cat)
##      Df    AIC    BIC  logLik deviance   Chisq Chi Df Pr(>Chisq)
## m0.3  5 1598.5 1624.4 -794.24   1588.5
## m1    6 1588.1 1619.2 -788.07   1576.1 12.3583      1   0.000439 ***
## m1.5  7 1586.3 1622.6 -786.16   1572.3  3.8111      1   0.050915 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(m1,m2)
```

```
## Data: dataloan2
## Models:
## m1: bor15.cat ~ phonlengthscale + (1 | cat) + (0 + AoAscale | cat) +
## m1:     (0 + subtlexzipfscale | cat) + (0 + phonlengthscale | cat)
## m2: bor15.cat ~ phonlengthscale + AoAscale + (1 | cat) + (0 + AoAscale |
## m2:     cat) + (0 + subtlexzipfscale | cat) + (0 + phonlengthscale |
## m2:     cat)
##    Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## m1  6 1588.1 1619.2 -788.07   1576.1
## m2  7 1578.6 1614.9 -782.29   1564.6 11.557      1  0.0006751 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(m2,m2.5)
```

```
## Data: dataloan2
## Models:
## m2: bor15.cat ~ phonlengthscale + AoAscale + (1 | cat) + (0 + AoAscale |
## m2:     cat) + (0 + subtlexzipfscale | cat) + (0 + phonlengthscale |
## m2:     cat)
## m2.5: bor15.cat ~ phonlengthscale + AoAscale + I(AoAscale^2) + (1 |
## m2.5:     cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale | cat) +
## m2.5:     (0 + phonlengthscale | cat)
##      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## m2    7 1578.6 1614.9 -782.29   1564.6
## m2.5  8 1577.6 1619.1 -780.81   1561.6 2.9478      1     0.086 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(m2,m3,m3.5)
```

```
## Data: dataloan2
```

```
## Models:
## m2: bor15.cat ~ phonlengthscale + AoAscale + (1 | cat) + (0 + AoAscale |
## m2:     cat) + (0 + subtlexzipfscale | cat) + (0 + phonlengthscale |
## m2:     cat)
## m3: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + (1 |
## m3:     cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale | cat) +
## m3:     (0 + phonlengthscale | cat)
## m3.5: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.5:     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m3.5:     cat) + (0 + phonlengthscale | cat)
##      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## m2    7 1578.6 1614.9 -782.29   1564.6
## m3    8 1575.3 1616.8 -779.64   1559.3  5.2912      1    0.02143 *
## m3.5  9 1558.8 1605.4 -770.39   1540.8 18.5063      1  1.693e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

... all are significant (but not quadratic effect of AoA nor quadratic effect of length).

Main effect of concreteness

```
m3.6 = glmer(bor15.cat ~
                phonlengthscale +
                AoAscale +
                subtlexzipfscale +
                I(subtlexzipfscale^2)  +
                concscale +
                (1 | cat) +
                (0 + AoAscale|cat) +
                (0 + subtlexzipfscale | cat) +
                (0 + phonlengthscale| cat),
            data=dataloan2, family = binomial)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : Model failed to converge with max|grad| = 0.00364188 (tol =
## 0.001, component 1)
```

```
anova(m3.5,m3.6)
```

```
## Data: dataloan2
## Models:
## m3.5: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.5:     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m3.5:     cat) + (0 + phonlengthscale | cat)
## m3.6: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.6:     concscale + (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m3.6:     cat) + (0 + phonlengthscale | cat)
##      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## m3.5  9 1558.8 1605.4 -770.39   1540.8
## m3.6 10 1558.8 1610.7 -769.43   1538.8 1.9273      1     0.1651
```

... not significant

Interaction between AoA and freq.

```
m4 = glmer(bor15.cat ~ phonlengthscale +
            AoAscale*subtlexzipfscale  +
            I(subtlexzipfscale^2)  +
```

```
                (1 | cat) +
                (0 + AoAscale|cat) +
                (0 + subtlexzipfscale | cat) +
                (0 + phonlengthscale| cat),
            data=dataloan2, family = binomial,
            control = glmerControl(optimizer = "bobyqa"))
anova(m3.5,m4)
```

```
## Data: dataloan2
## Models:
## m3.5: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.5:     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m3.5:     cat) + (0 + phonlengthscale | cat)
## m4: bor15.cat ~ phonlengthscale + AoAscale * subtlexzipfscale + I(subtlexzipfscale^2) +
## m4:     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m4:     cat) + (0 + phonlengthscale | cat)
##       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## m3.5   9 1558.8 1605.4 -770.39   1540.8
## m4    10 1559.7 1611.5 -769.85   1539.7 1.0796      1     0.2988
```

... not significant.

Interaction between AoA and length:

```
m5 = glmer(bor15.cat ~
              phonlengthscale*AoAscale +
              subtlexzipfscale+
              I(subtlexzipfscale^2)  +
              (1 | cat) +
              (0 + AoAscale|cat) +
              (0 + subtlexzipfscale | cat) +
              (0 + phonlengthscale| cat),
            data=dataloan2, family = binomial,
            control = glmerControl(optimizer = "bobyqa"))
anova(m3.5,m5)
```

```
## Data: dataloan2
## Models:
## m3.5: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.5:     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m3.5:     cat) + (0 + phonlengthscale | cat)
## m5: bor15.cat ~ phonlengthscale * AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m5:     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m5:     cat) + (0 + phonlengthscale | cat)
##       Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## m3.5   9 1558.8 1605.4 -770.39   1540.8
## m5    10 1559.4 1611.3 -769.72   1539.4 1.336      1     0.2477
```

Not significant.

Interaction between length and freq.

```
m6 = glmer(bor15.cat ~
              phonlengthscale +
              AoAscale+subtlexzipfscale  +
              I(subtlexzipfscale^2)  +
              subtlexzipfscale:phonlengthscale +
```

```
            (1 | cat) +
            (0 + AoAscale|cat) +
            (0 + subtlexzipfscale | cat) +
            (0 + phonlengthscale| cat),
          data=dataloan2, family = binomial,
          control = glmerControl(optimizer = "bobyqa"))
anova(m3.5,m6)
```

```
## Data: dataloan2
## Models:
## m3.5: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.5:      (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m3.5:      cat) + (0 + phonlengthscale | cat)
## m6: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m6:      subtlexzipfscale:phonlengthscale + (1 | cat) + (0 + AoAscale |
## m6:      cat) + (0 + subtlexzipfscale | cat) + (0 + phonlengthscale |
## m6:      cat)
##       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## m3.5   9 1558.8 1605.4 -770.39   1540.8
## m6    10 1560.7 1612.6 -770.37   1540.7 0.0433      1     0.8351
```

Not significant.

