

# Establishing the reliability and validity of measures extracted from long-form recordings

## Contents

Recalculate everything or not? . . . . .	1
Generate ICC tables . . . . .	1
Describe datasets . . . . .	1
Reliability analyses combining all corpora . . . . .	1
Paired analysis with minimum distance between recs . . . . .	3
Reliability analyses per corpus . . . . .	7
Reliability by child age . . . . .	7
Validity against age . . . . .	14
Formula stats . . . . .	14
Save information about packages used . . . . .	14

## Recalculate everything or not?

If RECALC is set to TRUE, then the ICC tables will be re-generated.

## Generate ICC tables

### Describe datasets

We are looking here at 6 corpora, 198 children, 1175 recordings, 36 many metrics.

Note that for LENA

### Reliability analyses combining all corpora

```
##  
## Call:  
## lm(formula = icc_child_id ~ subject + data_set, data = df.icc.mixed)  
##  
## Residuals:  
##      Min      1Q   Median      3Q      Max  
## -0.091268 -0.016615  0.000447  0.029711  0.075692  
##
```

Table 1: Measures with the lowest ICC attributed to children.

data_set	metric	icc_child_id	icc_corpus
lena	lena_CVC	0.29	0.26
aclew	avg_cry_voc_dur_chi	0.30	0.02
aclew	sc_adu_ph	0.32	0.15
aclew	pc_adu_ph	0.32	0.15
aclew	wc_adu_ph	0.32	0.15
lena	lena_CTC	0.33	0.27

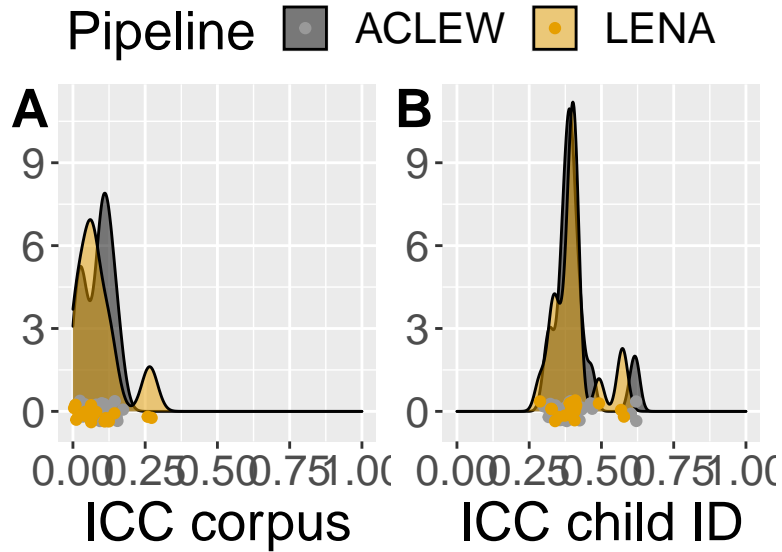


Figure 1: Distribution of ICC attributed to corpus (a) and children (b), when combining data from all corpora.

Table 2: Measures with the highest ICC attributed to children.

data_set	metric	icc_child_id	icc_corpus
lena	avg_voc_dur_och	0.49	0.00
lena	voc_och_ph	0.57	0.01
lena	voc_dur_och_ph	0.58	0.01
aclew	voc_och_ph	0.60	0.04
aclew	avg_voc_dur_och	0.62	0.01
aclew	voc_dur_och_ph	0.62	0.02

Table 3: Most commonly used metrics.

metric	LENA ICC	ACLEW ICC
wc_adu_ph	0.37	0.32
voc_chi_ph	0.41	0.39

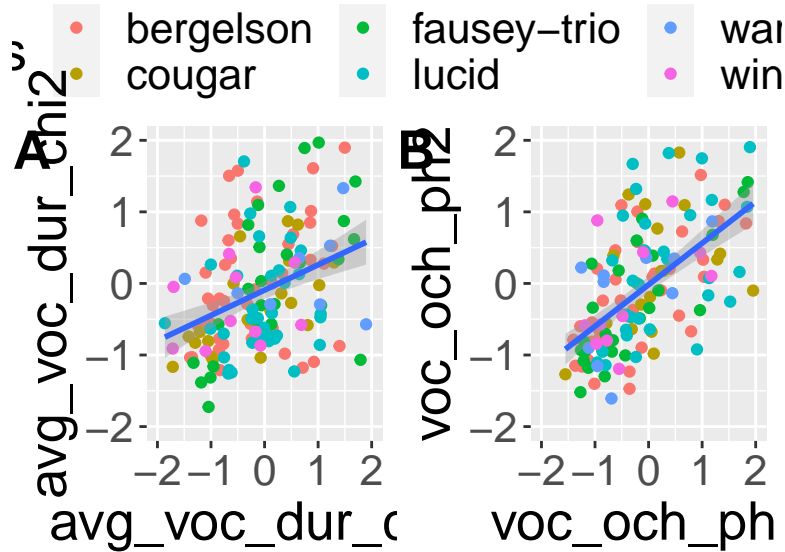


Figure 2: (A) scatterplot for one variable with relatively low ICCs versus (B) one with relatively higher ICCs (see Tables 1-2 for details)

```
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.39237   0.00835  46.993 < 2e-16 ***
## subjectoch   0.19492   0.01705  11.434 4.84e-15 ***
## subjectfem   0.01111   0.01404   0.791 0.43285
## subjectmal  -0.00820   0.01404  -0.584 0.56194
## subjectadu  -0.05512   0.01991  -2.768 0.00809 **
## data_setlena -0.01470   0.01084  -1.356 0.18168
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03663 on 46 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.7811, Adjusted R-squared:  0.7573
## F-statistic: 32.83 on 5 and 46 DF,  p-value: 4.213e-14
##
## Call:
## lm(formula = icc_child_id ~ subject + data_set, data = df.icc.mixed,
##     subset = c(metric %in% common_metrics))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.080765 -0.018157  0.003443  0.024985  0.055559
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.402753   0.011789  34.164 < 2e-16 ***
## subjectoch   0.185059   0.016922  10.936 1.29e-11 ***
## subjectfem   0.004236   0.015544   0.273  0.7872
## subjectmal  -0.013266   0.015544  -0.853  0.4007
## subjectadu  -0.046668   0.025384  -1.839  0.0766 .
```

```
## data_setlena -0.015732  0.011240  -1.400  0.1726
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03277 on 28 degrees of freedom
## Multiple R-squared:  0.8616, Adjusted R-squared:  0.8369
## F-statistic: 34.87 on 5 and 28 DF,  p-value: 3.432e-11
```

