Vowel duration and tongue root advancement in Italian and Polish

Stefano Coretta

18/09/2017

1 Import data

`tongues` contains splines coordinates at maximum displacement, `tongues_clos` has splines coordinates at acoustic closure, and `vowels` has durational data.

```
languages <- read_csv("data/languages.csv")
```

```
## Parsed with column specification:
## cols(
##   speaker = col_character(),
##   language = col_character()
## )
```

```
words <- read_csv("data/nonce.csv")
```

```
## Parsed with column specification:
## cols(
##   item = col_integer(),
##   word = col_character(),
##   ipa = col_character(),
##   c1 = col_character(),
##   c1phonation = col_character(),
##   vowel = col_character(),
##   anteropost = col_character(),
##   height = col_character(),
##   c2 = col_character(),
##   c2phonation = col_character(),
##   c2place = col_character(),
##   language = col_character()
## )
```

```
columns <- c(
    "speaker",
    "seconds",
    "rec.date",
    "prompt",
    "label",
    "TT.displacement.sm",
    "TT.velocity",
    "TT.velocity.abs",
    "TD.displacement.sm",
    "TD.velocity",
    "TD.velocity.abs"
)

aaa_files <- list.files(
```

```
    path = "./data/tongue",
    pattern = "*-tongue-cart.tsv",
    full.names = TRUE
)

tongues <- read_aaa(
    aaa_files,
    columns,
    na.rm = TRUE
) %>%
    mutate(word = word(prompt, 2)) %>%
    left_join(y = languages) %>%
    left_join(y = words) %>%
    mutate_if(is.character, as.factor) %>%
    group_by(speaker) %>%
    mutate(
        X.re = rescale(X),
        Y.re = rescale(Y)
    ) %>%
    ungroup() %>%
    mutate(
        vowel.ord = ordered(vowel, levels = c("a", "o", "u")),
        c2place.ord = ordered(c2place, levels = c("coronal", "velar")),
        c2phonation.ord = ordered(c2phonation, levels = c("voiceless", "voiced"))
    ) %>%
    filter(label %in% c("max_TT", "max_TD"), vowel != "u") %>%
    arrange(rec.date, fan.line) %>%
    create_event_start("rec.date")
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   speaker = col_character(),
##   rec.date = col_character(),
##   prompt = col_character(),
##   label = col_character()
## )
```

```
## See spec(...) for full column specifications.
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   speaker = col_character(),
##   rec.date = col_character(),
##   prompt = col_character(),
##   label = col_character()
## )
```

```
## See spec(...) for full column specifications.
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   speaker = col_character(),
##   rec.date = col_character(),
```

```
##   prompt = col_character(),
##   label = col_character()
## )
```

```
## See spec(...) for full column specifications.
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   speaker = col_character(),
##   rec.date = col_character(),
##   prompt = col_character(),
##   label = col_character()
## )
```

```
## See spec(...) for full column specifications.
```

```
## Joining, by = "speaker"
```

```
## Joining, by = c("word", "language")
```

```
aaa_files_clos <- list.files(
    path = "./data/tongue",
    pattern = "*-tongue-clos-cart.tsv",
    full.names = TRUE
)

tongues_clos <- read_aaa(
    aaa_files_clos,
    columns,
    na.rm = TRUE
) %>%
    mutate(word = word(prompt, 2)) %>%
    left_join(y = languages) %>%
    left_join(y = words) %>%
    mutate_if(is.character, as.factor) %>%
    group_by(speaker) %>%
    mutate(
        X.re = rescale(X),
        Y.re = rescale(Y)
    ) %>%
    ungroup() %>%
    mutate(
        vowel.ord = ordered(vowel, levels = c("a", "o", "u")),
        c2place.ord = ordered(c2place, levels = c("coronal", "velar")),
        c2phonation.ord = ordered(c2phonation, levels = c("voiceless", "voiced"))
    ) %>%
    filter(vowel != "u") %>%
    arrange(rec.date, fan.line) %>%
    create_event_start("rec.date")
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   speaker = col_character(),
##   rec.date = col_character(),
##   prompt = col_character(),
```

```
##   label = col_character(),
##   X_2 = col_character(),
##   Y_2 = col_character(),
##   X_3 = col_character(),
##   Y_3 = col_character(),
##   X_4 = col_character(),
##   Y_4 = col_character(),
##   X_5 = col_character(),
##   Y_5 = col_character(),
##   X_6 = col_character(),
##   Y_6 = col_character()
## )

## See spec(...) for full column specifications.

## Parsed with column specification:
## cols(
##   .default = col_double(),
##   speaker = col_character(),
##   rec.date = col_character(),
##   prompt = col_character(),
##   label = col_character(),
##   X_3 = col_character(),
##   Y_3 = col_character(),
##   X_4 = col_character(),
##   Y_4 = col_character(),
##   X_5 = col_character(),
##   Y_5 = col_character()
## )

## See spec(...) for full column specifications.

## Parsed with column specification:
## cols(
##   .default = col_double(),
##   speaker = col_character(),
##   rec.date = col_character(),
##   prompt = col_character(),
##   label = col_character()
## )

## See spec(...) for full column specifications.

## Parsed with column specification:
## cols(
##   .default = col_double(),
##   speaker = col_character(),
##   rec.date = col_character(),
##   prompt = col_character(),
##   label = col_character()
## )

## See spec(...) for full column specifications.

## Joining, by = "speaker"

## Joining, by = c("word", "language")
```

```
vowels <- list.files(path = "data/durations",
                     pattern = "*-vowel-durations.csv",
                     full.names = TRUE) %>%
    map_df(~read_csv(., na = "--undefined--")) %>%
    left_join(y = languages) %>%
    left_join(y = words) %>%
    mutate_if(is.character, as.factor) %>%
    mutate(
        index = as.factor(index),
        c2phonation = factor(c2phonation, levels = c("voiceless", "voiced"))
    ) %>%
    filter(!(file == "it04-002" | file == "pl03-020")) # filter obvious outliers
```

```
## Parsed with column specification:
## cols(
##   index = col_integer(),
##   speaker = col_character(),
##   file = col_character(),
##   word = col_character(),
##   time = col_double(),
##   word.duration = col_double(),
##   c1.duration = col_double(),
##   vowel.duration = col_double(),
##   closure.duration = col_double(),
##   rvot = col_double(),
##   c2.duration = col_double(),
##   v2.duration = col_double(),
##   sentence.duration = col_double()
## )
```

```
## Parsed with column specification:
## cols(
##   index = col_integer(),
##   speaker = col_character(),
##   file = col_character(),
##   word = col_character(),
##   time = col_double(),
##   word.duration = col_double(),
##   c1.duration = col_double(),
##   vowel.duration = col_double(),
##   closure.duration = col_double(),
##   rvot = col_double(),
##   c2.duration = col_double(),
##   v2.duration = col_double(),
##   sentence.duration = col_double()
## )
## Parsed with column specification:
## cols(
##   index = col_integer(),
##   speaker = col_character(),
##   file = col_character(),
##   word = col_character(),
##   time = col_double(),
##   word.duration = col_double(),
```

```
##   c1.duration = col_double(),
##   vowel.duration = col_double(),
##   closure.duration = col_double(),
##   rvot = col_double(),
##   c2.duration = col_double(),
##   v2.duration = col_double(),
##   sentence.duration = col_double()
## )
## Parsed with column specification:
## cols(
##   index = col_integer(),
##   speaker = col_character(),
##   file = col_character(),
##   word = col_character(),
##   time = col_double(),
##   word.duration = col_double(),
##   c1.duration = col_double(),
##   vowel.duration = col_double(),
##   closure.duration = col_double(),
##   rvot = col_double(),
##   c2.duration = col_double(),
##   v2.duration = col_double(),
##   sentence.duration = col_double()
## )
## Parsed with column specification:
## cols(
##   index = col_integer(),
##   speaker = col_character(),
##   file = col_character(),
##   word = col_character(),
##   time = col_double(),
##   word.duration = col_double(),
##   c1.duration = col_double(),
##   vowel.duration = col_double(),
##   closure.duration = col_double(),
##   rvot = col_double(),
##   c2.duration = col_double(),
##   v2.duration = col_double(),
##   sentence.duration = col_double()
## )
## Parsed with column specification:
## cols(
##   index = col_integer(),
##   speaker = col_character(),
##   file = col_character(),
##   word = col_character(),
##   time = col_double(),
##   word.duration = col_double(),
##   c1.duration = col_double(),
##   vowel.duration = col_double(),
##   closure.duration = col_double(),
##   rvot = col_double(),
##   c2.duration = col_double(),
##   v2.duration = col_double(),
```

```
##   sentence.duration = col_double()
## )
## Parsed with column specification:
## cols(
##   index = col_integer(),
##   speaker = col_character(),
##   file = col_character(),
##   word = col_character(),
##   time = col_double(),
##   word.duration = col_double(),
##   c1.duration = col_double(),
##   vowel.duration = col_double(),
##   closure.duration = col_double(),
##   rvot = col_double(),
##   c2.duration = col_double(),
##   v2.duration = col_double(),
##   sentence.duration = col_double()
## )
## Parsed with column specification:
## cols(
##   index = col_integer(),
##   speaker = col_character(),
##   file = col_character(),
##   word = col_character(),
##   time = col_double(),
##   word.duration = col_double(),
##   c1.duration = col_double(),
##   vowel.duration = col_double(),
##   closure.duration = col_double(),
##   rvot = col_double(),
##   c2.duration = col_double(),
##   v2.duration = col_double(),
##   sentence.duration = col_double()
## )

## Joining, by = "speaker"

## Joining, by = c("word", "language")
```

2 Vowel duration

2.1 Italian

```
it_vow_lm <- lmer(
    vowel.duration ~
        c2phonation +
        c2place +
        vowel +
        c2phonation:vowel +
        sentence.duration +
        (1+c2phonation|speaker) +
        (1|word),
    data = filter(vowels, language == "italian")
```

```
)

summary(it_vow_lm)
```

```
## Linear mixed model fit by REML t-tests use Satterthwaite approximations
##   to degrees of freedom [lmerMod]
## Formula:
## vowel.duration ~ c2phonation + c2place + vowel + c2phonation:vowel +
##     sentence.duration + (1 + c2phonation | speaker) + (1 | word)
##    Data: filter(vowels, language == "italian")
##
## REML criterion at convergence: 3289.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.8640 -0.6543 -0.0251  0.5852  3.2478
##
## Random effects:
##  Groups   Name              Variance Std.Dev. Corr
##  word     (Intercept)        34.65    5.887
##  speaker  (Intercept)        56.01    7.484
##           c2phonationvoiced  51.78    7.196   0.53
##  Residual                   191.02   13.821
## Number of obs: 406, groups:  word, 24; speaker, 4
##
## Fixed effects:
##                           Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)                 14.513     12.434 133.600   1.167  0.24520
## c2phonationvoiced           21.841      6.073  12.700   3.596  0.00335 **
## c2placevelar                -8.524      2.803  15.700  -3.041  0.00791 **
## vowelo                      -8.697      4.864  15.800  -1.788  0.09296 .
## vowelu                     -29.686      4.860  15.800  -6.108 1.62e-05 ***
## sentence.duration           77.009      6.662 336.600  11.559  < 2e-16 ***
## c2phonationvoiced:vowelo     2.561      6.869  15.700   0.373  0.71423
## c2phonationvoiced:vowelu   -15.577      6.866  15.700  -2.269  0.03777 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                           (Intr) c2phnt c2plcv vowelo vowelu sntnc.
## c2phontnvcd               -0.059
## c2placevelr               -0.112 -0.002
## vowelo                    -0.174  0.401 -0.003
## vowelu                    -0.182  0.402 -0.002  0.502
## sentnc.drtn               -0.905 -0.005  0.000 -0.024 -0.015
## c2phonationvoiced:vowelo   0.126 -0.566  0.002 -0.708 -0.355  0.014
## c2phonationvoiced:vowelu   0.136 -0.566  0.002 -0.355 -0.708  0.003
##                           c2phonationvoiced:vowelo
## c2phontnvcd
## c2placevelr
## vowelo
## vowelu
## sentnc.drtn
## c2phonationvoiced:vowelo
```

```
## c2phonationvoiced:vowelu  0.501
```