Sumamry

The best model is model 3.5 with all main effects and quadratic effect of frequency.

Summary of coefficients:

```
summary(m3.5)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
##     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
##     cat) + (0 + phonlengthscale | cat)
##    Data: dataloan2
##
##      AIC     BIC   logLik deviance df.resid
##   1558.8   1605.4   -770.4   1540.8     1308
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.3763 -0.7398 -0.4067  0.8523  4.2365
##
## Random effects:
##  Groups Name             Variance   Std.Dev.
##  cat    (Intercept)      1.110e+00 1.0534023
##  cat.1  AoAscale         2.946e-08 0.0001716
##  cat.2  subtlexzipfscale 2.289e-08 0.0001513
##  cat.3  phonlengthscale  1.098e-02 0.1047757
## Number of obs: 1317, groups:  cat, 12
```

```
## 
## Fixed effects:
##                    Estimate Std. Error z value Pr(>|z|)    
## (Intercept)        -1.33469    0.47005  -2.839  0.00452 ** 
## phonlengthscale     0.48507    0.08943   5.424 5.83e-08 ***
## AoAscale            0.58584    0.08701   6.733 1.66e-11 ***
## subtlexzipfscale    0.38898    0.09622   4.042 5.29e-05 ***
## I(subtlexzipfscale^2) -0.24455  0.05950  -4.110 3.95e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Correlation of Fixed Effects:
##            (Intr) phnlng AoAscl sbtlxz
## phnlngthscl -0.109                     
## AoAscale     0.004 -0.066              
## sbtlxzpfscl -0.190  0.151  0.555       
## I(sbtlxz^2) -0.271  0.064 -0.195  0.068
## convergence code: 0
## Model failed to converge with max|grad| = 0.00352504 (tol = 0.001, component 1)
```

Plot fixed effects and random effects:

```
sjp.glmer(m3.5, 'fe')
```


Fixed effects
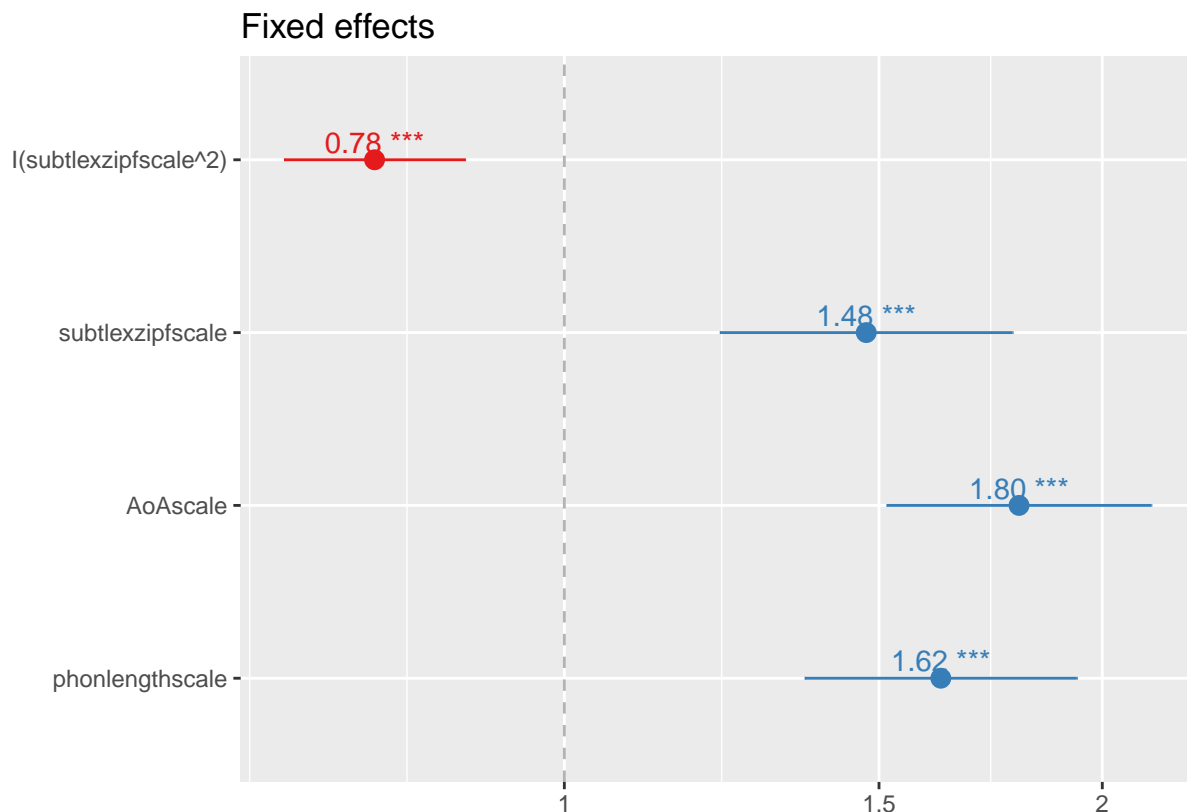
```
sjp.glmer(m3.5, 'eff')
```