### Paired analysis with minimum distance between recs

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      2.00  10.00   12.00   13.12  12.00   49.00

##
##      bergelson      cougar fausey-trio      lucid warlaumont      winnipeg
##              347              200              56              6              0              28

## [1] 637

##
##      bergelson      cougar fausey-trio      lucid warlaumont      winnipeg
##              44              26              28              6              0              12

## [1] 116

##      lena_CVC      lena_CTC      lena_wc_adu_ph      aclew_voc_chi_ph
##      Min.      :0.6073      Min.      :0.6329      Min.      :0.4143      Min.      :0.6984
##      1st Qu.:0.6742      1st Qu.:0.7069      1st Qu.:0.4904      1st Qu.:0.7338
##      Median :0.7040      Median :0.7232      Median :0.5449      Median :0.7556
##      Mean   :0.7015      Mean   :0.7262      Mean   :0.5431      Mean   :0.7550
##      3rd Qu.:0.7221      3rd Qu.:0.7579      3rd Qu.:0.5867      3rd Qu.:0.7758
##      Max.   :0.8239      Max.   :0.7878      Max.   :0.6649      Max.   :0.8075

##      aclew_wc_adu_ph
##      Min.      :0.3416
##      1st Qu.:0.5133
##      Median :0.5490
##      Mean   :0.5439
##      3rd Qu.:0.5850
##      Max.   :0.6559
```

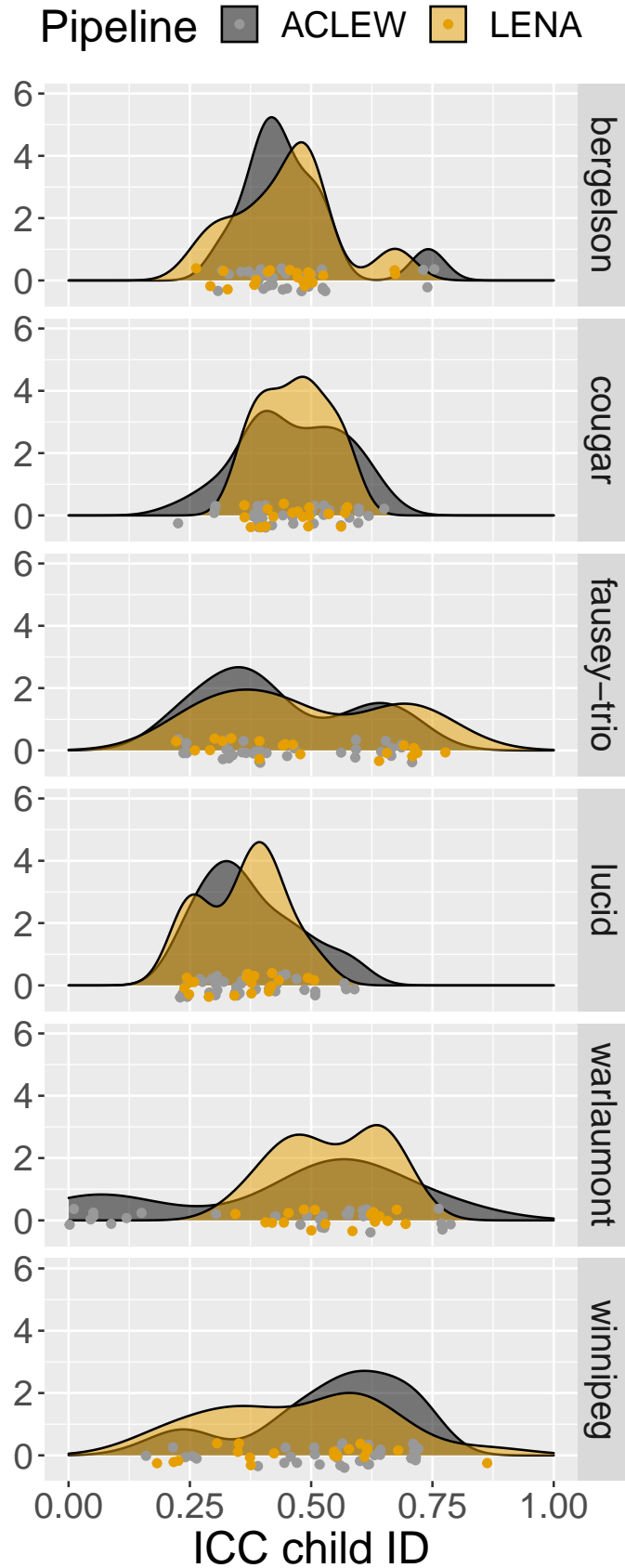


Figure 3: Distribution of ICC attributed to children in each separate corpus.

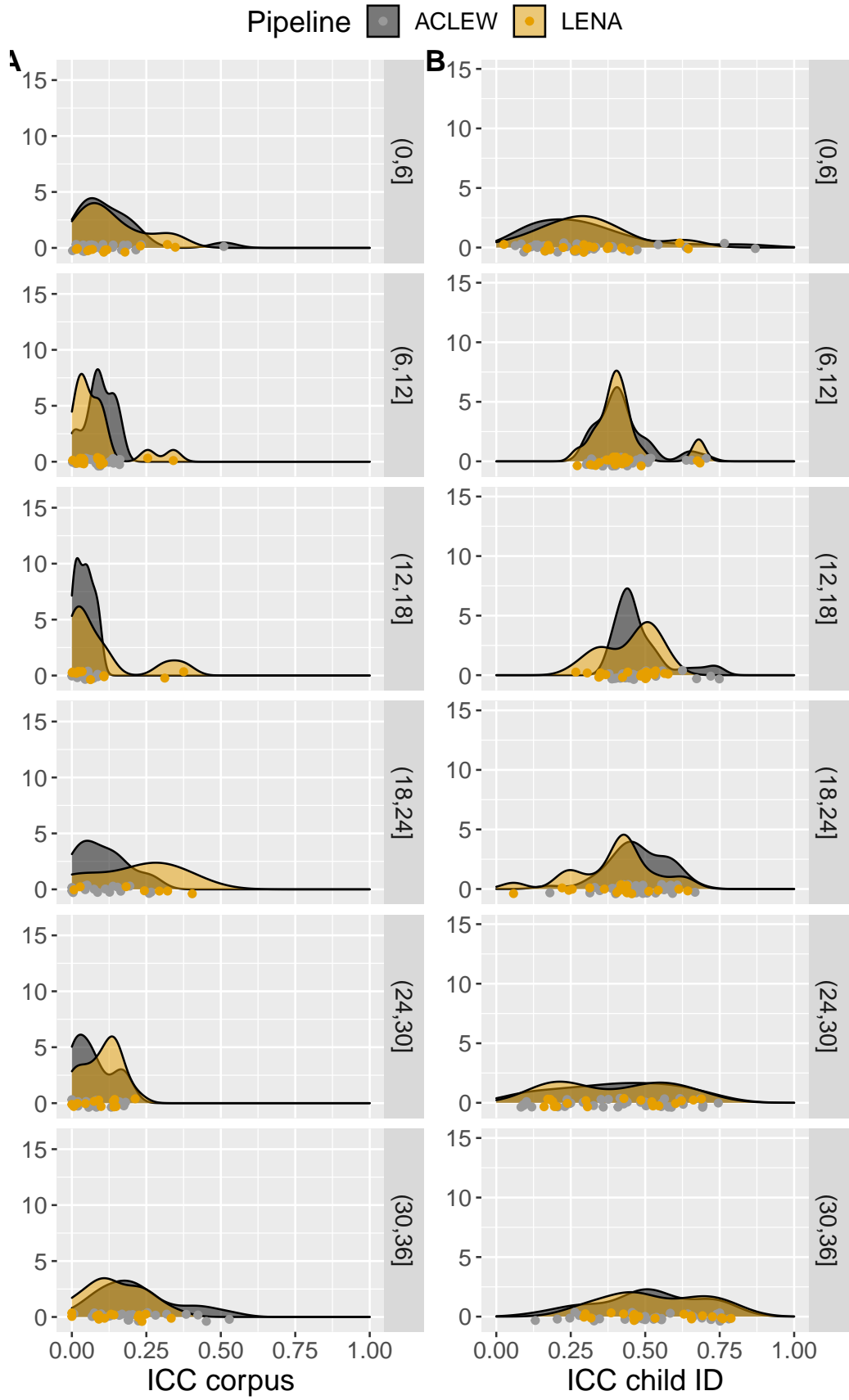


