Pilot study of Lombard voicing

Stefano Coretta

20/09/2017

1 Extract data

```
lmt("./lombard/code/get-measurements.praat.md")
praat_run("./lombard/code/get-measurements.praat")
lmt("./lombard/code/egg.praat.md")
praat_run("./lombard/code/get-voicing.praat", 40, 10000, 11, 0)
```

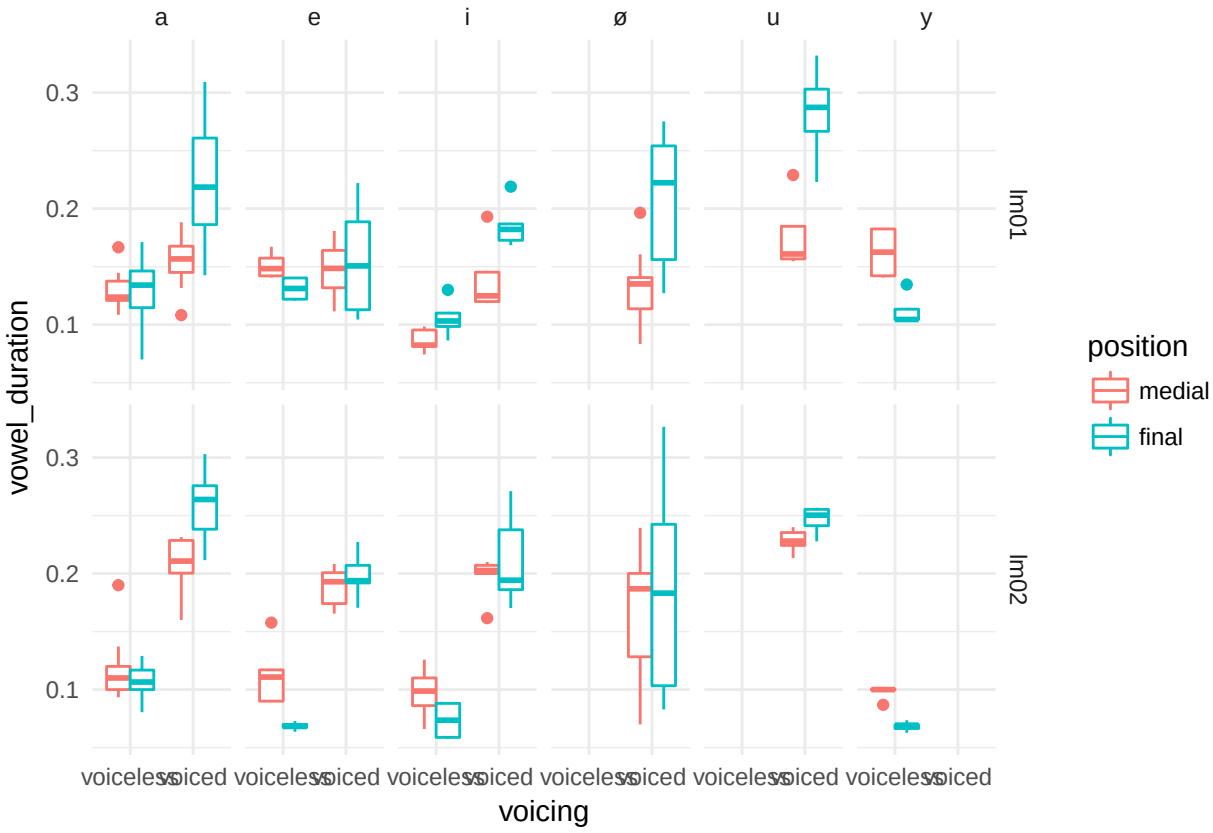
2 Import data

```
stimuli <- read_csv("./lombard/task/prompts.csv")

acoustics <- read_csv("./lombard/results/acoustics.csv", na = "--undefined--") %>%
    left_join(y = stimuli)

durations <- read_csv("./lombard/results/durations.csv", na = "--undefined--") %>%
    left_join(y = stimuli) %>%
    mutate_if(is.character, as.factor) %>%
    mutate(
        voicing = factor(voicing, levels = c("voiceless", "voiced")),
        position = factor(position, levels = c("medial", "final")),
        place = factor(place, levels = c("velar", "coronal", "labial"))
    )

voicing <- read_csv("./lombard/results/voicing.csv", na = "--undefined--") %>%
    left_join(y = stimuli) %>%
    mutate(
        voicing = factor(voicing, levels = c("voiceless", "voiced")),