Marginal effects of model predictors
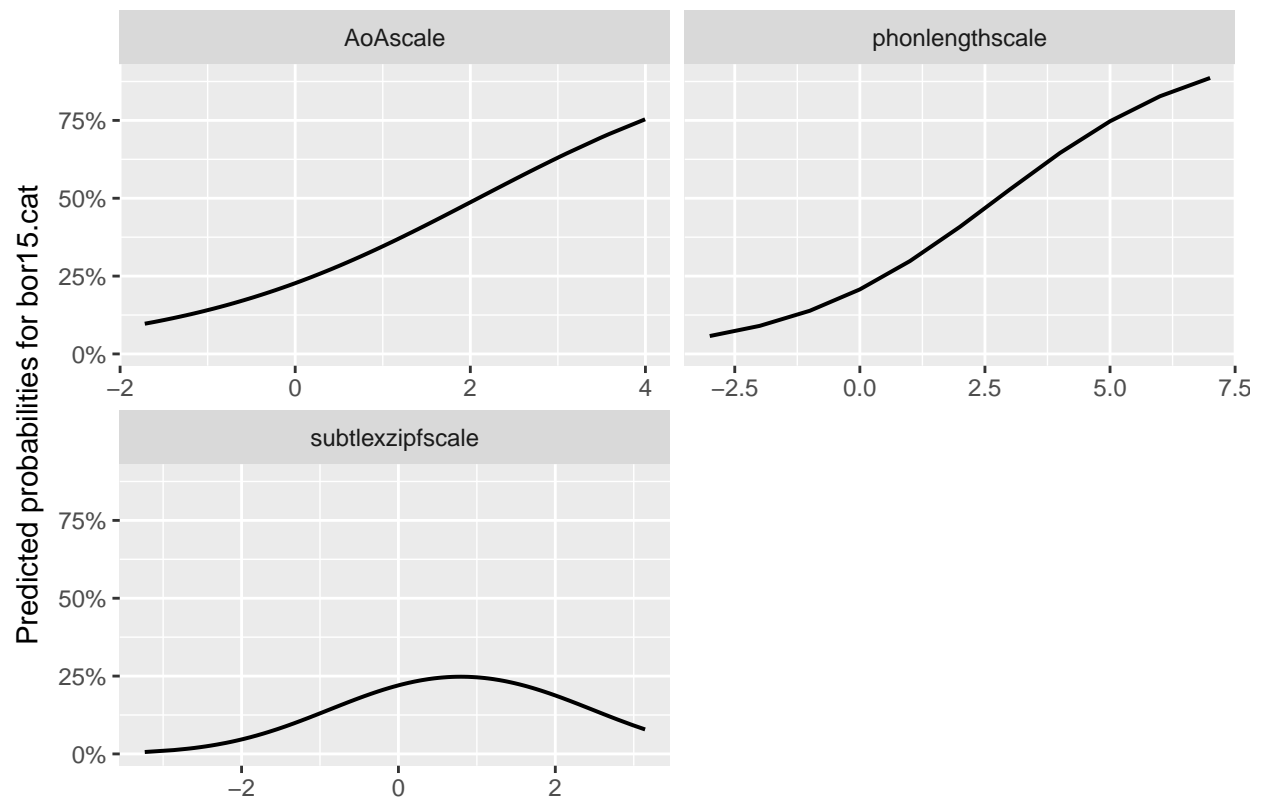

```
dotplot(ranef(m3.5))
```

```
## $cat
```

cat
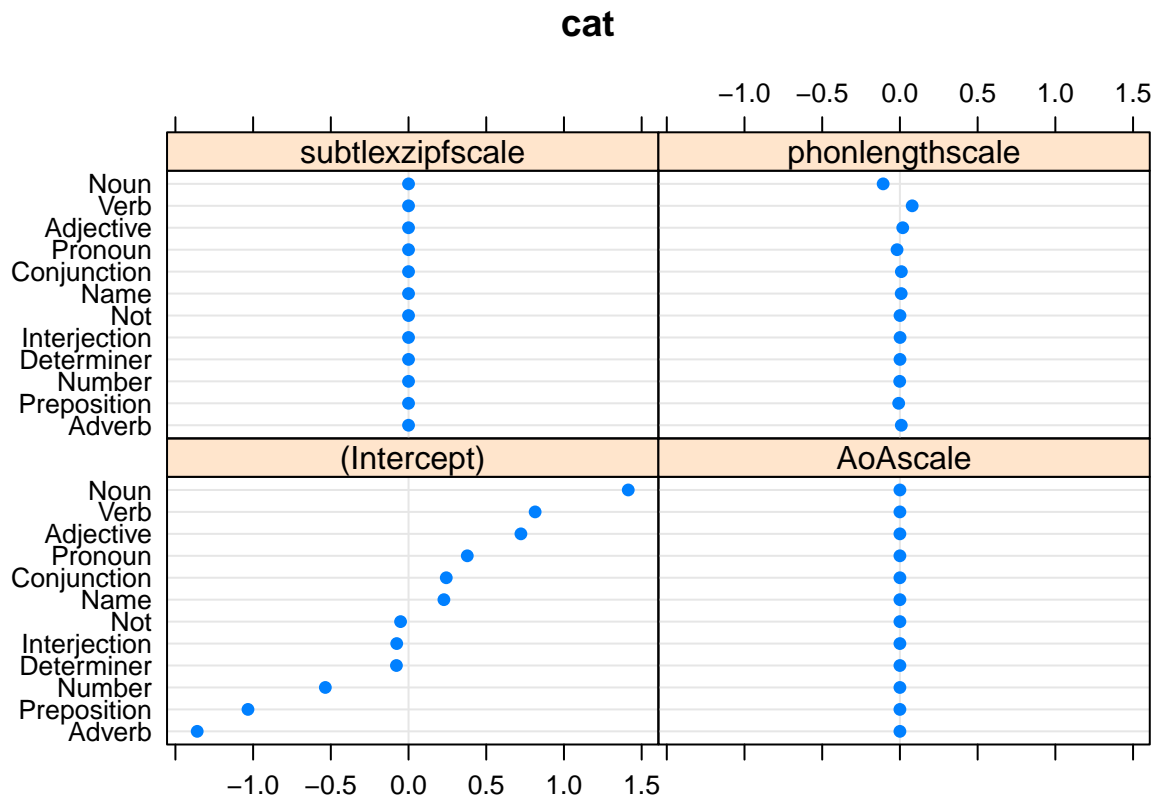
Main random effects are for the general level of borrowing prob., very little difference for slopes.

Check the effect of random intercept with the full model:

```
m3.5.nointercept = glmer(bor15.cat ~
                         phonlengthscale +
             AoAscale +
             subtlexzipfscale  +
             I(subtlexzipfscale^2)  +
                         (0 + AoAscale|cat) +
                         (0 + subtlexzipfscale | cat) +
                         (0 + phonlengthscale| cat),
             data=dataloan2, family = binomial)
anova(m3.5,m3.5.nointercept)
```

```
## Data: dataloan2
## Models:
## m3.5.nointercept: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.5.nointercept:     (0 + AoAscale | cat) + (0 + subtlexzipfscale | cat) + (0 +
## m3.5.nointercept:     phonlengthscale | cat)
## m3.5: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.5:     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m3.5:     cat) + (0 + phonlengthscale | cat)
##                  Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## m3.5.nointercept  8 1585.0 1626.5 -784.51   1569.0
## m3.5              9 1558.8 1605.4 -770.39   1540.8 28.25      1  1.066e-07
##
## m3.5.nointercept
## m3.5             ***
## ---
```

13

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

... significant: different POS vary in their baseline probability of being borrowed.