Figure 4: Distribution of ICC attributed to corpus (a) and children (b), when binning children's age.

Table 4: Most commonly used metrics by age bin (LENA).

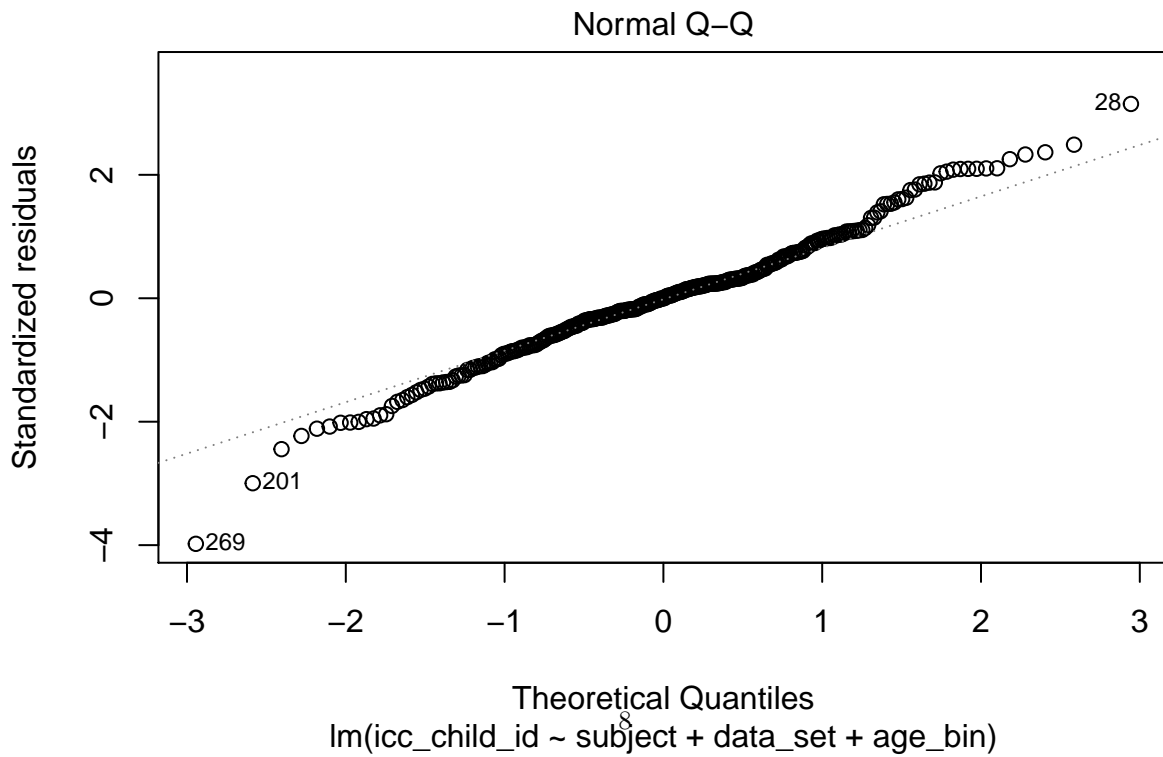
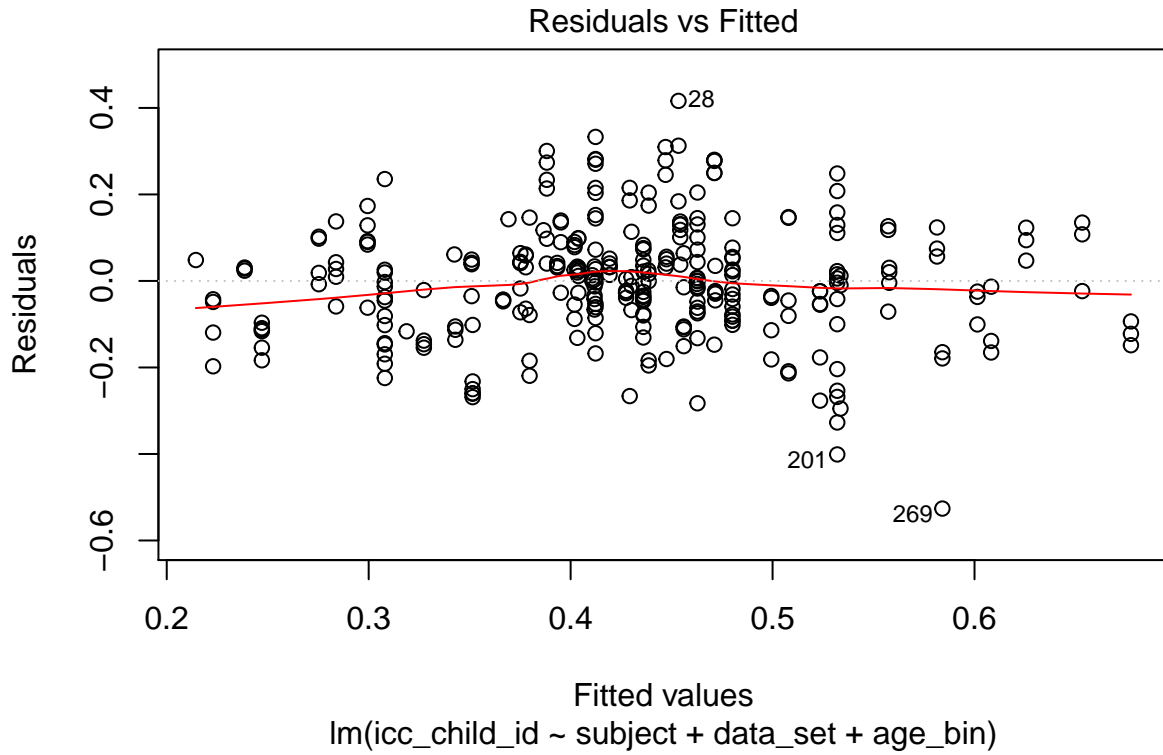
metric	age_bin	icc_child_id
wc_adu_ph	(0,6]	0.26
voc_chi_ph	(0,6]	0.29
wc_adu_ph	(6,12]	0.40
voc_chi_ph	(6,12]	0.42
wc_adu_ph	(12,18]	0.50
voc_chi_ph	(12,18]	0.34
wc_adu_ph	(18,24]	0.51
voc_chi_ph	(18,24]	0.61
wc_adu_ph	(24,30]	0.20
voc_chi_ph	(24,30]	0.62
wc_adu_ph	(30,36]	0.46
voc_chi_ph	(30,36]	0.43

Table 5: Most commonly used metrics by age bin (aclew).

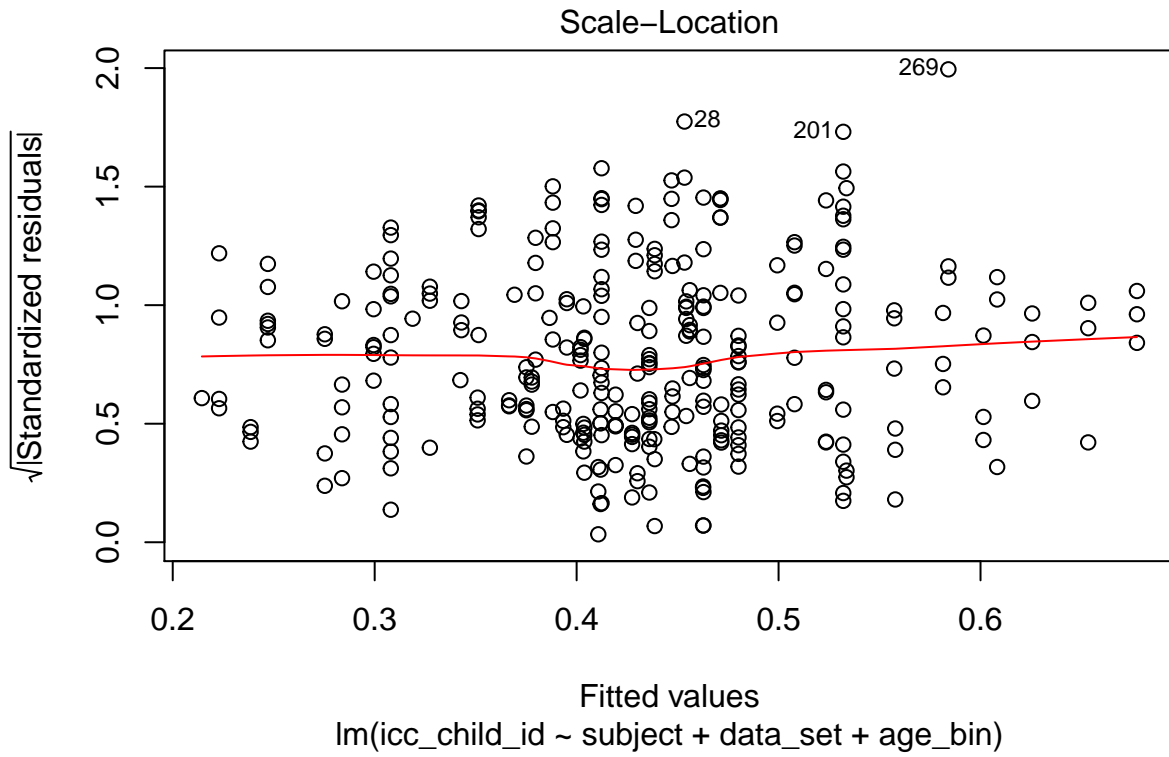
metric	age_bin	icc_child_id
wc_adu_ph	(0,6]	0.26
voc_chi_ph	(0,6]	0.33
wc_adu_ph	(6,12]	0.32