        position = factor(position, levels = c("medial", "final")),
        place = factor(place, levels = c("velar", "coronal", "labial")),
        devoicing = ifelse(voicing_duration > consonant_duration / 2, "voiced", "devoiced"),
        devoicing_3 = ifelse(voicing_duration > consonant_duration / 3, "voiced", "devoiced"),
        devoiced = ifelse(
            voicing_duration < consonant_duration / 5, "1_5",
            ifelse(
                voicing_duration < consonant_duration / 5 * 2, "2_5",
                ifelse(
                    voicing_duration < consonant_duration / 5 * 3, "3_5",
                    ifelse(
                        voicing_duration < consonant_duration / 5 * 4, "4_5",
                        "voiced"
                    )
                )
            )
```

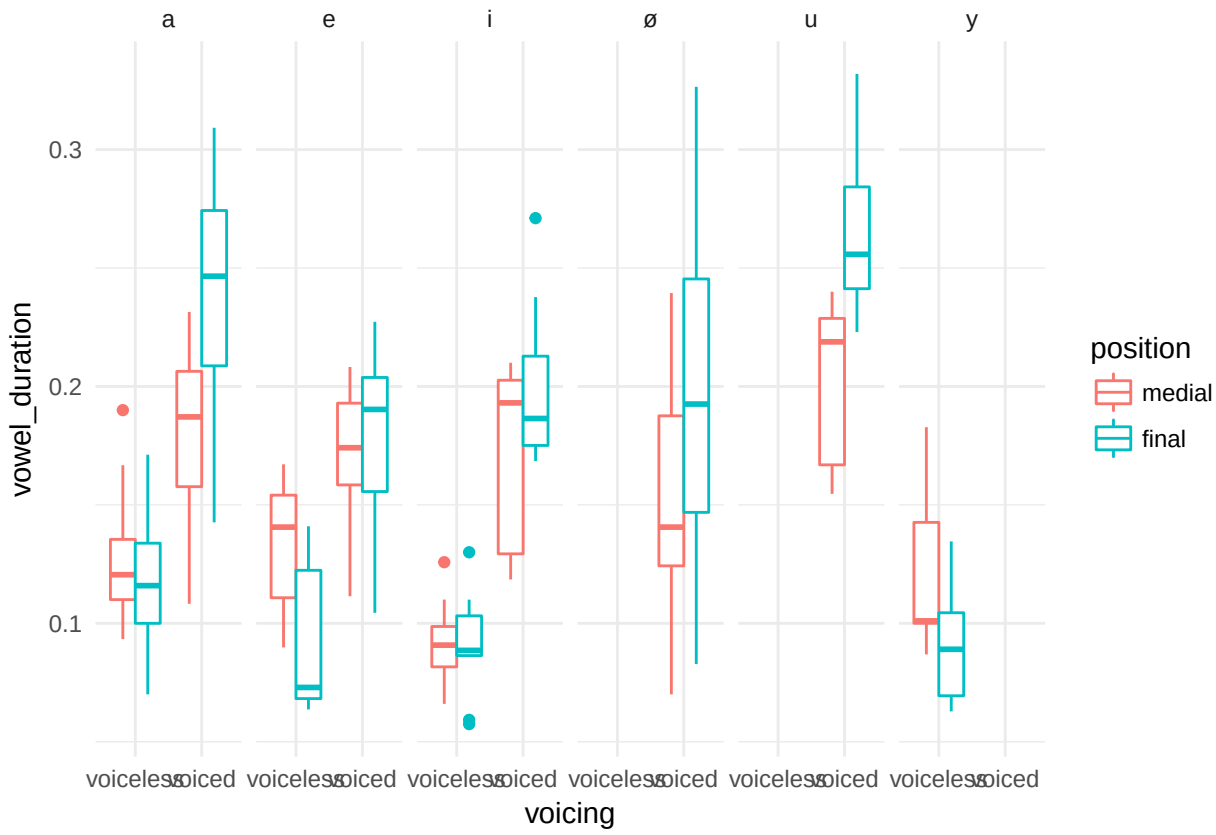
```
        )
    ) %>%
    mutate_if(is.character, as.factor)
```

3 Vowel duration

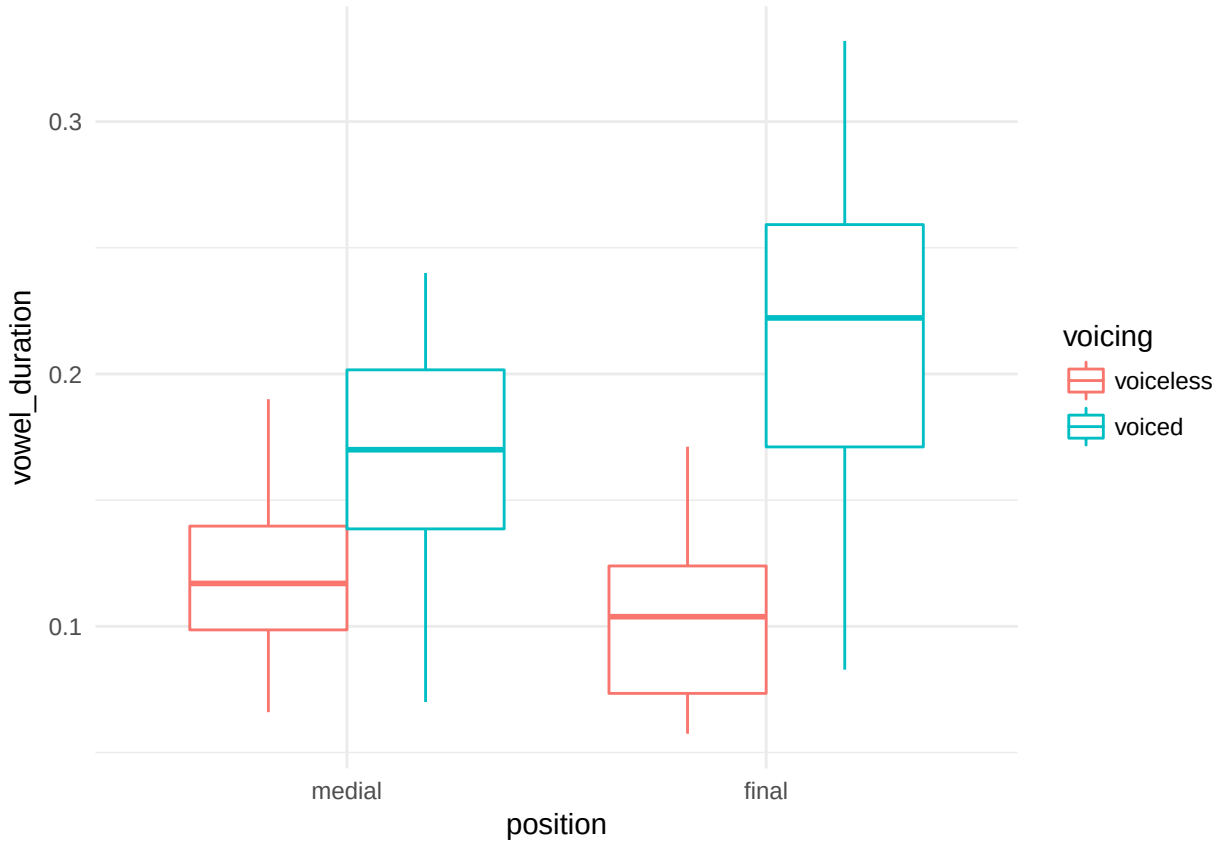
```
durations %>%
    ggplot(aes(voicing, vowel_duration, colour = position)) +
    geom_boxplot() +
    facet_grid(speaker ~ vowel)
```