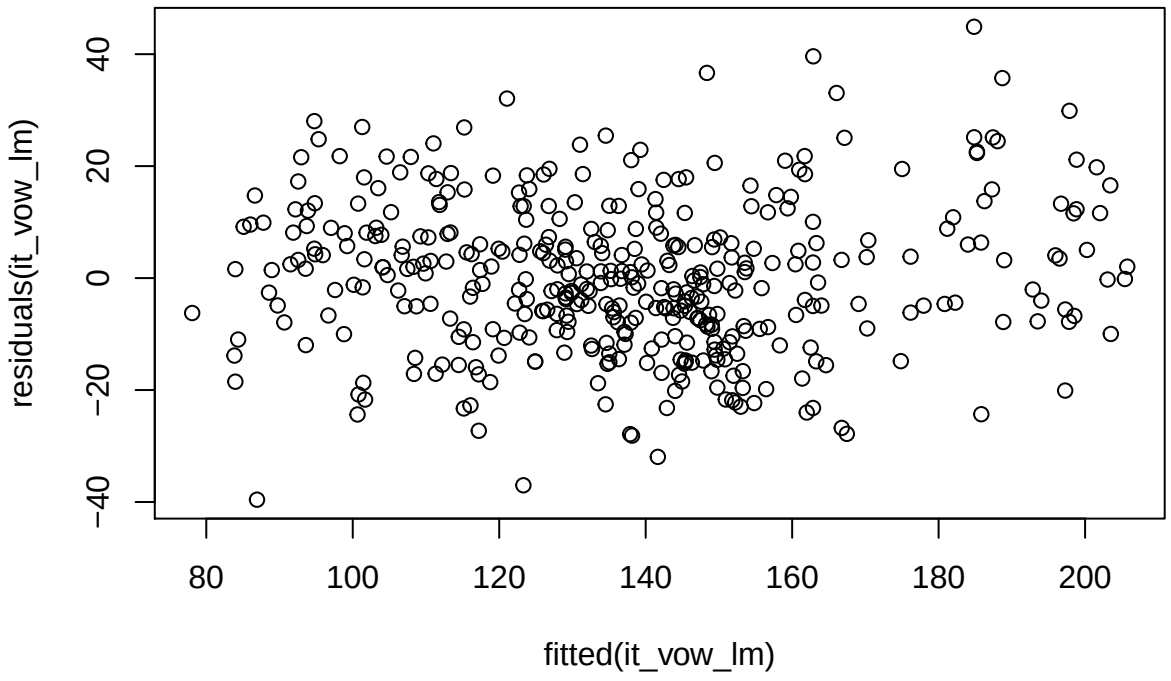
```
it_vow_lm_null <- lmer(
    vowel.duration ~
#         c2phonation +
        c2place +
        vowel +
        sentence.duration +
        (1+c2phonation|speaker) +
        (1|word),
    data = filter(vowels, language == "italian")
)

anova(it_vow_lm_null, it_vow_lm)
```

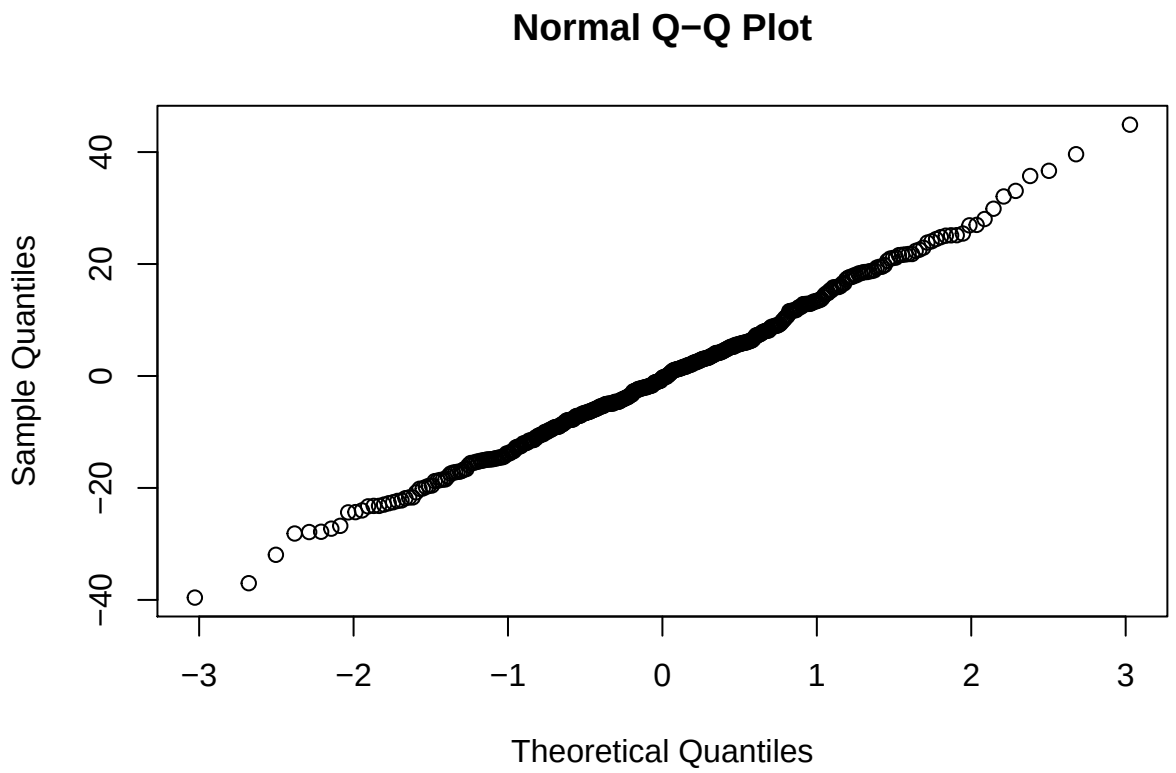
```
## refitting model(s) with ML (instead of REML)
```

```
## Data: filter(vowels, language == "italian")
## Models:
## object: vowel.duration ~ c2place + vowel + sentence.duration + (1 + c2phonation |
## object:     speaker) + (1 | word)
## ..1: vowel.duration ~ c2phonation + c2place + vowel + c2phonation:vowel +
## ..1:     sentence.duration + (1 + c2phonation | speaker) + (1 | word)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object 10 3363.5 3403.6 -1671.8   3343.5
## ..1    13 3352.9 3405.0 -1663.4   3326.9 16.609      3  0.0008505 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

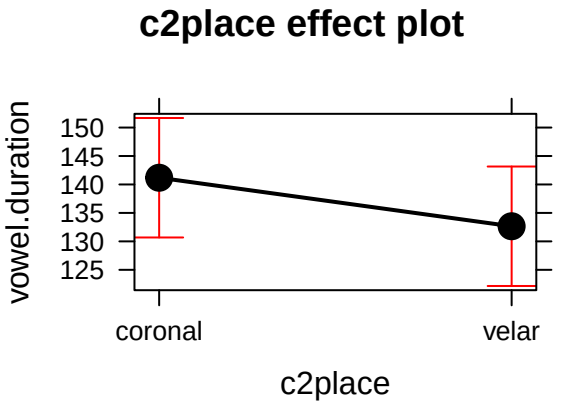
```
plot(fitted(it_vow_lm),residuals(it_vow_lm))
```

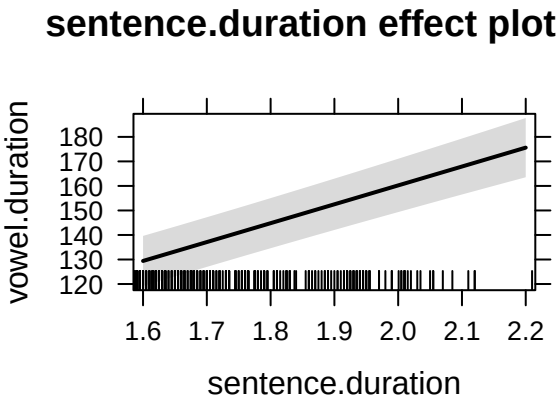


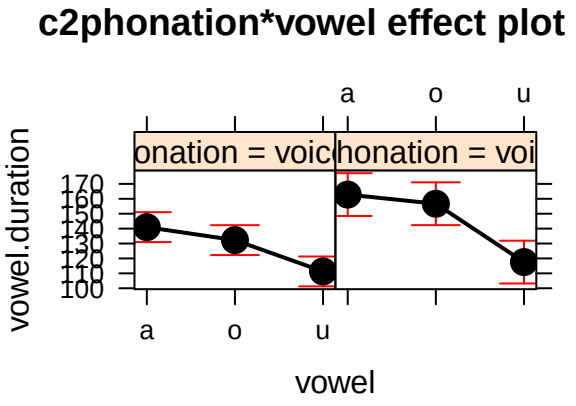
```
qqnorm(resid(it_vow_lm))
```




```
plot(allEffects(it_vow_lm))
```







2.2 Polish

```
pl_vow_lm <- lmer(
    vowel.duration ~
        c2phonation +
        vowel +
        c2place +
        sentence.duration +
        (1+c2phonation|speaker) +
        (1|word),
    data = filter(vowels, language == "polish")
)

summary(pl_vow_lm)
```

```
## Linear mixed model fit by REML t-tests use Satterthwaite approximations
##   to degrees of freedom [lmerMod]
## Formula:
## vowel.duration ~ c2phonation + vowel + c2place + sentence.duration +
##     (1 + c2phonation | speaker) + (1 | word)
##    Data: filter(vowels, language == "polish")
##
## REML criterion at convergence: 2317.9
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.77796 -0.72472  0.00879  0.59730  2.90314
##
## Random effects:
##  Groups   Name              Variance Std.Dev. Corr
##  word     (Intercept)        14.16    3.763
##  speaker  (Intercept)        39.31    6.270
##           c2phonationvoiced  18.37    4.286   1.00
##  Residual                   100.69   10.034
## Number of obs: 311, groups:  word, 12; speaker, 4
##
## Fixed effects:
##                   Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)         22.928     10.437 127.130   2.197  0.02985 *
## c2phonationvoiced    7.881      3.259   6.860   2.418  0.04691 *
## vowelo             -11.795      3.005   7.000  -3.925  0.00571 **
## vowelu             -29.276      3.016   7.100  -9.707 2.37e-05 ***
## c2placevelar        -5.577      2.453   7.000  -2.273  0.05722 .
## sentence.duration   70.813      9.743 261.040   7.268 4.22e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) c2phnt vowelo vowelu c2plcv
## c2phontnvcd  0.139
## vowelo      -0.106  0.001
## vowelu      -0.063  0.002  0.501
## c2placevelr -0.098  0.000  0.001  0.003
## sentnc.drtn -0.917 -0.033 -0.041 -0.087 -0.022
```

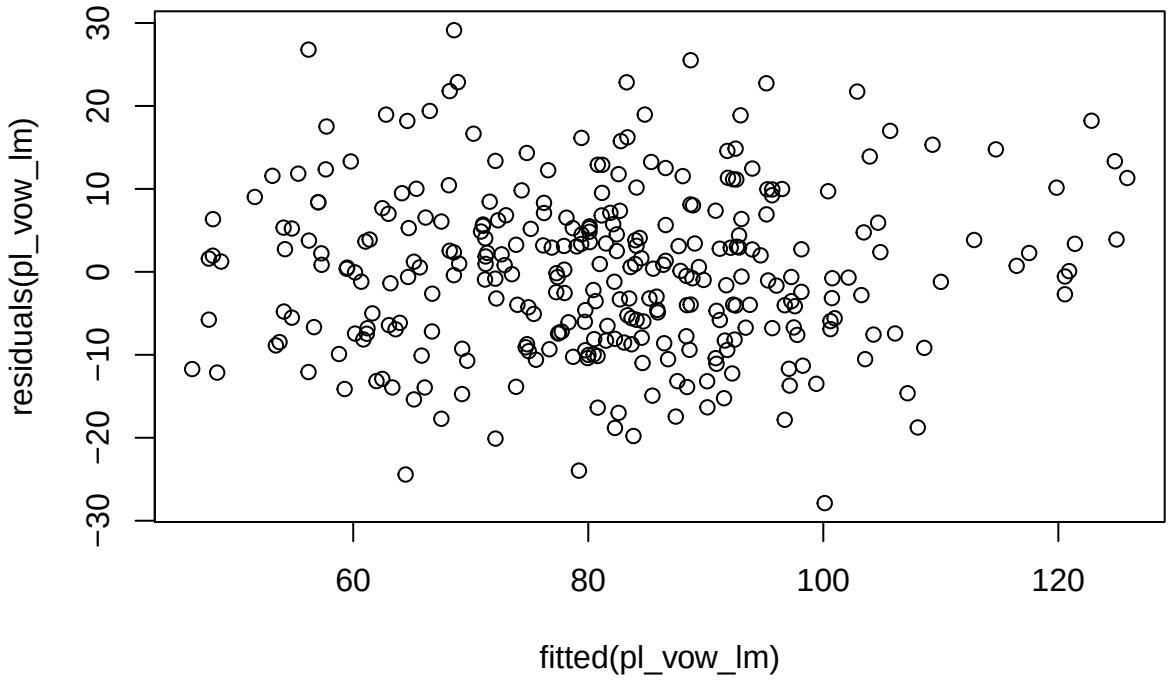
```
pl_vow_lm_null <- lmer(
    vowel.duration ~
#         c2phonation +
        vowel +
        c2place +
        sentence.duration +
        (1+c2phonation|speaker) +
        (1|word),
    data = filter(vowels, language == "polish")
)