voc_chi_ph	(6,12]	0.52
wc_adu_ph	(12,18]	0.41
voc_chi_ph	(12,18]	0.39
wc_adu_ph	(18,24]	0.43
voc_chi_ph	(18,24]	0.46
wc_adu_ph	(24,30]	0.24
voc_chi_ph	(24,30]	0.44
wc_adu_ph	(30,36]	0.47
voc_chi_ph	(30,36]	0.13

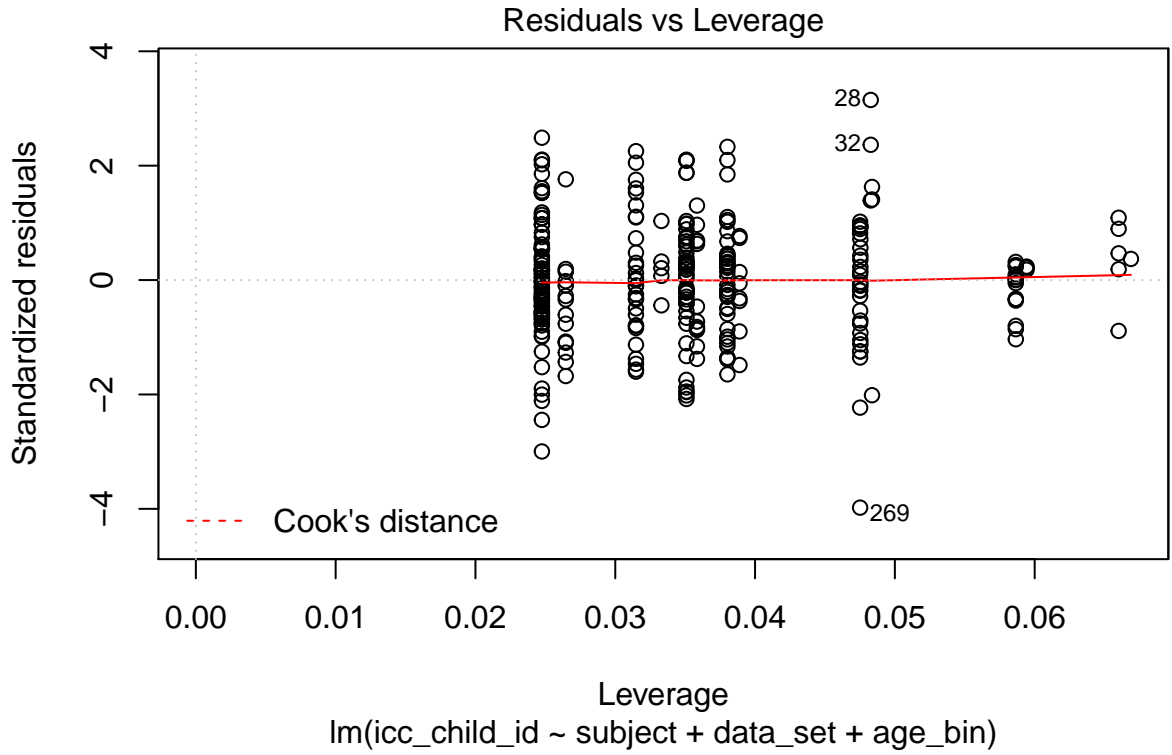
# Reliability analyses per corpus

## Reliability by child age



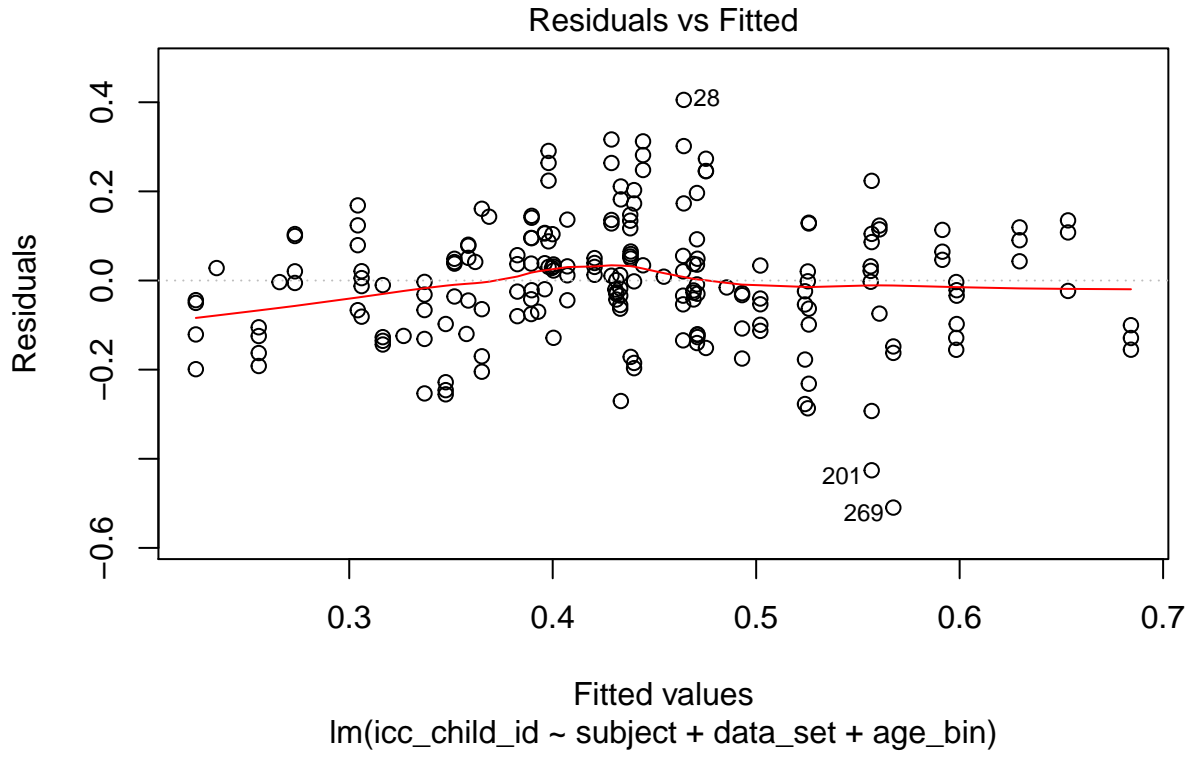


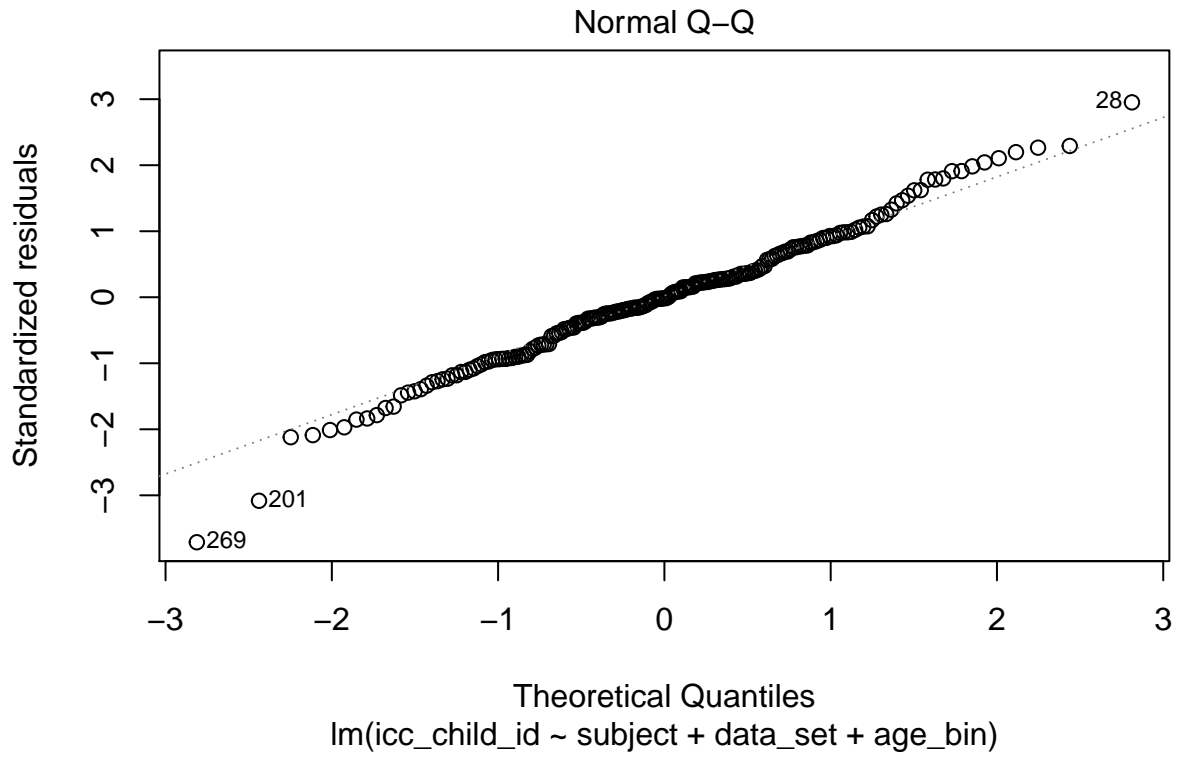


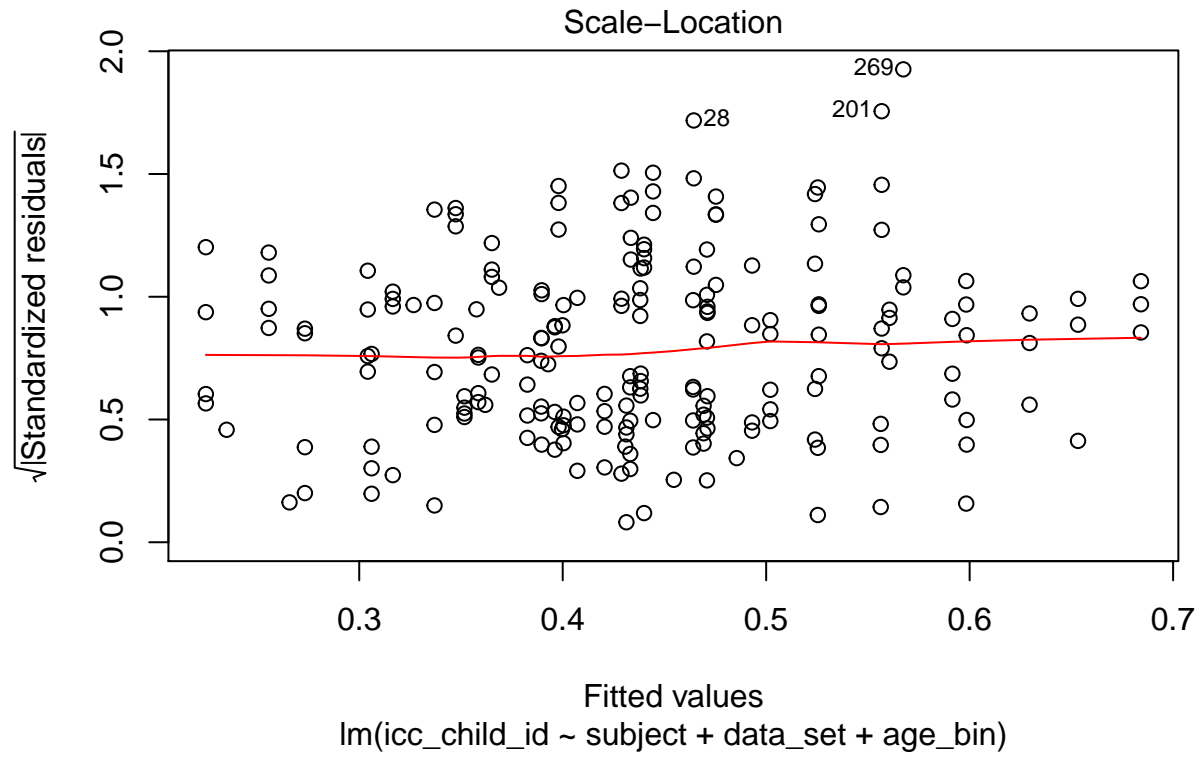


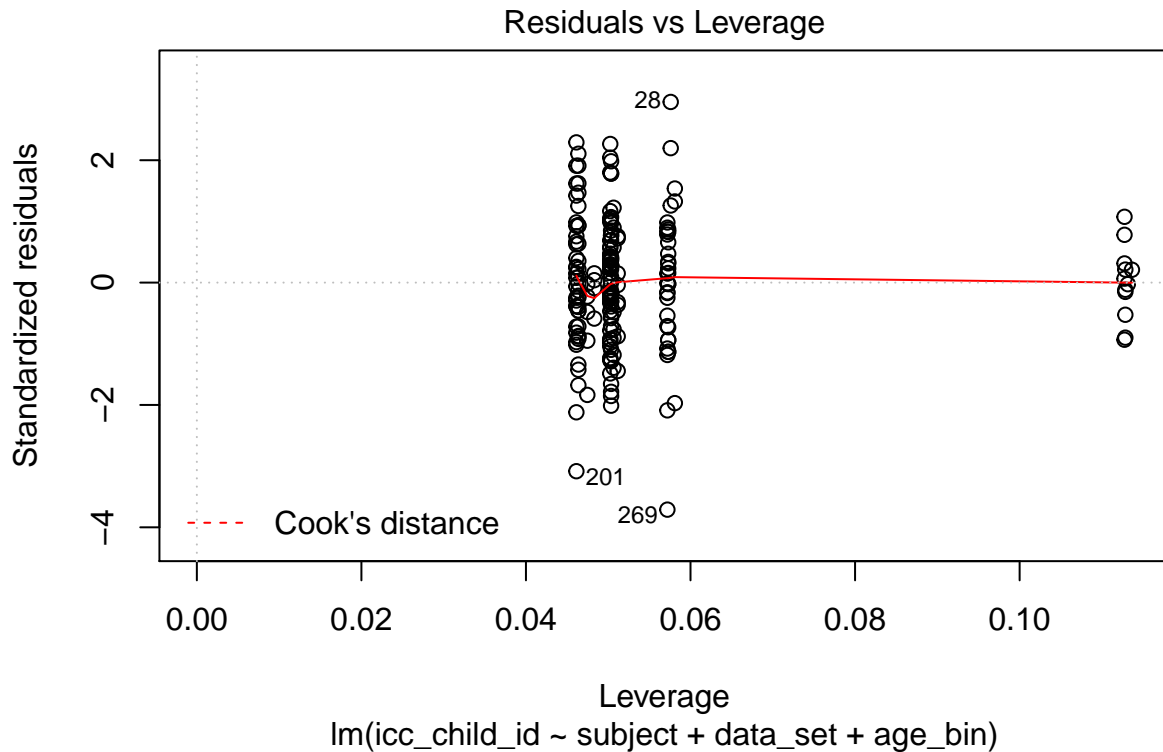
```
##
## Call:
## lm(formula = icc_child_id ~ subject + data_set + age_bin, data = df.icc.age)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.52618 -0.07743 -0.00015  0.07301  0.41625
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.307981   0.022042  13.973 < 2e-16 ***
## subjectoch    0.145425   0.025820   5.632 4.12e-08 ***
## subjectfem   -0.008532   0.021284  -0.401 0.688819
## subjectmal   -0.060873   0.021284  -2.860 0.004535 **
## subjectadu   -0.069427   0.030127  -2.304 0.021885 *
## data_setlena -0.024144   0.016456  -1.467 0.143376
## age_bin(6,12] 0.128023   0.026984   4.744 3.25e-06 ***
## age_bin(12,18] 0.172171   0.026984   6.381 6.73e-10 ***
## age_bin(18,24] 0.154826   0.026984   5.738 2.36e-08 ***
## age_bin(24,30] 0.104368   0.026984   3.868 0.000135 ***
## age_bin(30,36] 0.224057   0.026984   8.303 3.58e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1355 on 298 degrees of freedom
## (9 observations deleted due to missingness)
```