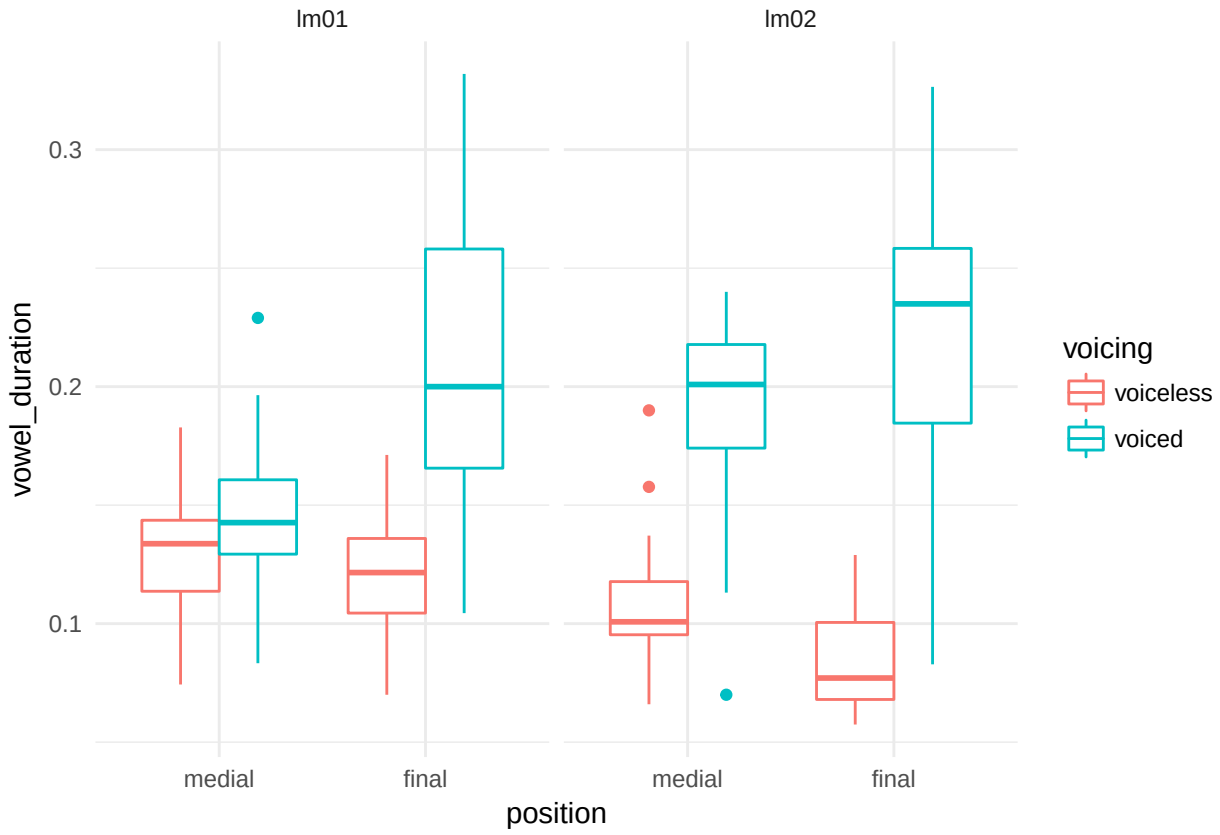

```
durations %>%
    ggplot(aes(voicing, vowel_duration, colour = position)) +
    geom_boxplot() +
    facet_grid( ~ vowel)
```




```
durations %>%
    ggplot(aes(position, vowel_duration, colour = voicing)) +
    geom_boxplot()
```




```
durations %>%
    ggplot(aes(position, vowel_duration, colour = voicing)) +
    geom_boxplot() +
    facet_grid(. ~ speaker)
```




```
vowel_dur_lm <- lmer(
    vowel_duration ~
        voicing *
        position +
        manner +
        place +
        (1|speaker) +
        (1|word),
    data = durations
)

summary(vowel_dur_lm)
```

```
## Linear mixed model fit by REML t-tests use Satterthwaite approximations
##   to degrees of freedom [lmerMod]
## Formula:
## vowel_duration ~ voicing * position + manner + place + (1 | speaker) +
##     (1 | word)
##    Data: durations
##
## REML criterion at convergence: -880.6
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.70947 -0.75344  0.03827  0.70662  2.59817
##
## Random effects:
##  Groups   Name        Variance  Std.Dev.
```