Check the effects of random slopes with the full model (i.e. do the strength of the main effects vary by POS?):

```
m3.5.noAoASlope = glmer(bor15.cat ~
                        phonlengthscale +
            AoAscale +
            subtlexzipfscale  +
            I(subtlexzipfscale^2)  +
                    (1 | cat) +
                    (0 + subtlexzipfscale | cat) +
                    (0 + phonlengthscale| cat),
                        data=dataloan2, family = binomial)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : Model failed to converge with max|grad| = 0.00505438 (tol =
## 0.001, component 1)
```

```
anova(m3.5, m3.5.noAoASlope )
```

```
## Data: dataloan2
## Models:
## m3.5.noAoASlope: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.5.noAoASlope:     (1 | cat) + (0 + subtlexzipfscale | cat) + (0 + phonlengthscale |
## m3.5.noAoASlope:     cat)
## m3.5: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.5:     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m3.5:     cat) + (0 + phonlengthscale | cat)
##                 Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## m3.5.noAoASlope  8 1556.8 1598.2 -770.39   1540.8
## m3.5             9 1558.8 1605.4 -770.39   1540.8     0      1     0.9985
```

```
m3.5.nofreqSlope = glmer(bor15.cat ~
                        phonlengthscale +
            AoAscale +
            subtlexzipfscale  +
            I(subtlexzipfscale^2)  +
                    (1 | cat) +
                    (0 + AoAscale | cat) +
                    (0 + phonlengthscale| cat),
                data=dataloan2, family = binomial,
                control = glmerControl(optimizer = 'bobyqa'))
anova(m3.5, m3.5.nofreqSlope )
```

```
## Data: dataloan2
## Models:
## m3.5.nofreqSlope: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.5.nofreqSlope:     (1 | cat) + (0 + AoAscale | cat) + (0 + phonlengthscale |
## m3.5.nofreqSlope:     cat)
## m3.5: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2) +
## m3.5:     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m3.5:     cat) + (0 + phonlengthscale | cat)
##                  Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## m3.5.nofreqSlope  8 1556.8 1598.2 -770.39   1540.8
## m3.5              9 1558.8 1605.4 -770.39   1540.8     0      1          1
```

```
m3.5.nolengthSlope = glmer(bor15.cat ~
                           phonlengthscale +
              AoAscale +
              subtlexzipfscale  +
              I(subtlexzipfscale^2)  +
                           (1 | cat) +
                           (0 + AoAscale | cat) +
                           (0 + subtlexzipfscale| cat),
                    data=dataloan2, family = binomial)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : Model failed to converge with max|grad| = 0.00668295 (tol =
## 0.001, component 1)
```

```
anova(m4, m3.5.nolengthSlope )
```

```
## Data: dataloan2
## Models:
## m3.5.nolengthSlope: bor15.cat ~ phonlengthscale + AoAscale + subtlexzipfscale + I(subtlexzipfscale^2)
## m3.5.nolengthSlope:     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m3.5.nolengthSlope:     cat)
## m4: bor15.cat ~ phonlengthscale + AoAscale * subtlexzipfscale + I(subtlexzipfscale^2) +
## m4:     (1 | cat) + (0 + AoAscale | cat) + (0 + subtlexzipfscale |
## m4:     cat) + (0 + phonlengthscale | cat)
##                    Df    AIC    BIC  logLik deviance  Chisq Chi Df
## m3.5.nolengthSlope  8 1558.4 1599.9 -771.22   1542.4
## m4                 10 1559.7 1611.5 -769.85   1539.7 2.7345      2
##                    Pr(>Chisq)
## m3.5.nolengthSlope
## m4                     0.2548
```

. . . random slopes do not improve the fit of the model over the fixed effects. So strength does not vary by
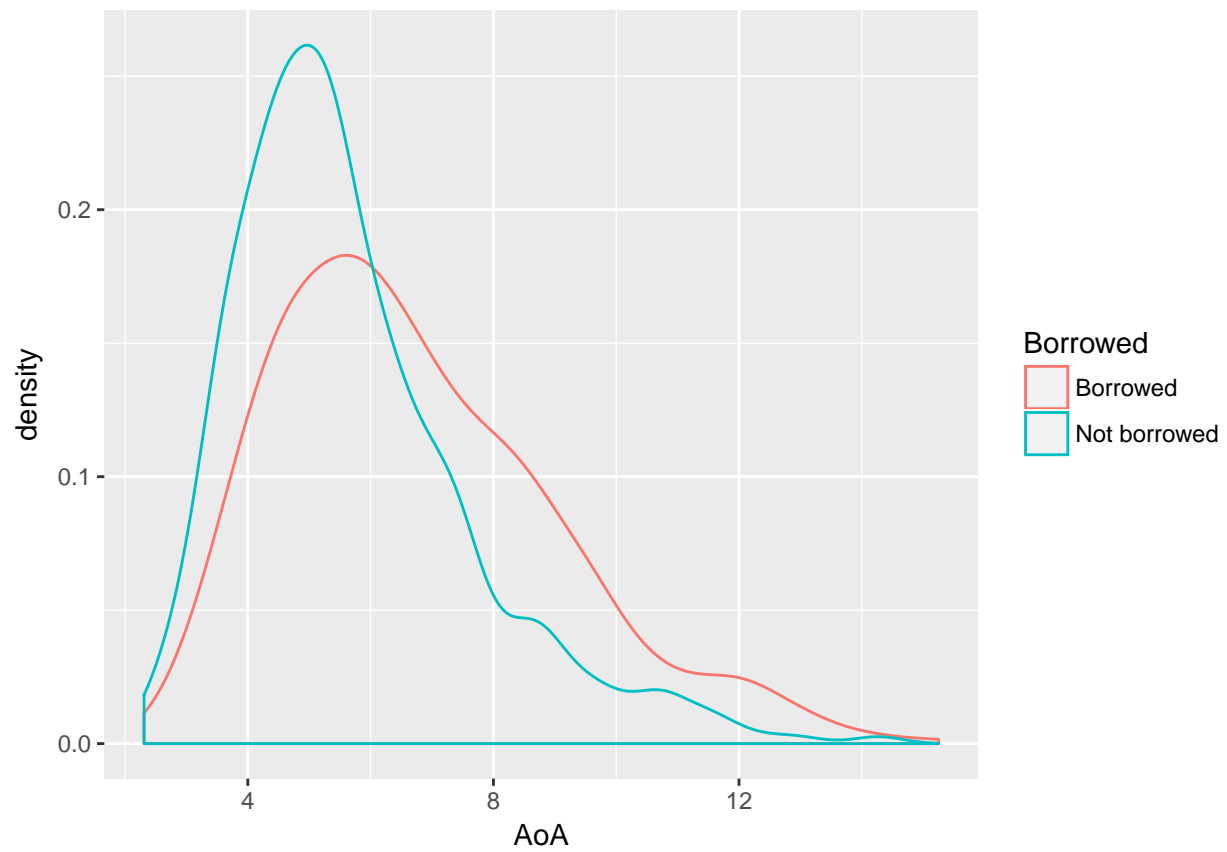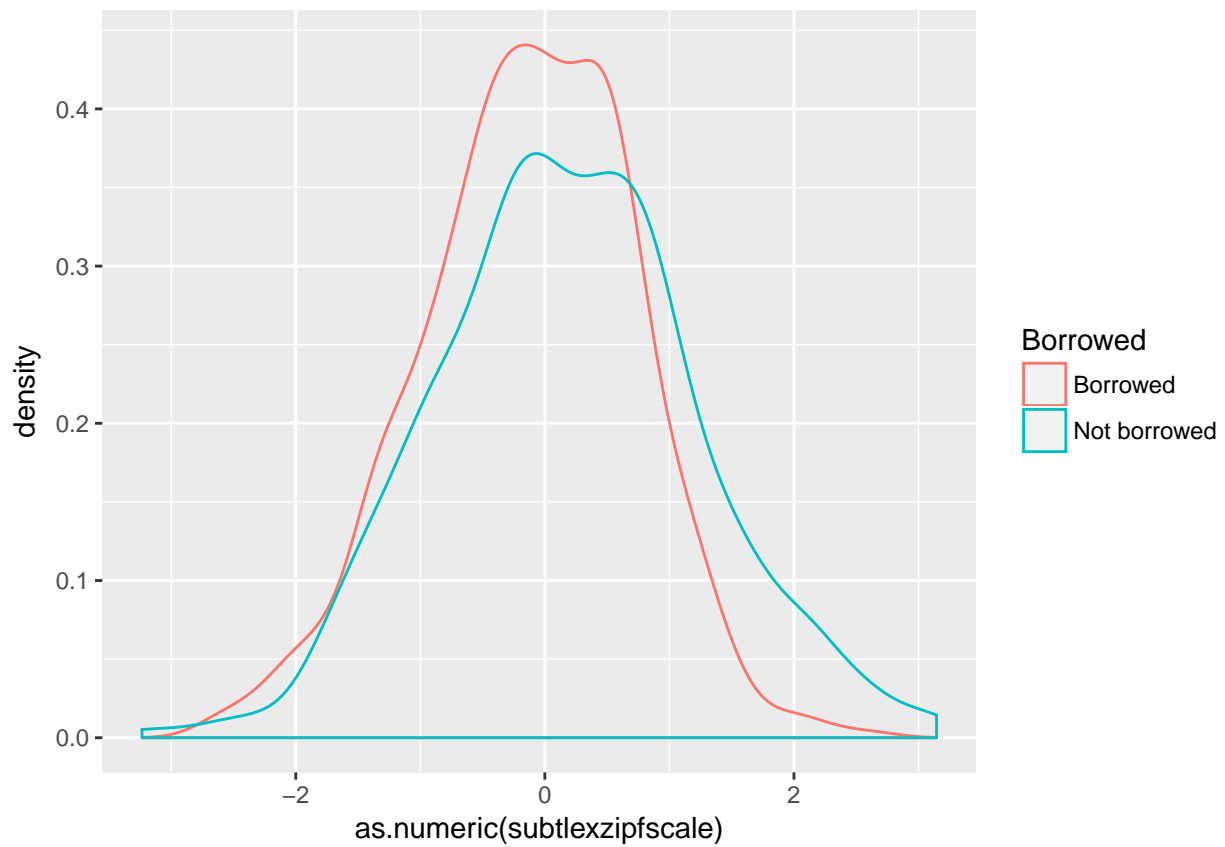POS.