```
## Multiple R-squared:  0.3183, Adjusted R-squared:  0.2955
## F-statistic: 13.92 on 10 and 298 DF,  p-value: < 2.2e-16
```









```
##
## Call:
## lm(formula = icc_child_id ~ subject + data_set + age_bin, data = df.icc.age,
##     subset = c(metric %in% common_metrics))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.50949 -0.08059 -0.00195  0.08698  0.40524
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.33700    0.03080  10.940 < 2e-16 ***
## subjectoch     0.12741    0.02991   4.260 3.20e-05 ***
## subjectfem    -0.03279    0.02749  -1.193 0.234428
## subjectmal    -0.08147    0.02749  -2.964 0.003424 **
## subjectadu    -0.07125    0.04478  -1.591 0.113220
## data_setlena  -0.03091    0.01985  -1.557 0.121033
## age_bin(6,12]  0.12710    0.03456   3.678 0.000305 ***
## age_bin(12,18] 0.16505    0.03456   4.776 3.54e-06 ***
## age_bin(18,24] 0.13391    0.03456   3.875 0.000146 ***
## age_bin(24,30] 0.09189    0.03456   2.659 0.008501 **
## age_bin(30,36] 0.21978    0.03456   6.360 1.44e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1414 on 192 degrees of freedom
```