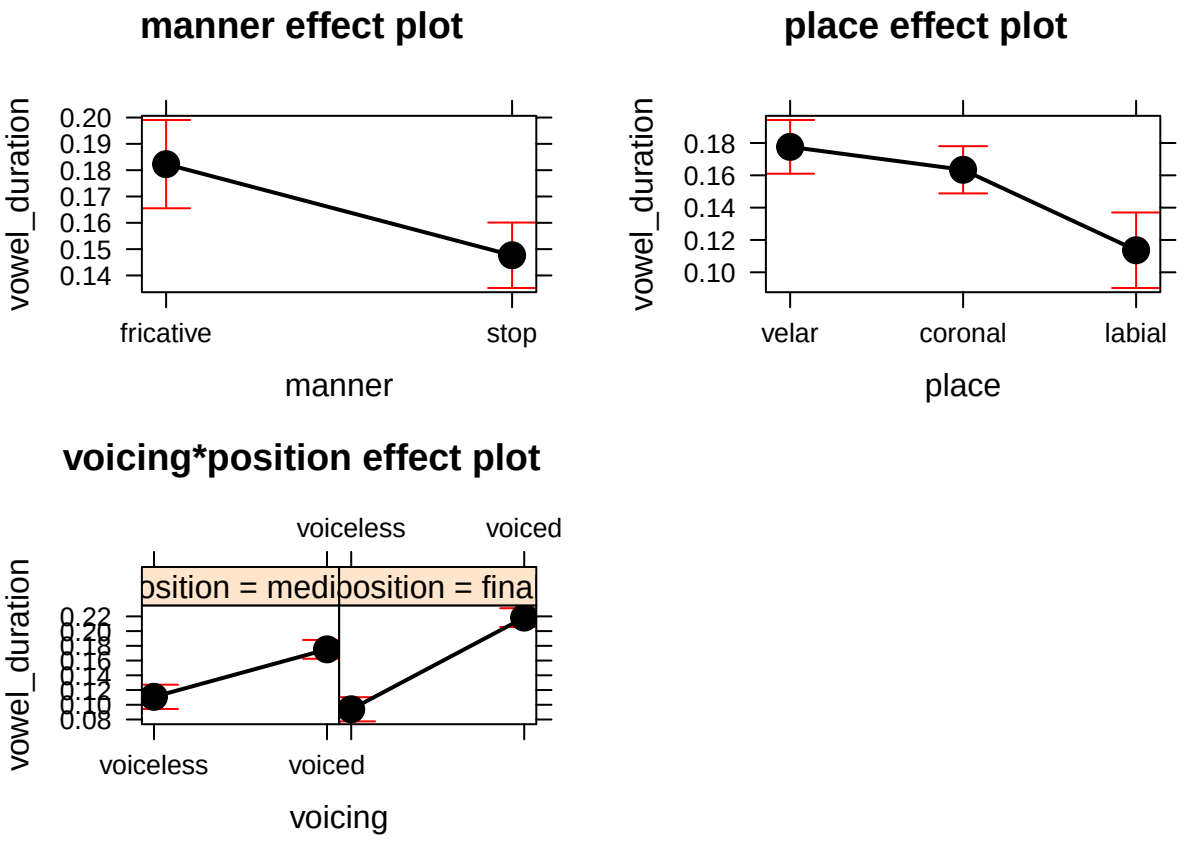
```
##  word     (Intercept) 1.831e-04 0.013530
##  speaker  (Intercept) 4.122e-07 0.000642
##  Residual             1.338e-03 0.036576
## Number of obs: 250, groups:  word, 13; speaker, 2
##
## Fixed effects:
##                              Estimate Std. Error        df t value
## (Intercept)                   0.14874    0.01539   8.95000   9.668
## voicingvoiced                 0.06446    0.01081  12.30000   5.963
## positionfinal                -0.01691    0.00751 234.12000  -2.251
## mannerstop                   -0.03462    0.01176   7.96000  -2.944
## placecoronal                 -0.01419    0.01247   7.96000  -1.138
## placelabial                  -0.06398    0.01500   7.82000  -4.264
## voicingvoiced:positionfinal   0.06001    0.00954 234.08000   6.290
##                             Pr(>|t|)
## (Intercept)                 4.92e-06 ***
## voicingvoiced               5.94e-05 ***
## positionfinal                0.02532 *
## mannerstop                   0.01870 *
## placecoronal                 0.28803
## placelabial                  0.00289 **