anova(pl_vow_lm_null, pl_vow_lm)
```

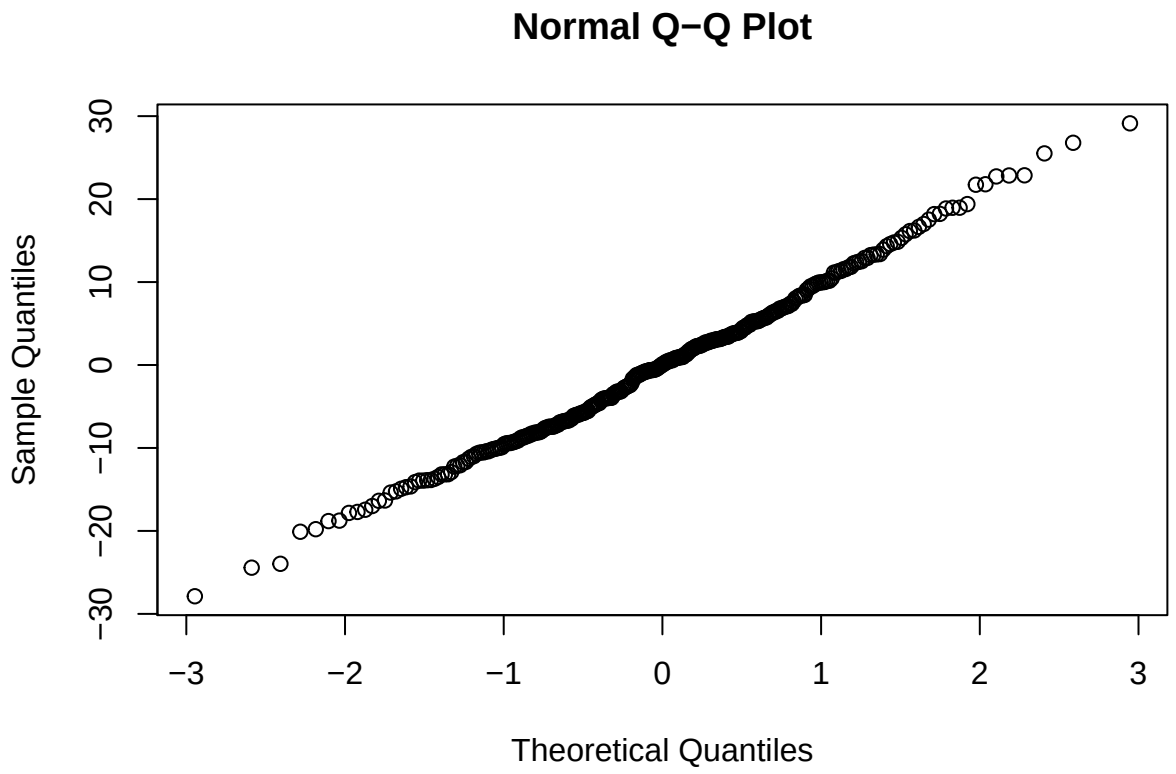
```
## refitting model(s) with ML (instead of REML)
```

```
## Data: filter(vowels, language == "polish")
## Models:
## object: vowel.duration ~ vowel + c2place + sentence.duration + (1 + c2phonation |
## object:     speaker) + (1 | word)
## ..1: vowel.duration ~ c2phonation + vowel + c2place + sentence.duration +
## ..1:     (1 + c2phonation | speaker) + (1 | word)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object 10 2368.4 2405.8 -1174.2   2348.4
## ..1    11 2365.0 2406.2 -1171.5   2343.0 5.3949      1     0.0202 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

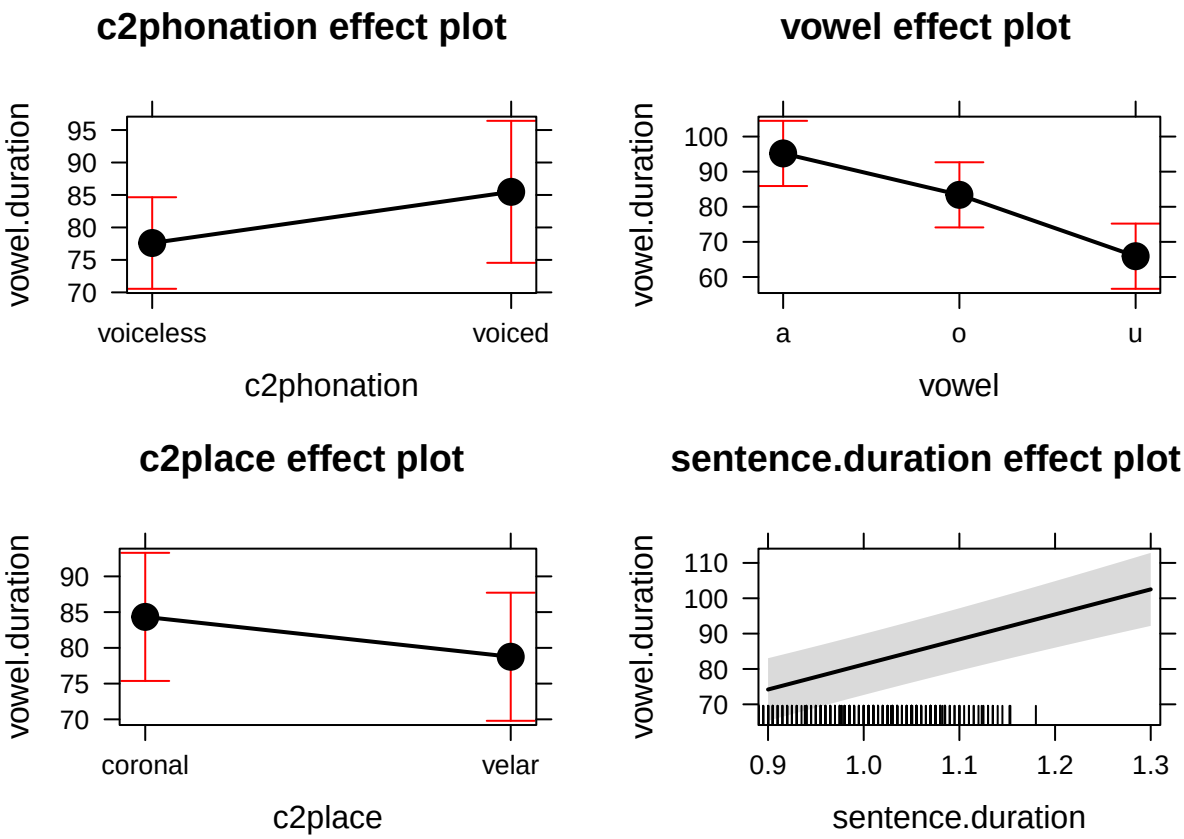
```
plot(fitted(pl_vow_lm),residuals(pl_vow_lm))
```




```
qqnorm(resid(pl_vow_lm))
```




```
plot(allEffects(pl_vow_lm))
```



3 Tongue root advancement at maximum displacement

3.1 Italian

3.1.1 IT01

```
it01_max <- filter(tongues, speaker == "it01")

it01_gamm <- bam(
    Y ~
        X.re +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it01_max,
    method = "fREML"
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
rho <- start_value_rho(it01_gamm)

it01_gamm_ar <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it01_max,
    method = "ML",
    rho = rho,
    AR.start = it01_max$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
summary(it01_gamm_ar)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Parametric coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.73141    0.30682  -15.42   <2e-16 ***
## X            0.76532    0.02666   28.70   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                               edf  Ref.df       F  p-value
## s(X)                        7.764   7.944 181.554  < 2e-16 ***
## s(X):c2phonation.ordvoiced  5.037   6.136   9.480 2.06e-10 ***
## s(X):c2place.ordvelar       8.822   8.979 183.531  < 2e-16 ***
## s(X):vowel.ordo             6.862   7.808  11.566 1.19e-15 ***
## s(X,rec.date)              92.680 225.000   1.976  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 267/268
## R-sq.(adj) =  0.979   Deviance explained = 91.4%
## -ML = 3488.2  Scale est. = 4.0701    n = 1932
```

```
it01_gamm_ar_null <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
#        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it01_max,
    method = "ML",
    rho = rho,
    AR.start = it01_max$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

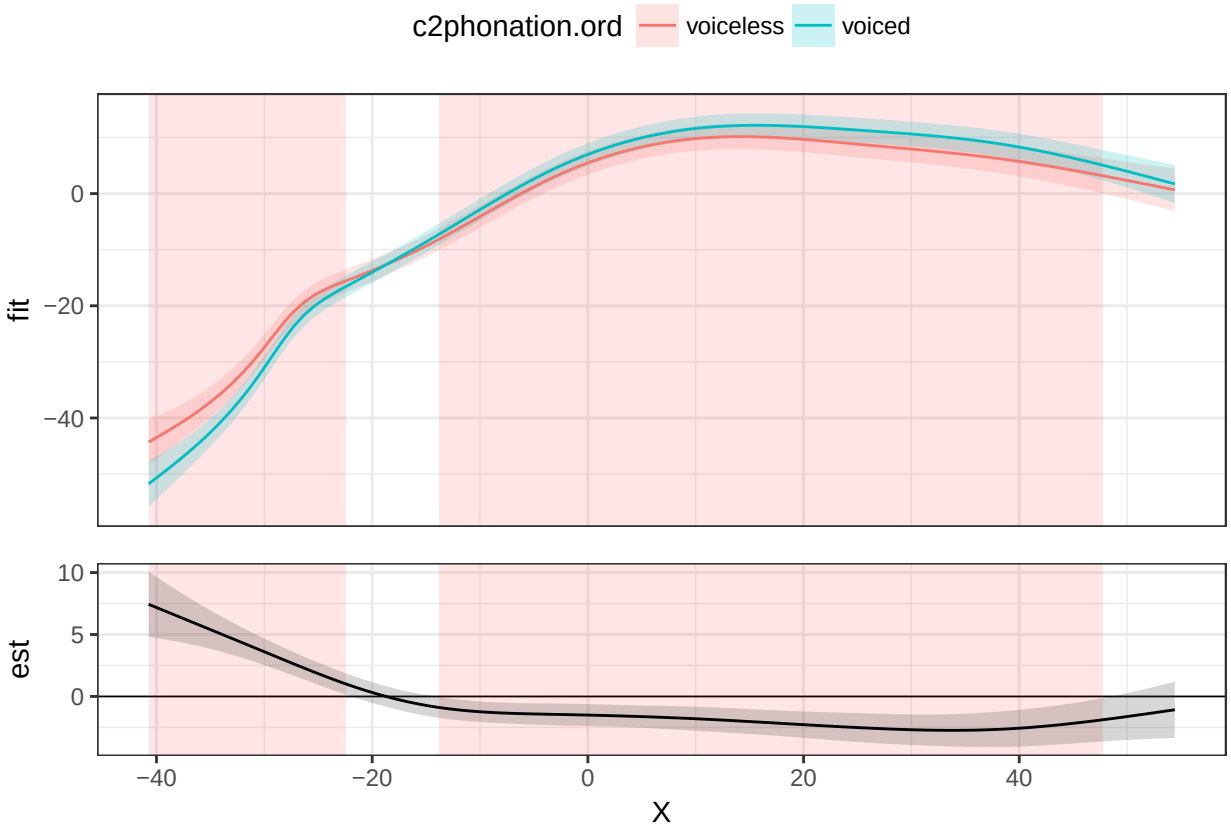
```
compareML(it01_gamm_ar_null, it01_gamm_ar)
```