Plots

Raw data

```
dataloan2$Borrowed = c("Not borrowed", "Borrowed")[dataloan2$bor15+1]
ggplot(dataloan2[!is.na(dataloan2$Borrowed),], aes(x=AoA, colour=Borrowed)) +
  geom_density()
```
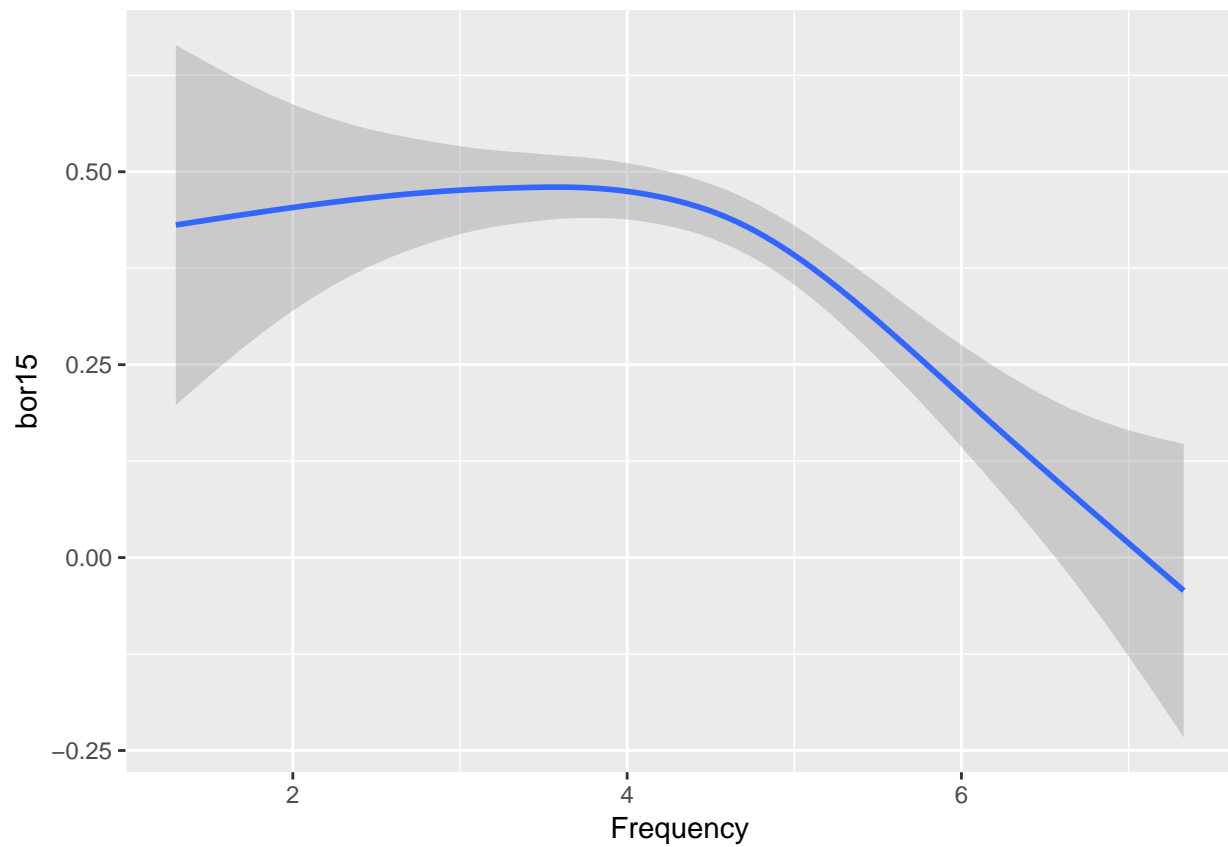
```
ggplot(dataloan2[!is.na(dataloan2$Borrowed),], aes(x=as.numeric(subtlexzipfscale), colour=Borrowed)) +
  geom_density()
```
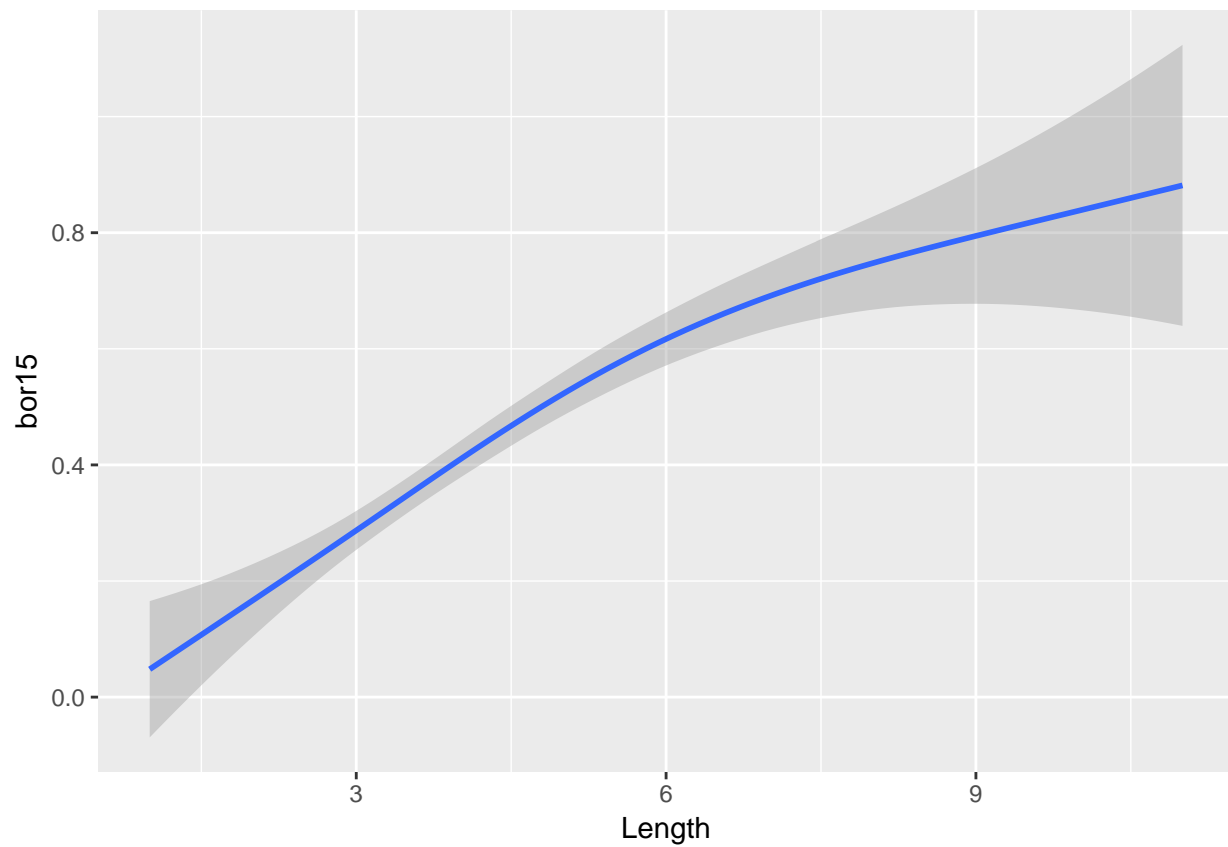
```
ggplot(dataloan2[!is.na(dataloan2$Borrowed),], aes(x=as.numeric(subtlexzipf), y=bor15)) +
  stat_smooth()+
  xlab("Frequency")
```

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

```
ggplot(dataloan2[!is.na(dataloan2$Borrowed),], aes(x=as.numeric(phonlength), y=bor15)) +
  stat_smooth() +
  xlab("Length")
```