## voicingvoiced:positionfinal 1.54e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) vcngvc pstnfn mnnrst plccrn plclbl
## voicingvocd -0.424
## positionfnl -0.247  0.350
## mannerstop  -0.765  0.003  0.006
## placecoronl -0.772  0.065 -0.004  0.630
## placelabial -0.420 -0.233 -0.001  0.392  0.497
## vcngvcd:pst  0.197 -0.444 -0.787 -0.011  0.004  0.001
```

```
vowel_dur_lm_null <- lmer(
    vowel_duration ~
        voicing *
#        position +
        manner +
        place +
        (1|speaker) +
        (1|word),
    data = durations
)

anova(vowel_dur_lm_null, vowel_dur_lm)
```

```
## Data: durations
## Models:
## object: vowel_duration ~ voicing * manner + place + (1 | speaker) + (1 |
## object:     word)
## ..1: vowel_duration ~ voicing * position + manner + place + (1 | speaker) +
## ..1:     (1 | word)
##        Df     AIC     BIC logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  9 -869.30 -837.61 443.65  -887.30
```

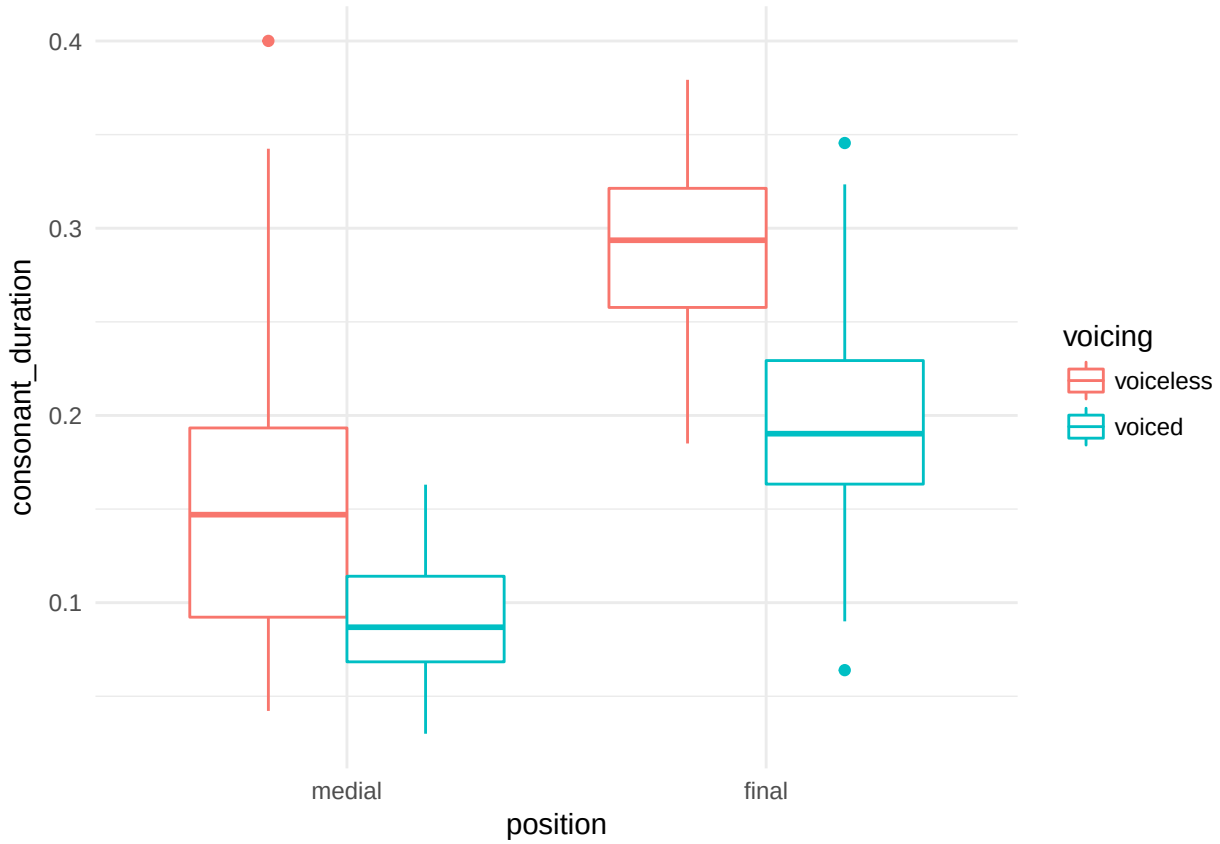
```
## ..1    10 -916.25 -881.03 468.12  -936.25 48.946      1  2.631e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(allEffects(vowel_dur_lm))
```