```
## it01_gamm_ar_null: Y ~ X + s(X, bs = "cr") + s(X, by = c2place.ord, bs = "cr") +
##     s(X, by = vowel.ord, bs = "cr") + s(X, rec.date, bs = "fs",
##     xt = "cr", m = 1, k = 5)
##
## it01_gamm_ar: Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Chi-square test of ML scores
## -----
##               Model    Score Edf Difference    Df   p.value Sig.
## 1 it01_gamm_ar_null 3509.547  11
## 2      it01_gamm_ar 3488.231  13     21.316 2.000 5.527e-10  ***
##
## AIC difference: 23.12, model it01_gamm_ar has lower AIC.
```

```
## Warning in compareML(it01_gamm_ar_null, it01_gamm_ar): AIC might not be
## reliable, as an AR1 model is included (rho1 = 0.745137, rho2 = 0.745137).
```

```
plot_gamsd(
    it01_gamm_ar,
    view = "X",
    comparison = list(c2phonation.ord = c("voiceless", "voiced")),
    conditions = list(c2place.ord = "coronal")
)
```

```
## Summary:
##  * X : numeric predictor; with 100 values ranging from -40.731200 to 54.425500.
##  * c2phonation.ord : factor; set to the value(s): voiced, voiceless.
##  * c2place.ord : factor; set to the value(s): coronal.
##  * vowel.ord : factor; set to the value(s): a.
##  * rec.date : factor; set to the value(s): 29/11/2016 15:10:52.
```



3.1.2 IT02

```
it02_max <- filter(tongues, speaker == "it02")

it02_gamm <- bam(
    Y ~
        X.re +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
```

```
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it02_max,
    method = "fREML"
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
rho <- start_value_rho(it02_gamm)

it02_gamm_ar <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it02_max,
    method = "ML",
    rho = rho,
    AR.start = it02_max$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
summary(it02_gamm_ar)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.67360    0.52437  -3.192  0.00145 **
## X            0.65159    0.03157  20.640  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                               edf  Ref.df      F  p-value
## s(X)                        6.386   7.189 57.403  < 2e-16 ***
## s(X):c2phonation.ordvoiced  6.589   7.613  4.927 9.93e-06 ***
## s(X):c2place.ordvelar       8.630   8.910 52.382  < 2e-16 ***
## s(X):vowel.ordo             5.937   7.020  9.062 6.16e-11 ***
## s(X,rec.date)              72.997 185.000  1.401  < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 227/228
## R-sq.(adj) =  0.916   Deviance explained =   71%
## -ML = 2803.6  Scale est. = 9.778     n = 1233
```

```
it02_gamm_ar_null <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
#        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it02_max,
    method = "ML",
    rho = rho,
    AR.start = it02_max$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
compareML(it02_gamm_ar_null, it02_gamm_ar)
```

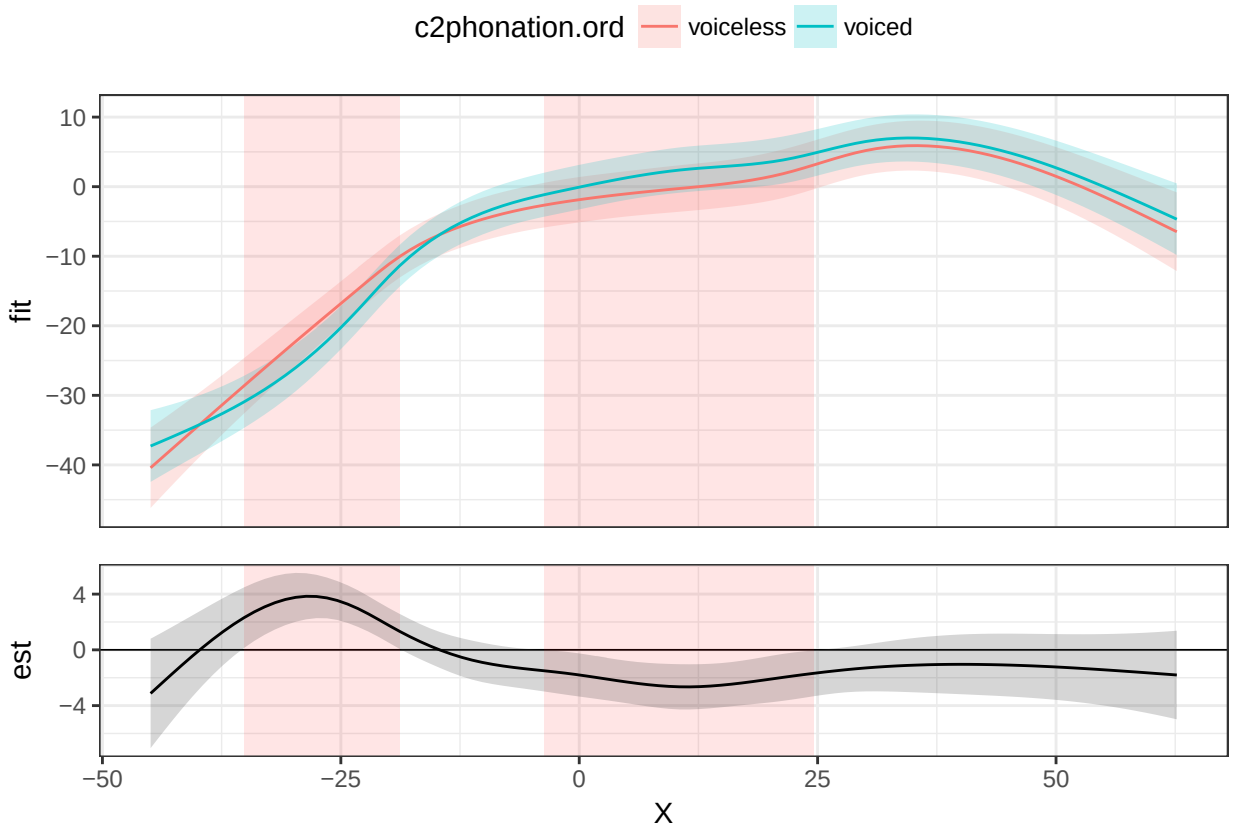
```
## it02_gamm_ar_null: Y ~ X + s(X, bs = "cr") + s(X, by = c2place.ord, bs = "cr") +
##     s(X, by = vowel.ord, bs = "cr") + s(X, rec.date, bs = "fs",
##     xt = "cr", m = 1, k = 5)
##
## it02_gamm_ar: Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Chi-square test of ML scores
## -----
##                Model    Score Edf Difference    Df   p.value Sig.
## 1 it02_gamm_ar_null 2813.488  11
## 2      it02_gamm_ar 2803.592  13      9.897 2.000 5.034e-05  ***
##
## AIC difference: 23.42, model it02_gamm_ar has lower AIC.
```

```
## Warning in compareML(it02_gamm_ar_null, it02_gamm_ar): AIC might not be
## reliable, as an AR1 model is included (rho1 = 0.735050, rho2 = 0.735050).
```

```
plot_gamsd(
    it02_gamm_ar,
    view = "X",
    comparison = list(c2phonation.ord = c("voiceless", "voiced")),
    conditions = list(c2place.ord = "coronal")
)
```

```
## Summary:
##  * X : numeric predictor; with 100 values ranging from -44.949000 to 62.666500.
##  * c2phonation.ord : factor; set to the value(s): voiced, voiceless.
##  * c2place.ord : factor; set to the value(s): coronal.
##  * vowel.ord : factor; set to the value(s): a.
```

```
##   * rec.date : factor; set to the value(s): 12/12/2016 14:45:14.
```



3.2 Polish

3.2.1 PL02

```
pl02_max <- filter(tongues, speaker == "pl02", X > -20)

pl02_gamm <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = pl02_max,
    method = "ML"
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
rho <- start_value_rho(pl02_gamm)

pl02_gamm_ar <- bam(
    Y ~
```

```
        X +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = pl02_max,
    method = "ML",
    rho = rho,
    AR.start = pl02_max$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
summary(pl02_gamm_ar)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.16593    0.50234   6.302 3.52e-10 ***
## X           -0.08800    0.03015  -2.919  0.00355 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                edf  Ref.df       F p-value
## s(X)                         7.112   7.721  53.045 < 2e-16 ***
## s(X):c2phonation.ordvoiced   7.570   8.492   4.417 0.00067 ***
## s(X):c2place.ordvelar        8.821   8.976 209.352 < 2e-16 ***
## s(X):vowel.ordo              7.675   8.550  12.914 < 2e-16 ***
## s(X,rec.date)              135.219 300.000   2.126 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 342/343
## R-sq.(adj) =  0.936   Deviance explained =   75%
## -ML =   4239  Scale est. = 3.4834    n = 2391
```

```
pl02_gamm_ar_null <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
```

```
    data = pl02_max,
    method = "ML"
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