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

```
dataloan2$subtlexzipf.cat = cut(
  dataloan2$subtlexzipf,
  breaks = quantile(dataloan2$subtlexzipf,
                    prob=seq(0,1,length.out=4)),
  include.lowest = T)


#dataloan2[!is.na(dataloan2$AoA) & !is.na(dataloan2$subtlexzipf) & !is.na(dataloan2$bor15.catLabel),] %>

ggplot(dataloan2[!is.na(dataloan2$Borrowed),], aes(y=AoA, colour=Borrowed, x=subtlexzipf.cat)) +
  geom_boxplot()
```
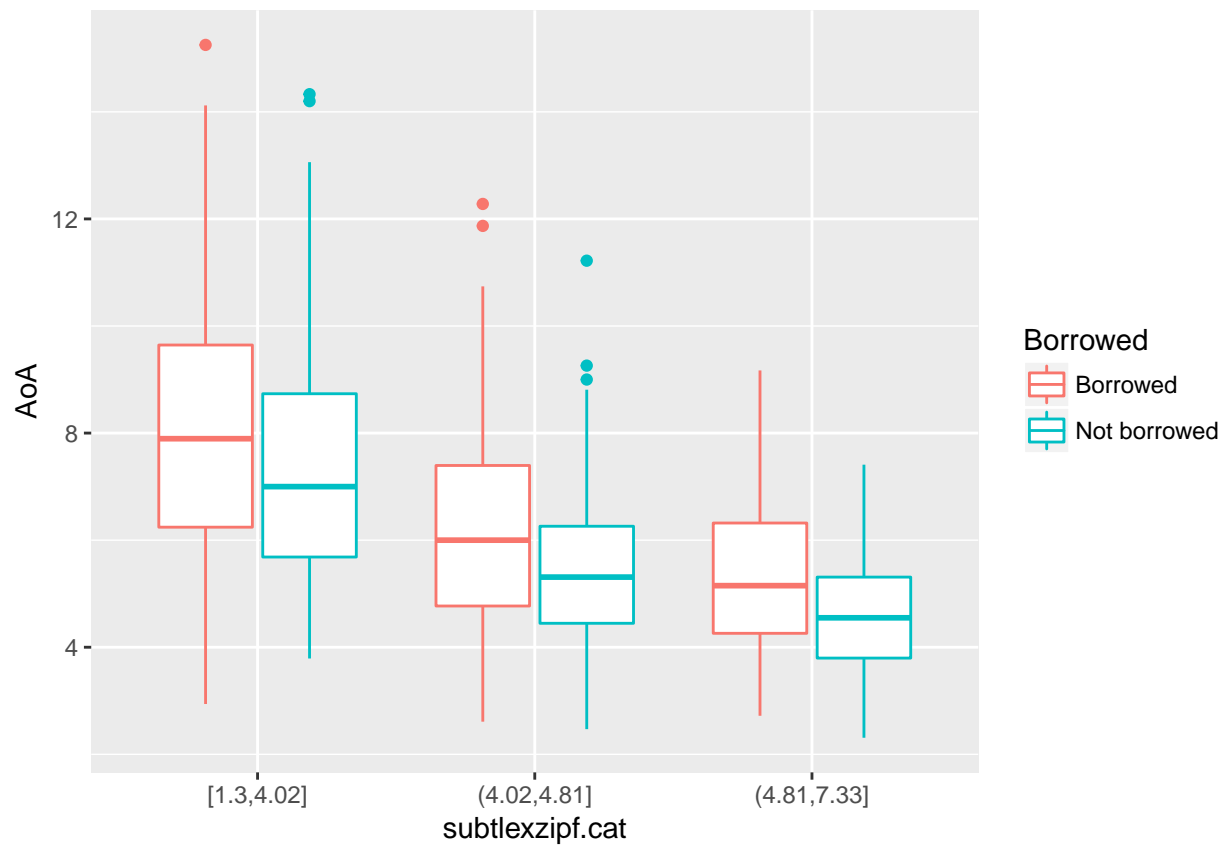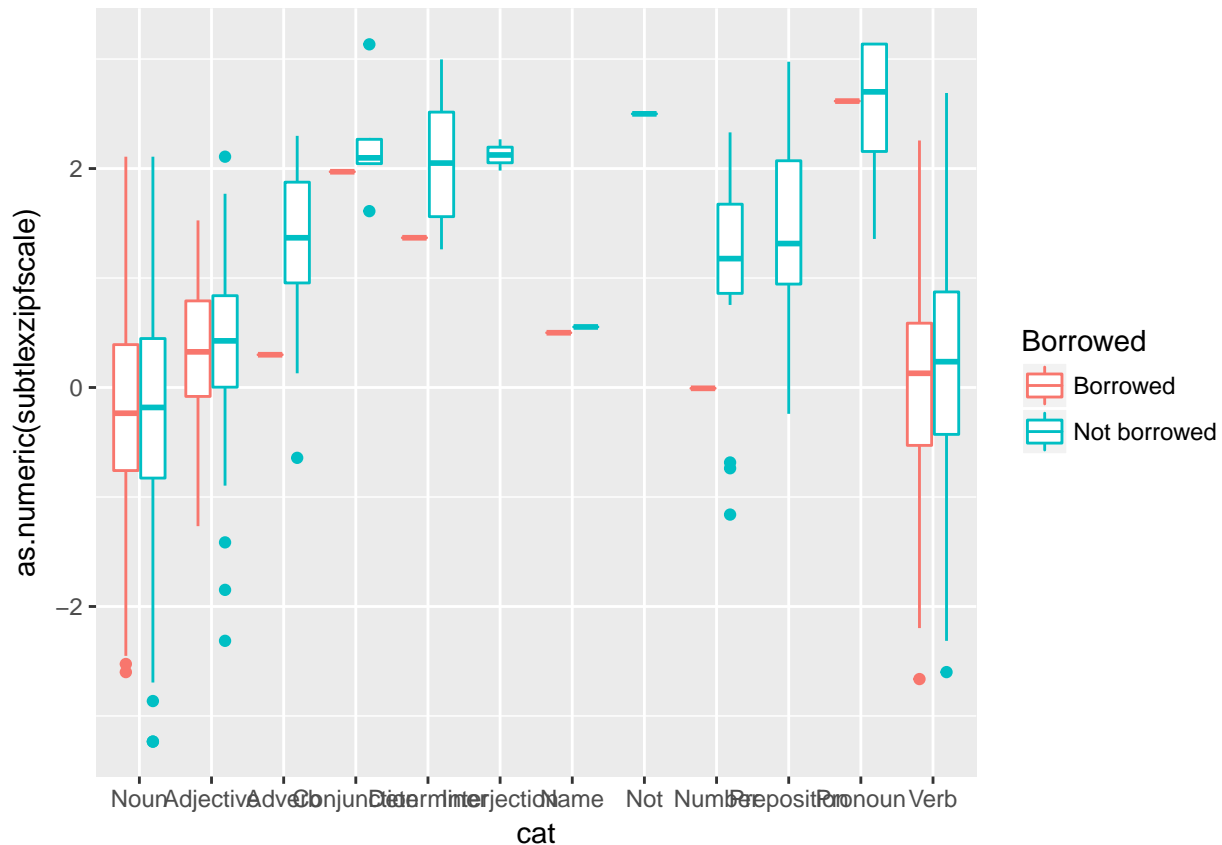
```
ggplot(dataloan2[!is.na(dataloan2$Borrowed),], aes(y=AoA, colour=Borrowed, x=cat)) +
  geom_boxplot()
```
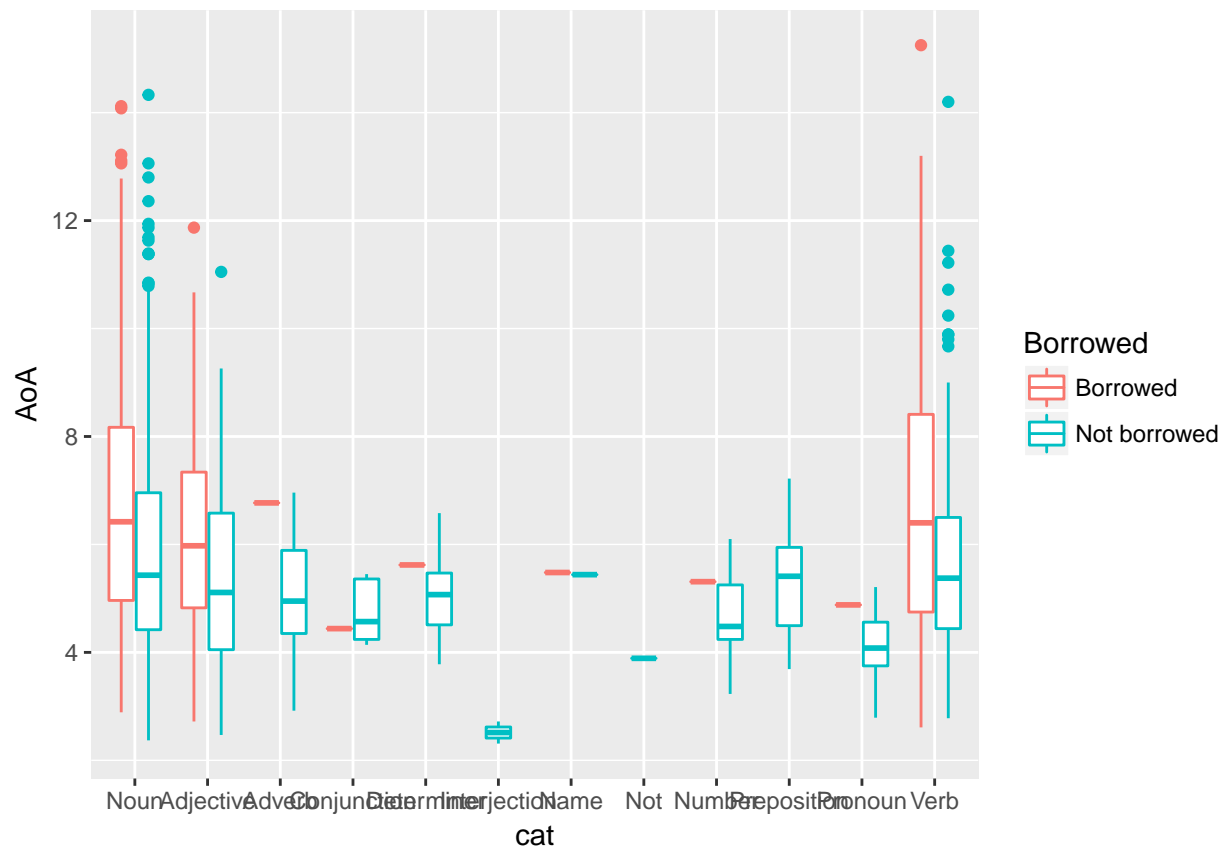
```
ggplot(dataloan2[!is.na(dataloan2$Borrowed),], aes(y=as.numeric(subtlexzipfscale), colour=Borrowed, x=ca
  geom_boxplot()
```