```
## (1 observation deleted due to missingness)
## Multiple R-squared: 0.3383, Adjusted R-squared: 0.3039
## F-statistic: 9.817 on 10 and 192 DF, p-value: 3.379e-13
```

## Validity against age

### Formula stats

Mixed models formula summary:

Var1	Freq
full	50
no_exp	3

Corpus models formula summary:

Var1	Freq
no_chi_effect	4
no_exp	314

Age models formula summary:

Var1	Freq
15	53
41	13
42	40
44	5
45	11
46	37
59	1
62	6
63	11
64	32
91	5
92	23
93	25
117	1
120	5
121	4
122	4
123	28
124	11

## Save information about packages used

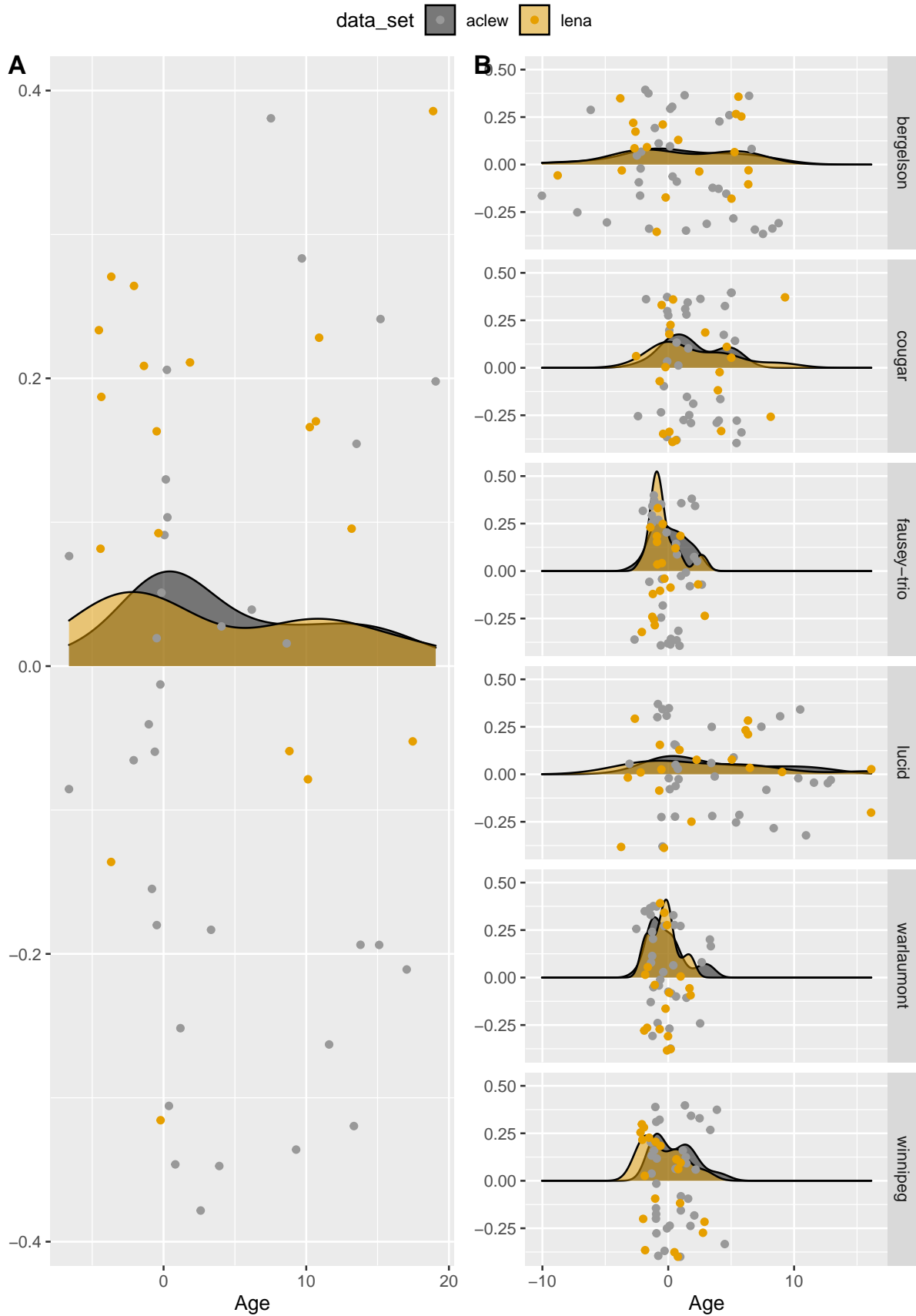


Figure 5: Distribution of  $t$  for age when all corpora are analyzed together (a) or for each corpus separately (b).