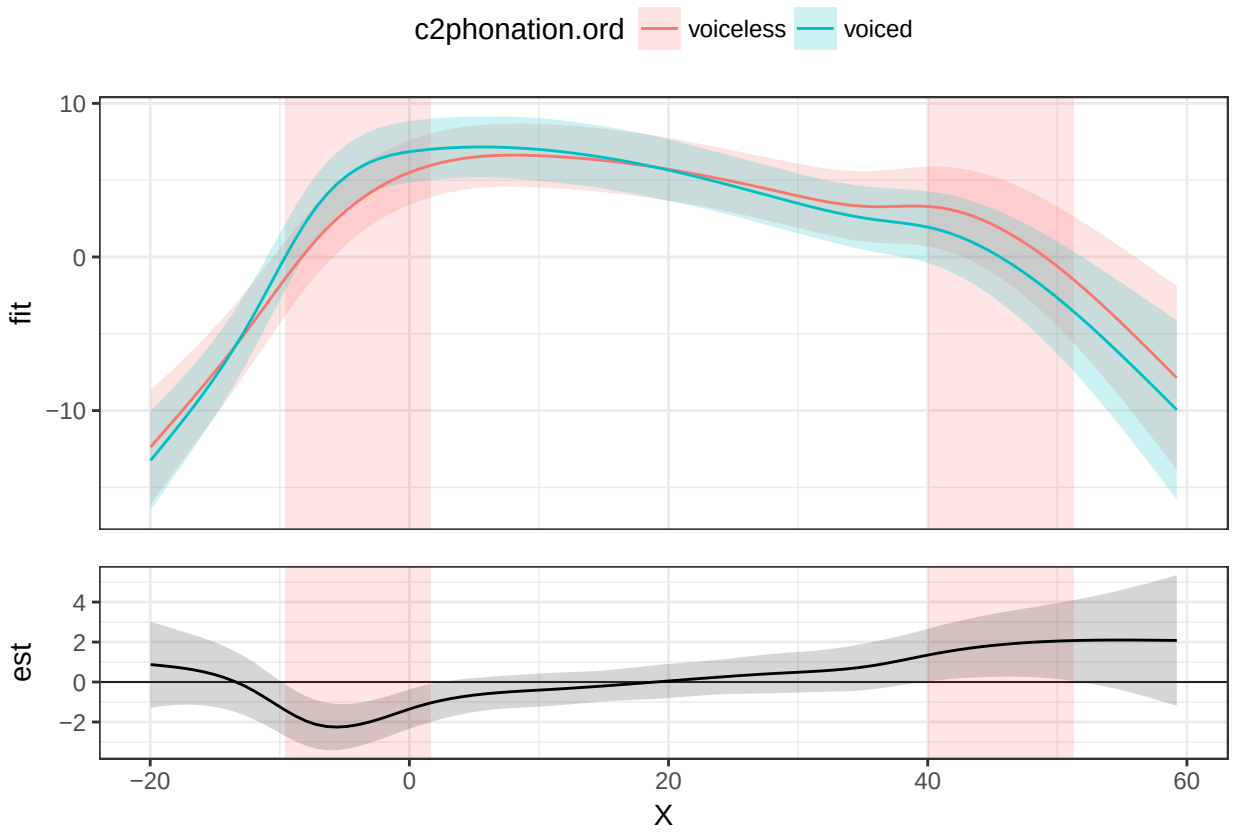
```
compareML(pl02_gamm_ar_null, pl02_gamm_ar)
```

```
## pl02_gamm_ar_null: Y ~ X + s(X, bs = "cr") + s(X, by = c2place.ord, bs = "cr") +
##     s(X, by = vowel.ord, bs = "cr") + s(X, rec.date, bs = "fs",
##     xt = "cr", m = 1, k = 5)
##
## pl02_gamm_ar: Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Chi-square test of ML scores
## -----
##               Model    Score Edf Difference    Df  p.value Sig.
## 1 pl02_gamm_ar_null 5557.001  11
## 2      pl02_gamm_ar 4238.954  13   1318.048 2.000  < 2e-16  ***
##
## AIC difference: 2368.80, model pl02_gamm_ar has lower AIC.
```

```
## Warning in compareML(pl02_gamm_ar_null, pl02_gamm_ar): AIC might not be
## reliable, as an AR1 model is included (rho1 = 0.000000, rho2 = 0.717567).
```

```
plot_gamsd(
    pl02_gamm_ar,
    view = "X",
    comparison = list(c2phonation.ord = c("voiceless", "voiced")),
    conditions = list(c2place.ord = "coronal")
)
```

```
## Summary:
##  * X : numeric predictor; with 100 values ranging from -19.967800 to 59.219800.
##  * c2phonation.ord : factor; set to the value(s): voiced, voiceless.
##  * c2place.ord : factor; set to the value(s): coronal.
##  * vowel.ord : factor; set to the value(s): a.
##  * rec.date : factor; set to the value(s): 07/02/2017 16:29:14.
```



3.2.2 PL04

```
pl04_max <- filter(tongues, speaker == "pl04", X > -25)

pl04_gamm <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = pl04_max,
    method = "ML"
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
rho <- start_value_rho(pl04_gamm)

pl04_gamm_ar <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
```

```
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = pl04_max,
    method = "ML",
    rho = rho,
    AR.start = pl04_max$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
summary(pl04_gamm_ar)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.80103    0.38089   9.979   <2e-16 ***
## X           -0.04730    0.03513  -1.347    0.178
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                edf  Ref.df      F  p-value
## s(X)                         7.436   7.785 21.721  < 2e-16 ***
## s(X):c2phonation.ordvoiced   4.455   5.521  0.650    0.716
## s(X):c2place.ordvelar        8.454   8.853 29.487  < 2e-16 ***
## s(X):vowel.ordo              8.061   8.670  9.105 1.49e-12 ***
## s(X,rec.date)              196.514 230.000 27.988  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 272/273
## R-sq.(adj) =  0.977   Deviance explained = 96.6%
## -ML = 1671.6  Scale est. = 0.65314   n = 1165
```

```
pl04_gamm_ar_null <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = pl04_max,
    method = "ML"
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

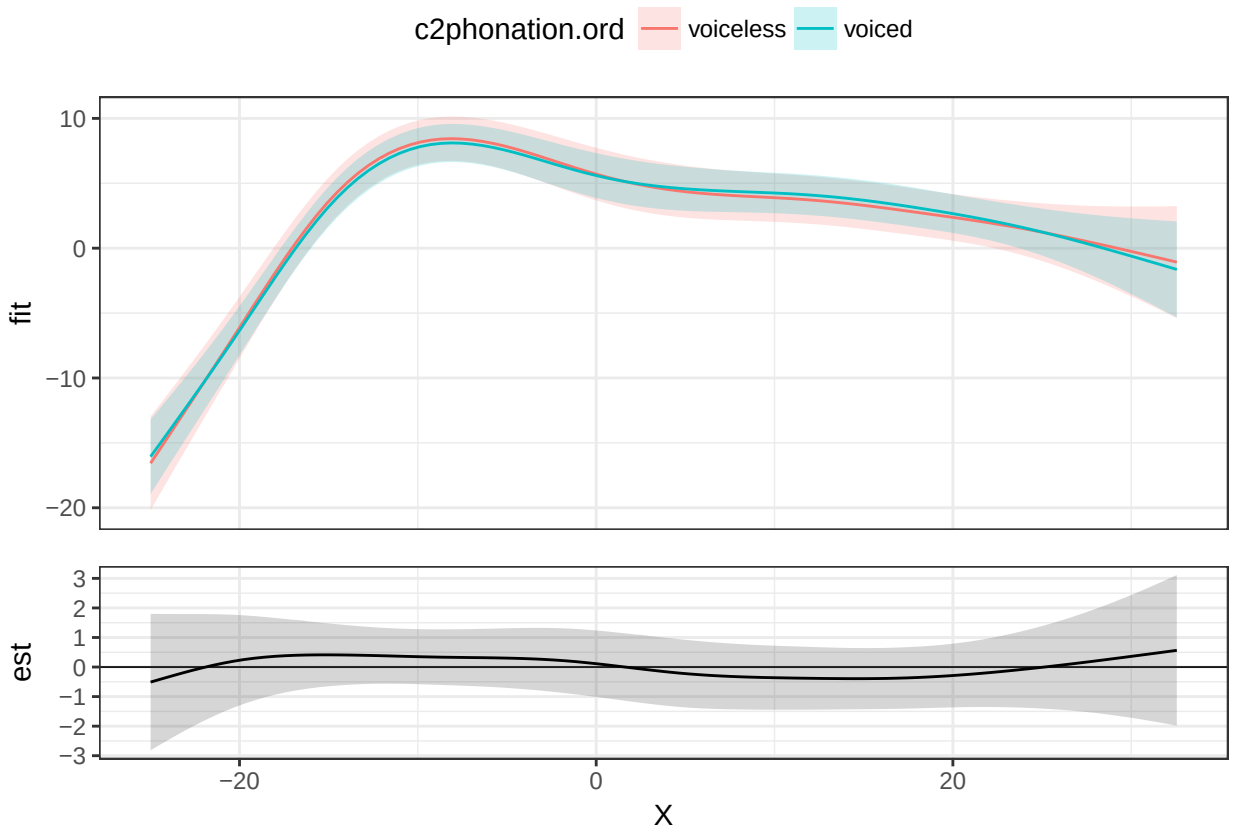
```
compareML(pl04_gamm_ar_null, pl04_gamm_ar)
```

```
## pl04_gamm_ar_null: Y ~ X + s(X, bs = "cr") + s(X, by = c2place.ord, bs = "cr") +
##     s(X, by = vowel.ord, bs = "cr") + s(X, rec.date, bs = "fs",
##     xt = "cr", m = 1, k = 5)
##
## pl04_gamm_ar: Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Chi-square test of ML scores
## -----
##               Model    Score Edf Difference    Df  p.value Sig.
## 1 pl04_gamm_ar_null 1937.166  11
## 2      pl04_gamm_ar 1671.593  13    265.573 2.000  < 2e-16  ***
##
## AIC difference: 352.64, model pl04_gamm_ar has lower AIC.
```

```
## Warning in compareML(pl04_gamm_ar_null, pl04_gamm_ar): AIC might not be
## reliable, as an AR1 model is included (rho1 = 0.000000, rho2 = 0.468632).
```

```
plot_gamsd(
    pl04_gamm_ar,
    view = "X",
    comparison = list(c2phonation.ord = c("voiceless", "voiced")),
    conditions = list(c2place.ord = "coronal")
)
```

```
## Summary:
##  * X : numeric predictor; with 100 values ranging from -24.985100 to 32.556800.
##  * c2phonation.ord : factor; set to the value(s): voiced, voiceless.
##  * c2place.ord : factor; set to the value(s): coronal.
##  * vowel.ord : factor; set to the value(s): a.
##  * rec.date : factor; set to the value(s): 26/05/2017 17:58:11.
```



4 Tongue root advancement at closure

4.1 Italian

4.1.1 IT01

```
it01_clos <- filter(tongues_clos, speaker == "it01")

it01_gamm <- bam(
    Y ~
        X.re +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it01_clos,
    method = "fREML"
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
rho <- start_value_rho(it01_gamm)
```

```r
it01_gamm_ar <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it01_clos,
    method = "ML",
    rho = rho,
    AR.start = it01_clos$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```r
summary(it01_gamm_ar)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.61066    0.39198   1.558    0.119
## X            0.84888    0.02509  33.831   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                               edf  Ref.df       F  p-value
## s(X)                        7.549   7.867 198.722  < 2e-16 ***
## s(X):c2phonation.ordvoiced  4.683   5.594  13.824 3.56e-14 ***
## s(X):c2place.ordvelar       8.641   8.924 174.974  < 2e-16 ***
## s(X):vowel.ordo             6.049   6.977  14.770  < 2e-16 ***
## s(X,rec.date)              96.251 230.000   1.789  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 272/273
## R-sq.(adj) =  0.981   Deviance explained = 93.6%
## -ML = 2983.9  Scale est. = 3.7241    n = 1648
```

```r
it01_gamm_ar_null <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
#        s(X, by = c2phonation.ord, bs = "cr") +
```

```
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it01_clos,
    method = "ML",
    rho = rho,
    AR.start = it01_clos$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