Model predictions: plot the interaction between frequency and age of acquisition.

```r
aoarange = range(dataloan2$AoAscale)

pd = data.frame(
  AoAscale = seq(aoarange[1],aoarange[2], length.out = 100),
  subtlexzipfscale = 0,
  concscale = 0,
  phonlengthscale = 0)

# scale some ages
year3 = (3/aoaSD) -aoaMean
year6 = (6/aoaSD) - aoaMean
year9 = (9/aoaSD) -aoaMean

pd$subtlexzipfscale = year6
predMed = logit2per(predict(m4,pd, re.form=NA))
pd$subtlexzipfscale = year3
predLow = logit2per(predict(m4,pd, re.form=NA))
pd$subtlexzipfscale = year9
predHigh = logit2per(predict(m4,pd, re.form=NA))

pd$AoA = (pd$AoAscale + aoaMean) * aoaSD

plot(pd$AoA, predMed, type='l', ylim=c(0,1),
     xlab='Age of acquisition',
     ylab = "Probability of borrowing",
```
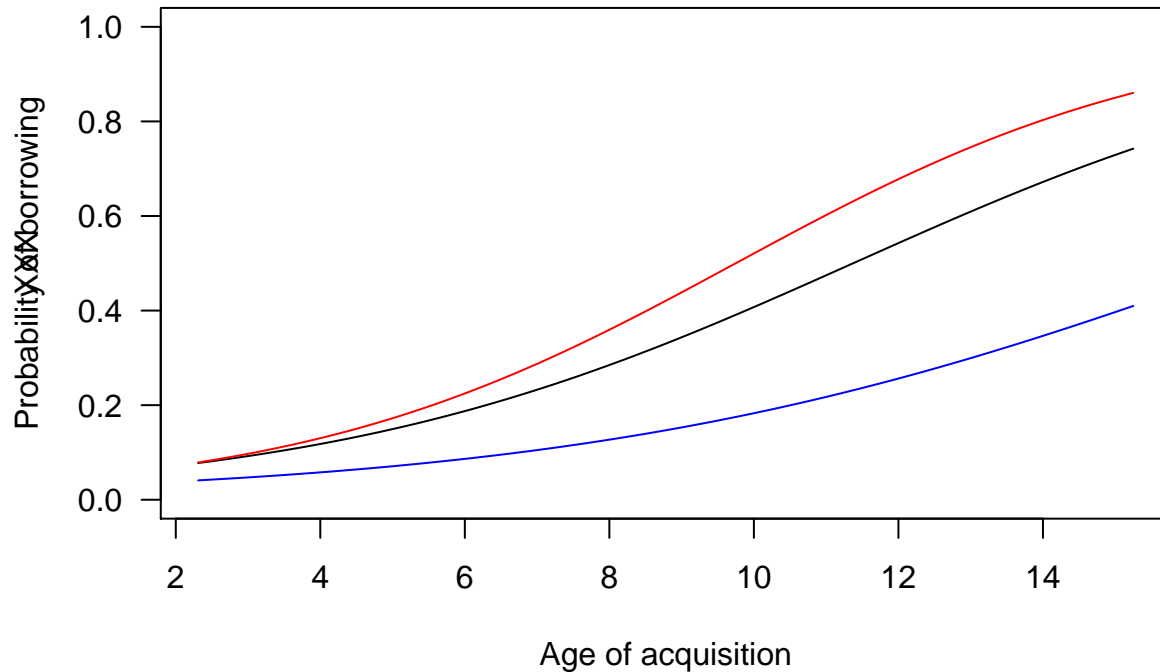
```
    las=1)
title(ylab='XXX',srt=90)
points(pd$AoA, predLow, type='l', col='blue')
points(pd$AoA, predHigh, type='l', col='red')
```



```
#effects::effect("subtlexzipfscale",m4)

# Use effect to predict values for prob borrowing for
#  3 age groups
intdf  = data.frame(
  effects::effect("AoAscale:subtlexzipfscale",m4, xlevels=list(AoAscale=c(year3,year6,year9), subtlexzi
#colnames(intdf) <- c("grp", "x", "y", "se", "conf.low",
#                 "conf.high")
intdf$AoAscale = factor(intdf$AoAscale)

baseplot <- ggplot(intdf,
                   aes_string(x = "subtlexzipfscale", y = "fit",
             colour = "AoAscale"))
baseplot <- baseplot +
  geom_ribbon(
    aes_string(
      ymin = "lower",
      ymax = "upper",
      colour = NULL,
      fill = "AoAscale"),
    alpha = 0.3,
    show.legend = FALSE) +
  geom_line(size = 2) +
  scale_color_discrete(breaks=c(year9,year6,year3),
                       labels=c("9y","6y",'3y'),
                       name="AoA") +
  xlab("Frequency (log, scaled)") +
```

```
ylab("Probability of borrowing")
```

baseplot