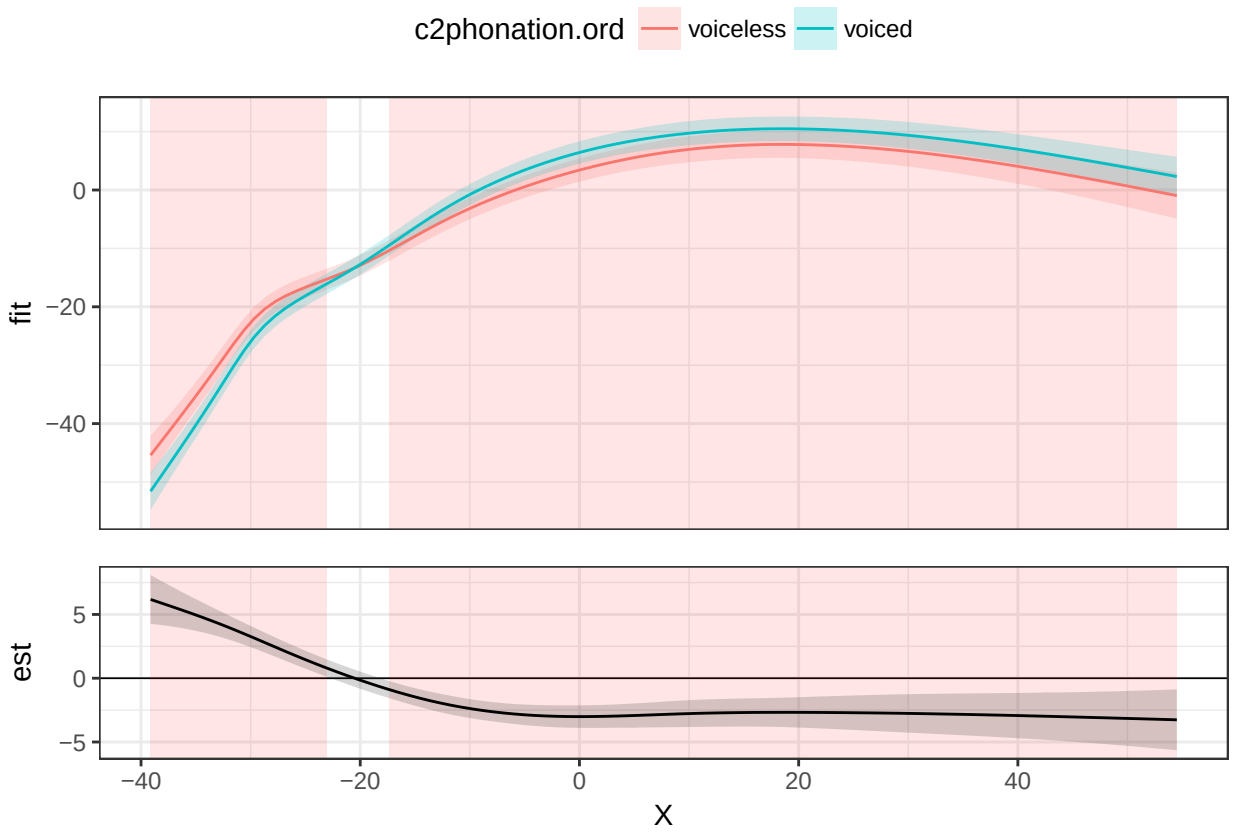
```
compareML(it01_gamm_ar_null, it01_gamm_ar)
```

```
## it01_gamm_ar_null: Y ~ X + s(X, bs = "cr") + s(X, by = c2place.ord, bs = "cr") +
##     s(X, by = vowel.ord, bs = "cr") + s(X, rec.date, bs = "fs",
##     xt = "cr", m = 1, k = 5)
##
## it01_gamm_ar: Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Chi-square test of ML scores
## -----
##               Model    Score Edf Difference    Df   p.value Sig.
## 1 it01_gamm_ar_null 3014.072  11
## 2      it01_gamm_ar 2983.854  13     30.218 2.000 7.523e-14  ***
##
## AIC difference: 28.11, model it01_gamm_ar has lower AIC.
```

```
## Warning in compareML(it01_gamm_ar_null, it01_gamm_ar): AIC might not be
## reliable, as an AR1 model is included (rho1 = 0.721734, rho2 = 0.721734).
```

```
plot_gamsd(
    it01_gamm_ar,
    view = "X",
    comparison = list(c2phonation.ord = c("voiceless", "voiced")),
    conditions = list(c2place.ord = "coronal")
)
```

```
## Summary:
##  * X : numeric predictor; with 100 values ranging from -39.125800 to 54.502000.
##  * c2phonation.ord : factor; set to the value(s): voiced, voiceless.
##  * c2place.ord : factor; set to the value(s): coronal.
##  * vowel.ord : factor; set to the value(s): a.
##  * rec.date : factor; set to the value(s): 29/11/2016 15:10:52.
```



4.1.2 IT02

```
it02_clos <- filter(tongues_clos, speaker == "it02")

it02_gamm <- bam(
    Y ~
        X.re +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it02_clos,
    method = "fREML"
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
rho <- start_value_rho(it02_gamm)

it02_gamm_ar <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
```

```
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it02_clos,
    method = "ML",
    rho = rho,
    AR.start = it02_clos$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
summary(it02_gamm_ar)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.56479    0.48605  -5.277 1.57e-07 ***
## X            0.64094    0.04079  15.714  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                edf  Ref.df      F  p-value
## s(X)                         7.047   7.612 40.215  < 2e-16 ***
## s(X):c2phonation.ordvoiced   7.191   8.131  5.061 2.41e-06 ***
## s(X):c2place.ordvelar        8.088   8.702 29.725  < 2e-16 ***
## s(X):vowel.ordo              8.257   8.776 10.910 2.49e-15 ***
## s(X,rec.date)              115.667 195.000  2.711  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 237/238
## R-sq.(adj) =   0.92   Deviance explained = 77.4%
## -ML = 2848.9  Scale est. = 6.7751    n = 1319
```

```
it02_gamm_ar_null <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
#        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it02_clos,
    method = "ML",
    rho = rho,
```

```
    AR.start = it02_clos$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

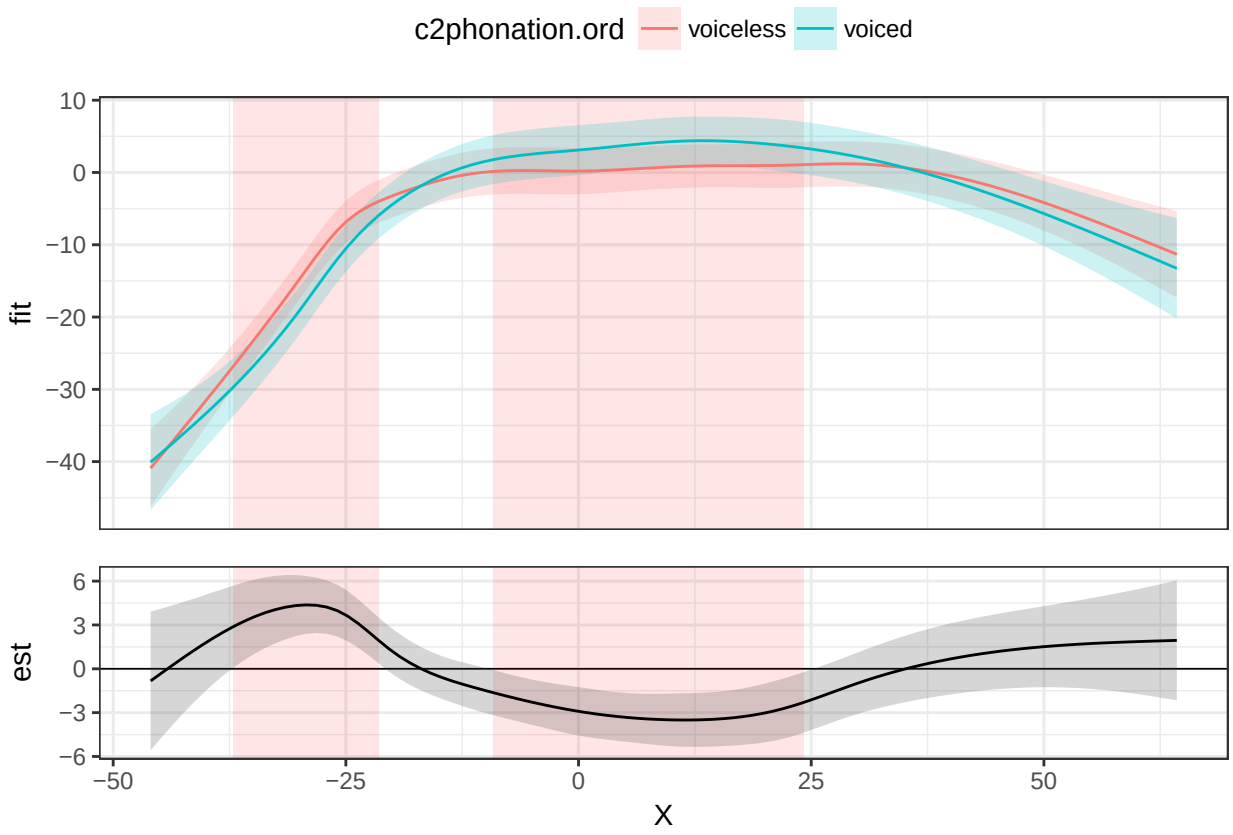
```
compareML(it02_gamm_ar_null, it02_gamm_ar)
```

```
## it02_gamm_ar_null: Y ~ X + s(X, bs = "cr") + s(X, by = c2place.ord, bs = "cr") +
##     s(X, by = vowel.ord, bs = "cr") + s(X, rec.date, bs = "fs",
##     xt = "cr", m = 1, k = 5)
##
## it02_gamm_ar: Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Chi-square test of ML scores
## -----
##               Model    Score Edf Difference    Df   p.value Sig.
## 1 it02_gamm_ar_null 2861.022  11
## 2      it02_gamm_ar 2848.944  13     12.078 2.000 5.680e-06  ***
##
## AIC difference: 18.39, model it02_gamm_ar has lower AIC.
```

```
## Warning in compareML(it02_gamm_ar_null, it02_gamm_ar): AIC might not be
## reliable, as an AR1 model is included (rho1 = 0.718876, rho2 = 0.718876).
```

```
plot_gamsd(
    it02_gamm_ar,
    view = "X",
    comparison = list(c2phonation.ord = c("voiceless", "voiced")),
    conditions = list(c2place.ord = "coronal")
)
```

```
## Summary:
##  * X : numeric predictor; with 100 values ranging from -45.980000 to 64.280100.
##  * c2phonation.ord : factor; set to the value(s): voiced, voiceless.
##  * c2place.ord : factor; set to the value(s): coronal.
##  * vowel.ord : factor; set to the value(s): o.
##  * rec.date : factor; set to the value(s): 12/12/2016 14:44:52.
```



4.2 Polish

4.2.1 PL02

```
pl02_clos <- filter(tongues_clos, speaker == "pl02", X > -20)

pl02_gamm <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = pl02_clos,
    method = "ML"
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
rho <- start_value_rho(pl02_gamm)

pl02_gamm_ar <- bam(
    Y ~
        X +
```

```
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = pl02_clos,
    method = "ML",
    rho = rho,
    AR.start = pl02_clos$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
summary(pl02_gamm_ar)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.32290    0.39224  13.570   <2e-16 ***
## X           -0.04301    0.03666  -1.173    0.241
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                edf  Ref.df       F  p-value
## s(X)                         6.573   7.431  15.117  < 2e-16 ***
## s(X):c2phonation.ordvoiced   1.006   1.009   0.372 0.544001
## s(X):c2place.ordvelar        8.705   8.935 122.916  < 2e-16 ***
## s(X):vowel.ordo              4.822   5.916   4.676 0.000103 ***
## s(X,rec.date)              173.512 315.000   5.860  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 357/358
## R-sq.(adj) =  0.913   Deviance explained = 85.6%
## -ML =   3918  Scale est. = 3.9958    n = 1844
```

```
pl02_gamm_ar_null <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = pl02_clos,
```

```
    method = "ML",
    rho = rho,
    AR.start = pl02_clos$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
compareML(pl02_gamm_ar_null, pl02_gamm_ar)
```

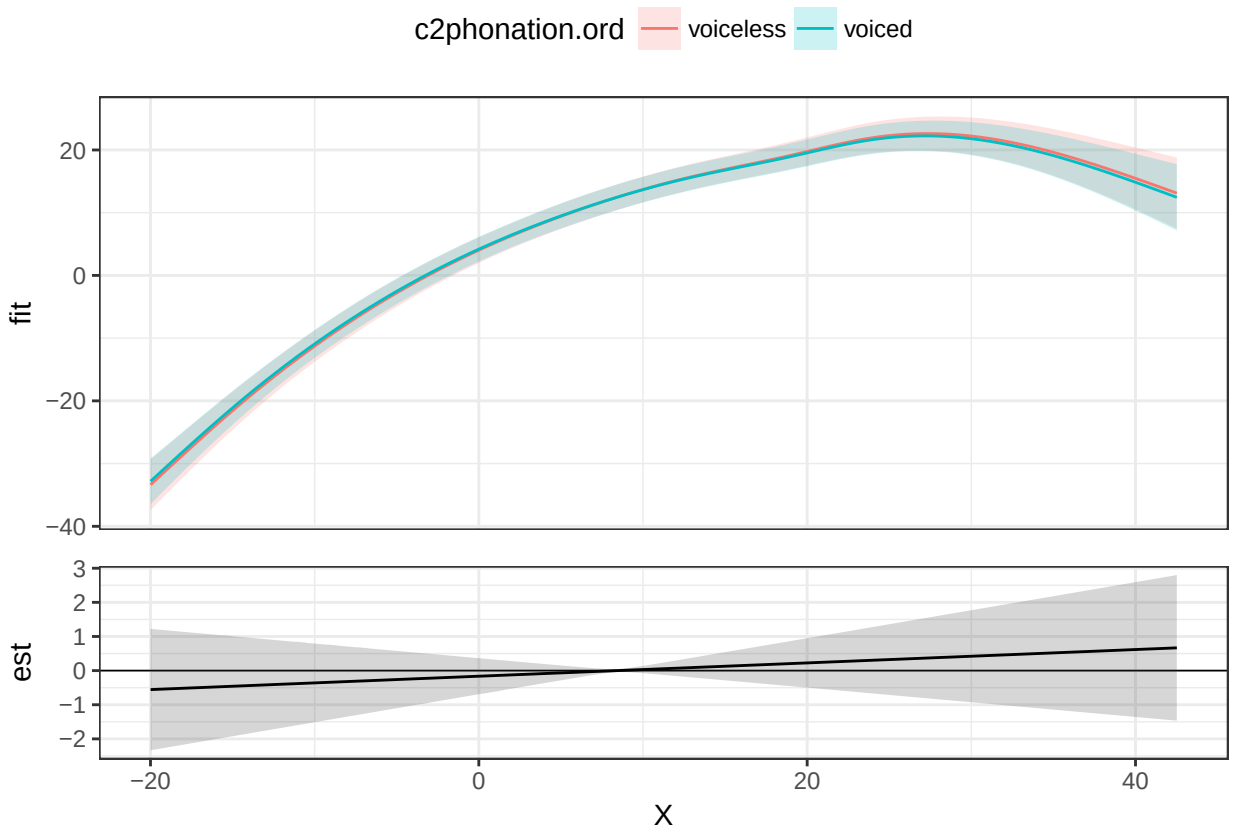
```
## pl02_gamm_ar_null: Y ~ X + s(X, bs = "cr") + s(X, by = c2place.ord, bs = "cr") +
##     s(X, by = vowel.ord, bs = "cr") + s(X, rec.date, bs = "fs",
##     xt = "cr", m = 1, k = 5)
##
## pl02_gamm_ar: Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Chi-square test of ML scores
## -----
##               Model    Score Edf Difference    Df p.value Sig.
## 1 pl02_gamm_ar_null 3918.139  11
## 2      pl02_gamm_ar 3917.953  13      0.186 2.000   0.830
##
## AIC difference: 0.01, model pl02_gamm_ar has lower AIC.
```

```
## Warning in compareML(pl02_gamm_ar_null, pl02_gamm_ar): AIC might not be
## reliable, as an AR1 model is included (rho1 = 0.460962, rho2 = 0.460962).
```

```
## Warning in compareML(pl02_gamm_ar_null, pl02_gamm_ar): Only small difference in ML...
```

```
plot_gamsd(
    pl02_gamm_ar,
    view = "X",
    comparison = list(c2phonation.ord = c("voiceless", "voiced")),
    conditions = list(c2place.ord = "coronal")
)
```

```
## Summary:
##  * X : numeric predictor; with 100 values ranging from -19.993600 to 42.515500.
##  * c2phonation.ord : factor; set to the value(s): voiced, voiceless.
##  * c2place.ord : factor; set to the value(s): coronal.
##  * vowel.ord : factor; set to the value(s): a.
##  * rec.date : factor; set to the value(s): 07/02/2017 16:21:39.
```



4.2.2 PL04

```
pl04_clos <- filter(tongues_clos, speaker == "pl04", X > -25)

pl04_gamm <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = pl04_clos,
    method = "ML"
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
rho <- start_value_rho(pl04_gamm)

pl04_gamm_ar <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
```

```
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = pl04_clos,
    method = "ML",
    rho = rho,
    AR.start = pl04_clos$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
summary(pl04_gamm_ar)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.82123    0.12145   6.762 2.28e-11 ***
## X           -0.14649    0.01221 -11.995  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                edf  Ref.df       F  p-value
## s(X)                         7.910   7.978 231.653  < 2e-16 ***
## s(X):c2phonation.ordvoiced   4.587   5.667   1.006    0.355
## s(X):c2place.ordvelar        6.586   7.722   6.921 1.44e-08 ***
## s(X):vowel.ordo              8.638   8.943  61.281  < 2e-16 ***
## s(X,rec.date)              177.603 235.000  16.079  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 277/278
## R-sq.(adj) =   0.99   Deviance explained = 98.2%
## -ML = 924.74  Scale est. = 0.20121   n = 1237
```

```
pl04_gamm_ar_null <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, rec.date, bs = "fs", xt = "cr", m = 1, k = 5),
    data = pl04_clos,
    method = "ML",
    rho = rho,
    AR.start = pl04_clos$start.event
```

```
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
compareML(pl04_gamm_ar_null, pl04_gamm_ar)
```

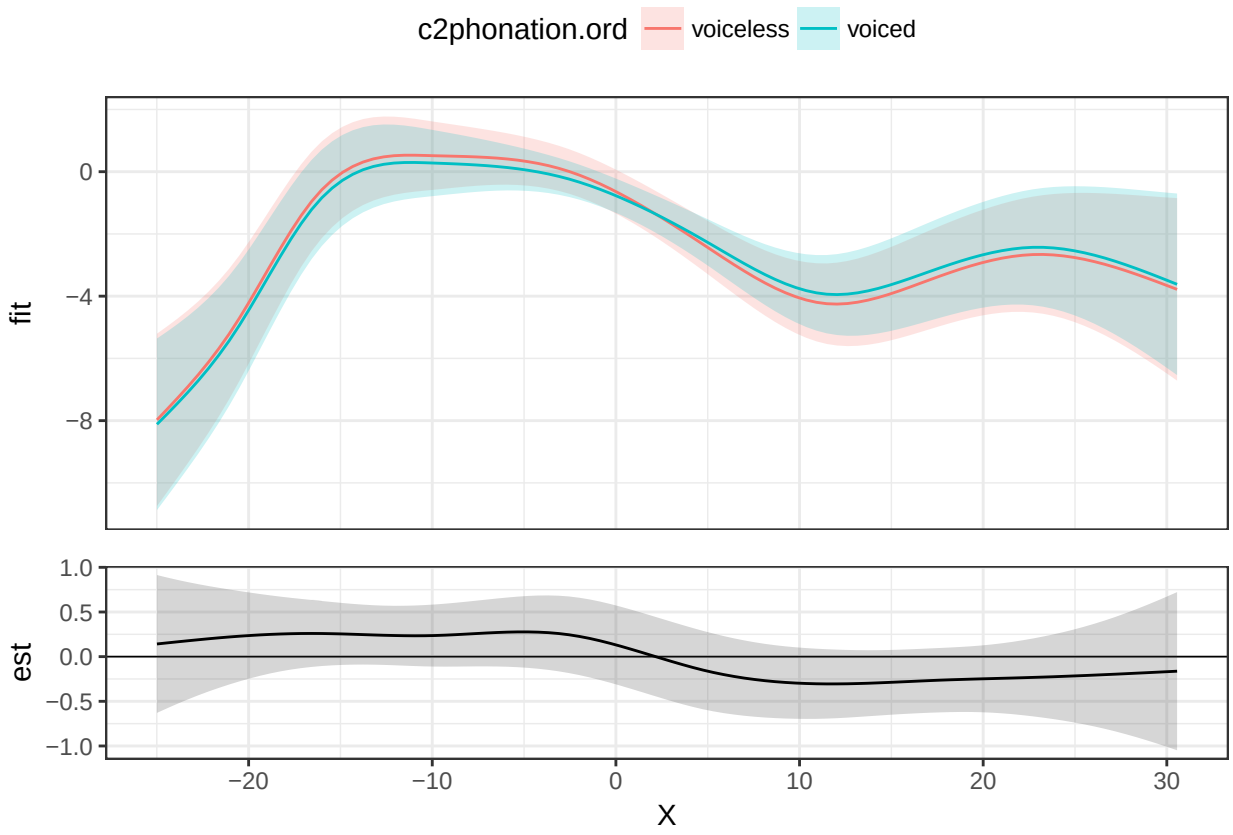
```
## pl04_gamm_ar_null: Y ~ X + s(X, bs = "cr") + s(X, by = c2place.ord, bs = "cr") +
##     s(X, by = vowel.ord, bs = "cr") + s(X, rec.date, bs = "fs",
##     xt = "cr", m = 1, k = 5)
##
## pl04_gamm_ar: Y ~ X + s(X, bs = "cr") + s(X, by = c2phonation.ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, rec.date, bs = "fs", xt = "cr", m = 1,
##     k = 5)
##
## Chi-square test of ML scores
## -----
##                Model    Score Edf Difference    Df p.value Sig.
## 1 pl04_gamm_ar_null 926.0129  11
## 2      pl04_gamm_ar 924.7407  13      1.272 2.000   0.280
##
## AIC difference: 3.22, model pl04_gamm_ar has lower AIC.
```

```
## Warning in compareML(pl04_gamm_ar_null, pl04_gamm_ar): AIC might not be
## reliable, as an AR1 model is included (rho1 = 0.472434, rho2 = 0.472434).
```

```
## Warning in compareML(pl04_gamm_ar_null, pl04_gamm_ar): Only small difference in ML...
```

```
plot_gamsd(
    pl04_gamm_ar,
    view = "X",
    comparison = list(c2phonation.ord = c("voiceless", "voiced")),
    conditions = list(c2place.ord = "coronal")
)
```

```
## Summary:
##  * X : numeric predictor; with 100 values ranging from -24.991700 to 30.560600.
##  * c2phonation.ord : factor; set to the value(s): voiced, voiceless.
##  * c2place.ord : factor; set to the value(s): coronal.
##  * vowel.ord : factor; set to the value(s): a.
##  * rec.date : factor; set to the value(s): 26/05/2017 18:00:08.
```



5 Comparison tongue at closure and maximum displacement in Italian

5.1 IT01

```
it01_voiced <- rbind(it01_max, it01_clos) %>%
    filter(c2phonation == "voiced") %>%
    mutate(
        position = ifelse(label %in% c("max_TT", "max_TD"), "maximum", "closure"),
        position_ord = ordered(position, levels = c("maximum", "closure"))
    ) %>%
    unite(item_no, seconds:rec.date) %>%
    mutate_if(is.character, as.factor)
```

```
it01_voiced_gamm <- bam(
    Y ~
        X.re +
        s(X, bs = "cr") +
        s(X, by = position_ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, item_no, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it01_voiced,
    method = "fREML"
```

```
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
rho <- start_value_rho(it01_voiced_gamm)

it01_voiced_gamm_ar <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = position_ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, item_no, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it01_voiced,
    method = "ML",
    rho = rho,
    AR.start = it01_voiced$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```
summary(it01_voiced_gamm_ar)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Y ~ X + s(X, bs = "cr") + s(X, by = position_ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, item_no, bs = "fs", xt = "cr", m = 1, k = 5)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.06151    0.32749  -6.295 3.92e-10 ***
## X            0.96067    0.02656  36.166  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                               edf  Ref.df       F  p-value
## s(X)                        7.594   7.881 313.331  < 2e-16 ***
## s(X):position_ordclosure    3.981   4.904  11.740 6.68e-11 ***
## s(X):c2place.ordvelar       8.186   8.585 219.239  < 2e-16 ***
## s(X):vowel.ordo             6.676   7.561  16.233  < 2e-16 ***
## s(X,item_no)               79.093 225.000   1.544  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 267/268
## R-sq.(adj) =   0.98   Deviance explained = 92.5%
## -ML = 3204.3  Scale est. = 3.8788    n = 1776
```

```
it01_voiced_gamm_ar_null <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
#        s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, item_no, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it01_voiced,
    method = "ML",
    rho = rho,
    AR.start = it01_voiced$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

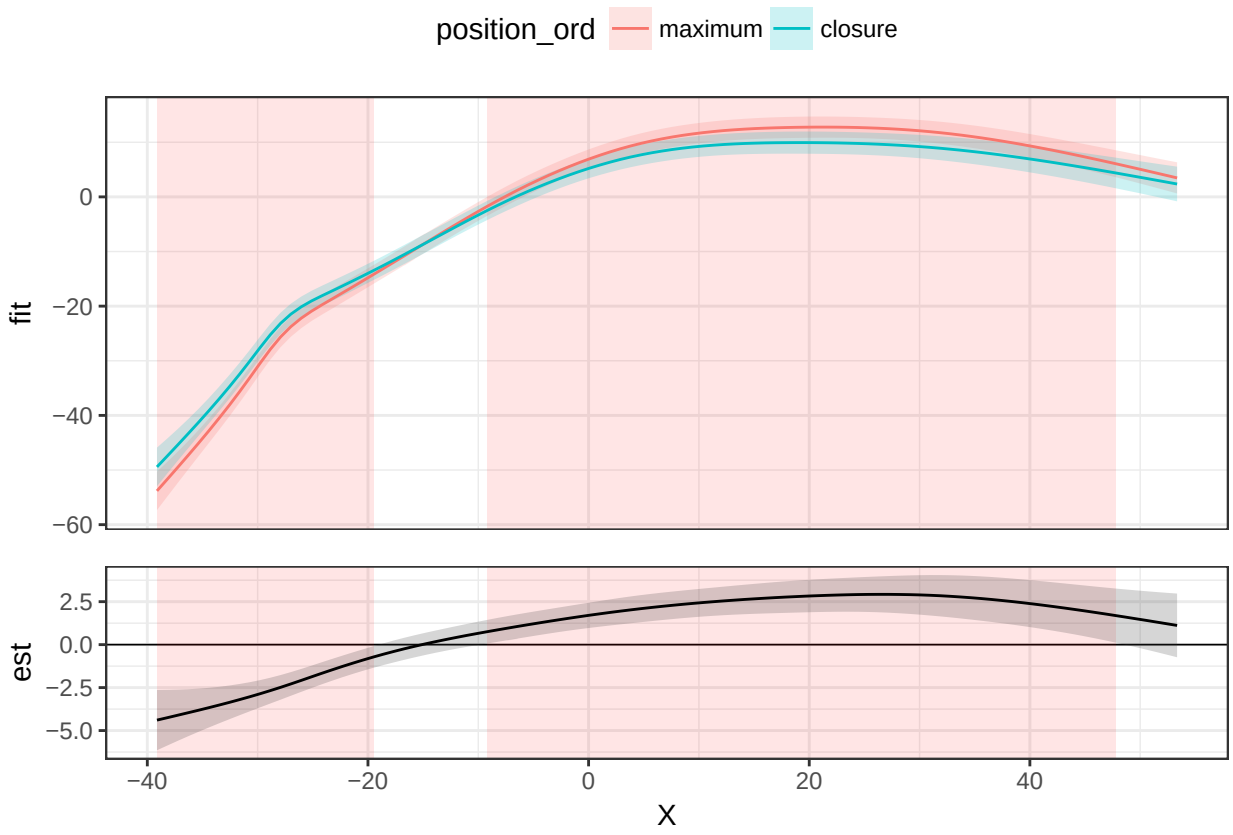
```
compareML(it01_voiced_gamm_ar_null, it01_voiced_gamm_ar)
```

```
## it01_voiced_gamm_ar_null: Y ~ X + s(X, bs = "cr") + s(X, by = c2place.ord, bs = "cr") +
##     s(X, by = vowel.ord, bs = "cr") + s(X, item_no, bs = "fs",
##     xt = "cr", m = 1, k = 5)
##
## it01_voiced_gamm_ar: Y ~ X + s(X, bs = "cr") + s(X, by = position_ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, item_no, bs = "fs", xt = "cr", m = 1, k = 5)
##
## Chi-square test of ML scores
## -----
##                       Model    Score Edf Difference    Df   p.value Sig.
## 1 it01_voiced_gamm_ar_null 3224.556  11
## 2      it01_voiced_gamm_ar 3204.308  13     20.248 2.000 1.608e-09  ***
##
## AIC difference: 8.12, model it01_voiced_gamm_ar has lower AIC.
```

```
## Warning in compareML(it01_voiced_gamm_ar_null, it01_voiced_gamm_ar): AIC
## might not be reliable, as an AR1 model is included (rho1 = 0.727873, rho2 =
## 0.727873).
```

```
plot_gamsd(
    it01_voiced_gamm_ar,
    view = "X",
    comparison = list(position_ord = c("maximum", "closure")),
    conditions = list(c2place.ord = "coronal")
)
```

```
## Summary:
##  * X : numeric predictor; with 100 values ranging from -39.125800 to 53.346500.
##  * position_ord : factor; set to the value(s): closure, maximum.
##  * c2place.ord : factor; set to the value(s): coronal.
##  * vowel.ord : factor; set to the value(s): a.
##  * item_no : factor; set to the value(s): 1.0035_29/11/2016 15:16:05.
```



5.2 IT02

```
it02_voiced <- rbind(it02_max, it02_clos) %>%
    filter(c2phonation == "voiced") %>%
    mutate(
        position = ifelse(label %in% c("max_TT", "max_TD"), "maximum", "closure"),
        position_ord = ordered(position, levels = c("maximum", "closure"))
    ) %>%
    unite(item_no, seconds:rec.date) %>%
    mutate_if(is.character, as.factor)
```

```
it02_voiced_gamm <- bam(
    Y ~
        X.re +
        s(X, bs = "cr") +
        s(X, by = position_ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, item_no, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it02_voiced,
    method = "fREML"
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```r
rho <- start_value_rho(it02_voiced_gamm)

it02_voiced_gamm_ar <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
        s(X, by = position_ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, item_no, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it02_voiced,
    method = "ML",
    rho = rho,
    AR.start = it02_voiced$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

```r
summary(it02_voiced_gamm_ar)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Y ~ X + s(X, bs = "cr") + s(X, by = position_ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, item_no, bs = "fs", xt = "cr", m = 1, k = 5)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.97114    0.44076  -6.741 2.46e-11 ***
## X            0.57698    0.03402  16.962  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                              edf  Ref.df      F  p-value
## s(X)                       7.127   7.649 62.981  < 2e-16 ***
## s(X):position_ordclosure   6.504   7.557  5.531 1.85e-06 ***
## s(X):c2place.ordvelar      8.345   8.814 51.634  < 2e-16 ***
## s(X):vowel.ordo            6.660   7.696 13.608  < 2e-16 ***
## s(X,item_no)              82.051 190.000  1.387  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 232/233
## R-sq.(adj) =  0.911   Deviance explained = 69.8%
## -ML = 2913.1  Scale est. = 9.7551    n = 1286
```

```r
it02_voiced_gamm_ar_null <- bam(
    Y ~
        X +
        s(X, bs = "cr") +
```

```
#         s(X, by = c2phonation.ord, bs = "cr") +
        s(X, by = c2place.ord, bs = "cr") +
        s(X, by = vowel.ord, bs = "cr") +
        s(X, item_no, bs = "fs", xt = "cr", m = 1, k = 5),
    data = it02_voiced,
    method = "ML",
    rho = rho,
    AR.start = it02_voiced$start.event
)
```

```
## Warning in gam.side(sm, X, tol = .Machine$double.eps^0.5): model has
## repeated 1-d smooths of same variable.
```

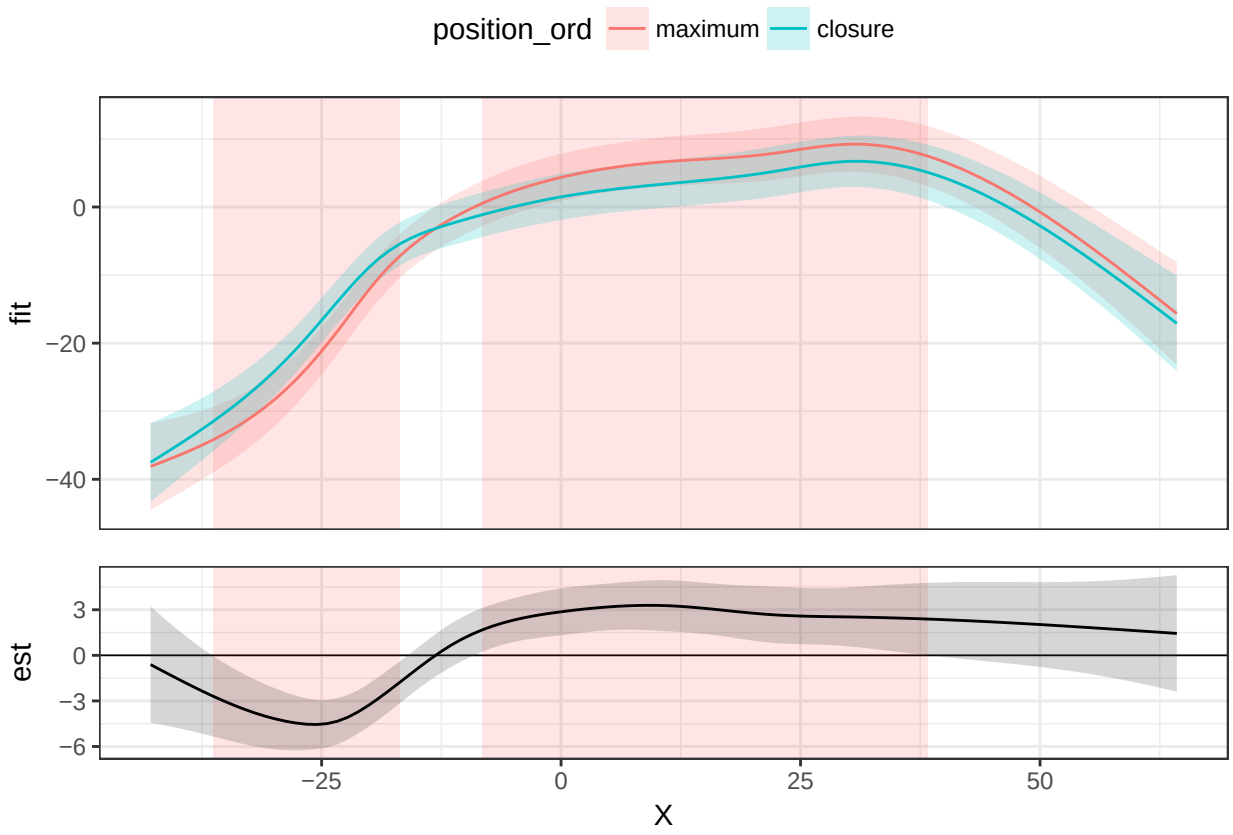
```
compareML(it02_voiced_gamm_ar_null, it02_voiced_gamm_ar)
```

```
## it02_voiced_gamm_ar_null: Y ~ X + s(X, bs = "cr") + s(X, by = c2place.ord, bs = "cr") +
##     s(X, by = vowel.ord, bs = "cr") + s(X, item_no, bs = "fs",
##     xt = "cr", m = 1, k = 5)
##
## it02_voiced_gamm_ar: Y ~ X + s(X, bs = "cr") + s(X, by = position_ord, bs = "cr") +
##     s(X, by = c2place.ord, bs = "cr") + s(X, by = vowel.ord,
##     bs = "cr") + s(X, item_no, bs = "fs", xt = "cr", m = 1, k = 5)
##
## Chi-square test of ML scores
## -----
##                       Model   Score Edf Difference    Df   p.value Sig.
## 1 it02_voiced_gamm_ar_null 2925.68  11
## 2      it02_voiced_gamm_ar 2913.15  13     12.531 2.000 3.614e-06  ***
##
## AIC difference: 20.47, model it02_voiced_gamm_ar has lower AIC.
```

```
## Warning in compareML(it02_voiced_gamm_ar_null, it02_voiced_gamm_ar): AIC
## might not be reliable, as an AR1 model is included (rho1 = 0.741853, rho2 =
## 0.741853).
```

```
plot_gamsd(
    it02_voiced_gamm_ar,
    view = "X",
    comparison = list(position_ord = c("maximum", "closure")),
    conditions = list(c2place.ord = "coronal")
)
```

```
## Summary:
##  * X : numeric predictor; with 100 values ranging from -42.852800 to 64.280100.
##  * position_ord : factor; set to the value(s): closure, maximum.
##  * c2place.ord : factor; set to the value(s): coronal.
##  * vowel.ord : factor; set to the value(s): a.
##  * item_no : factor; set to the value(s): 1.2585_12/12/2016 14:42:43.
```


position_ord
maximum
closure
fit
0
−20
−40
est
3
0
−3
−6
−25
0
25
50
X