4 Consonant duration

```
durations %>%
    ggplot(aes(position, consonant_duration, colour = voicing)) +
    geom_boxplot()
```




```
c_dur_lm <- lmer(
    consonant_duration ~
        voicing *
        position +
        manner +
        place +
        (1|speaker) +
        (1|word),
    data = durations
)

summary(c_dur_lm)
```

```
## Linear mixed model fit by REML t-tests use Satterthwaite approximations
##   to degrees of freedom [lmerMod]
## Formula: consonant_duration ~ voicing * position + manner + place + (1 |
##     speaker) + (1 | word)
##    Data: durations
##
## REML criterion at convergence: -756.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6746 -0.5739 -0.0443  0.4711  5.0673
##
## Random effects:
##  Groups   Name        Variance  Std.Dev.
```

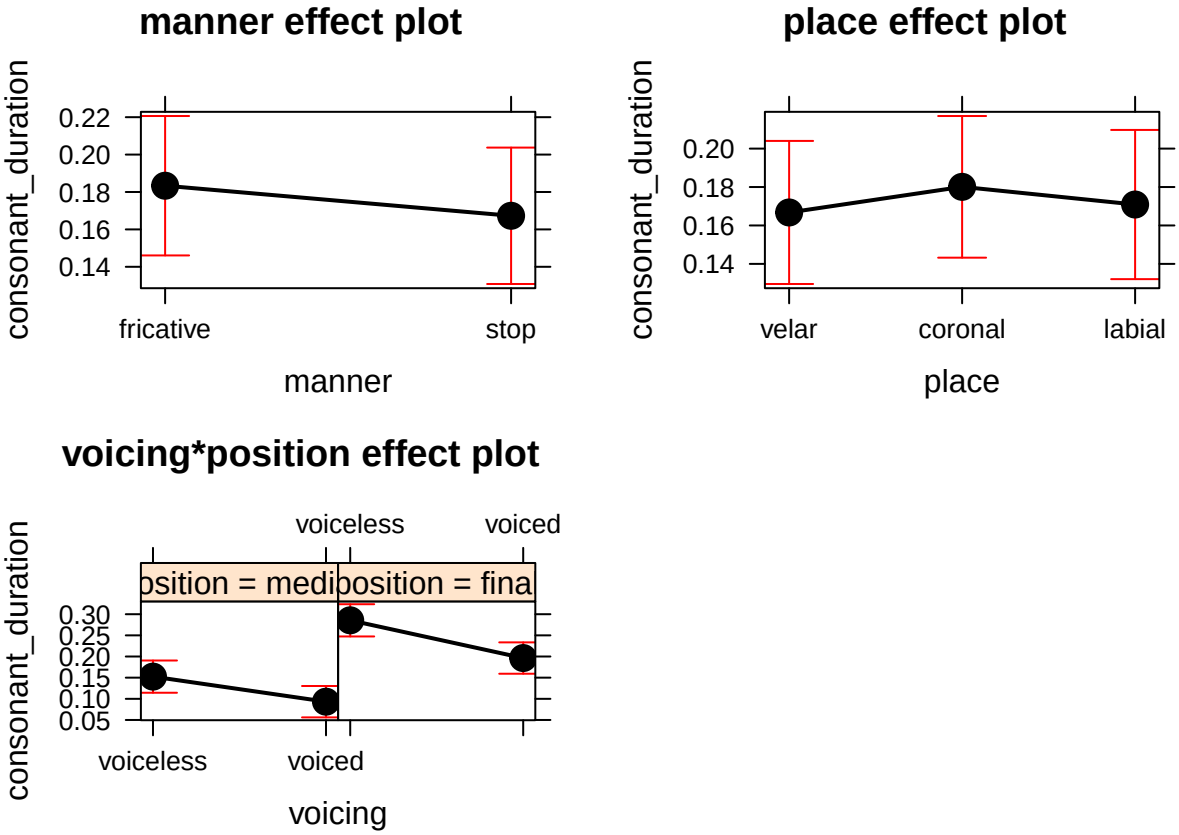
```
##  word     (Intercept) 0.0000000 0.00000
##  speaker  (Intercept) 0.0006479 0.02545
##  Residual             0.0022940 0.04790
## Number of obs: 250, groups:  word, 13; speaker, 2
##
## Fixed effects:
##                              Estimate Std. Error         df t value
## (Intercept)                  0.155670   0.021297   1.850000   7.309
## voicingvoiced               -0.059241   0.009167 242.000000  -6.462
## positionfinal                0.133090   0.009831 242.000000  13.537
## mannerstop                  -0.016071   0.008113 242.010000  -1.981
## placecoronal                 0.013331   0.008606 242.010000   1.549
## placelabial                  0.004132   0.010238 242.000000   0.404
## voicingvoiced:positionfinal -0.029952   0.012489 242.000000  -2.398
##                             Pr(>|t|)
## (Intercept)                   0.0223 *
## voicingvoiced               5.63e-10 ***
## positionfinal                < 2e-16 ***
## mannerstop                    0.0487 *
## placecoronal                  0.1227
## placelabial                   0.6868
## voicingvoiced:positionfinal   0.0172 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) vcngvc pstnfn mnnrst plccrn plclbl
## voicingvocd -0.262
## positionfnl -0.234  0.541
## mannerstop  -0.382  0.008  0.011
## placecoronl -0.381  0.044 -0.009  0.633
## placelabial -0.210 -0.195 -0.002  0.398  0.502
## vcngvcd:pst  0.186 -0.686 -0.787 -0.020  0.007  0.002
```

```
c_dur_lm_null <- lmer(
    consonant_duration ~
        voicing *
#        position +
        manner +
        place +
        (1|speaker) +
        (1|word),
    data = durations
)

anova(c_dur_lm_null, c_dur_lm)
```

```
## Data: durations
## Models:
## object: consonant_duration ~ voicing * manner + place + (1 | speaker) +
## object:     (1 | word)
## ..1: consonant_duration ~ voicing * position + manner + place + (1 |
## ..1:     speaker) + (1 | word)
##        Df     AIC     BIC logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  9 -565.85 -534.16 291.92  -583.85
```

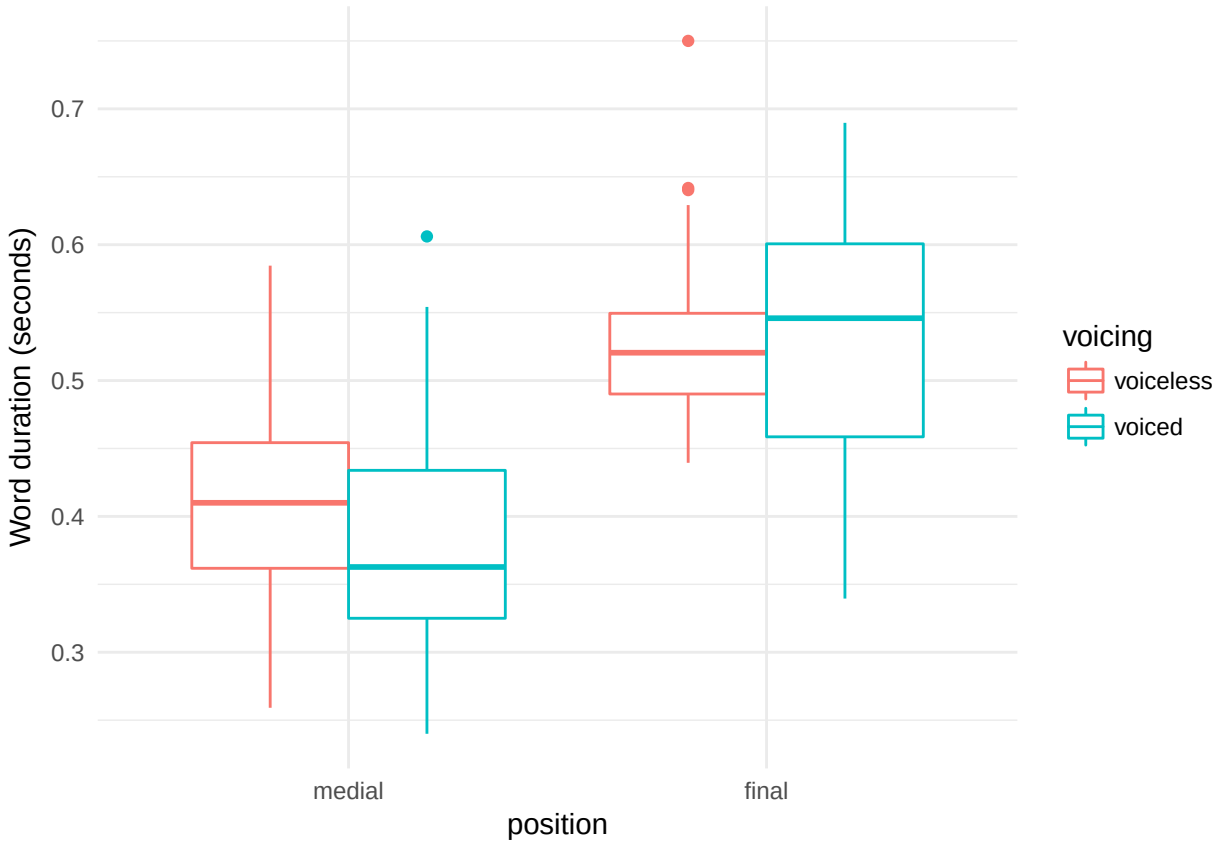
```
## ..1    10 -790.17 -754.95 405.08  -810.17 226.32      1  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(allEffects(c_dur_lm))
```



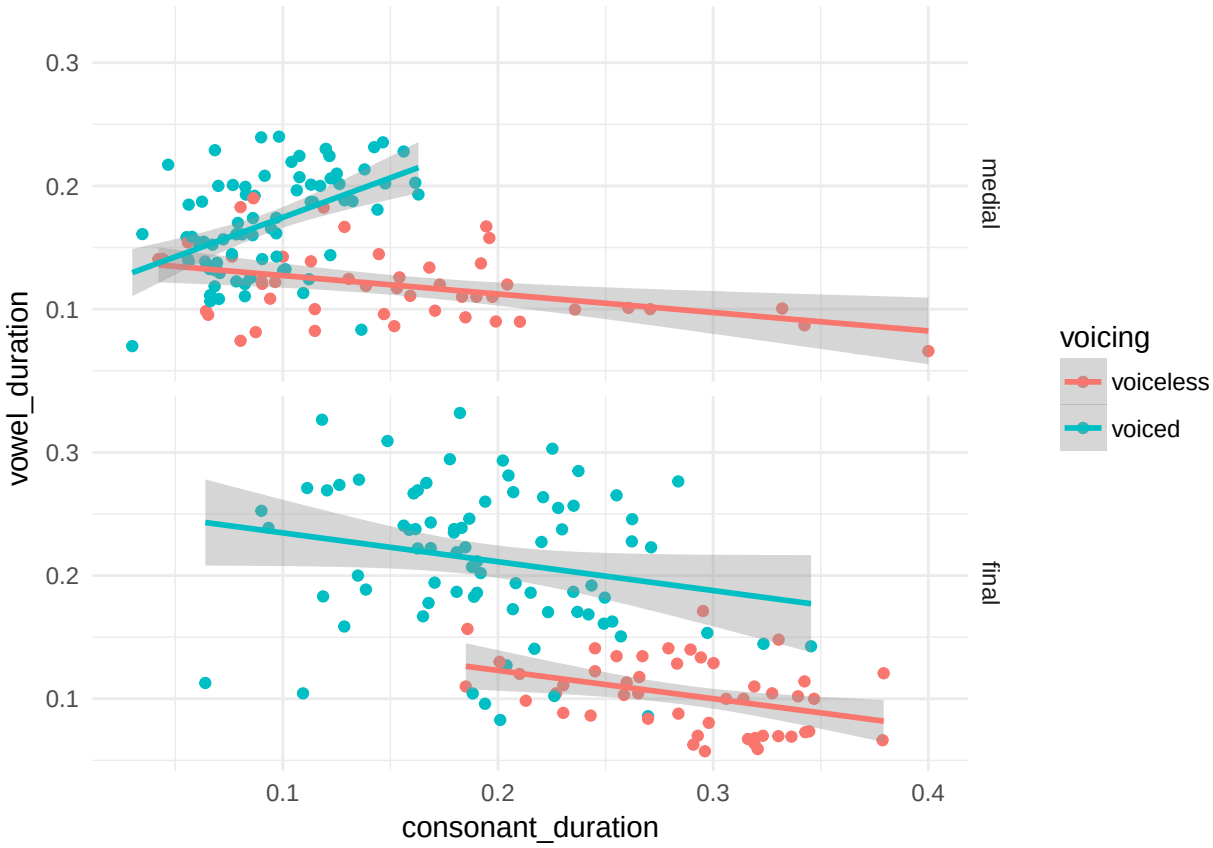
5 Word duration

```
durations %>%
    ggplot(aes(position, word_duration, colour = voicing)) +
    geom_boxplot() +
    ylab("Word duration (seconds)")
```



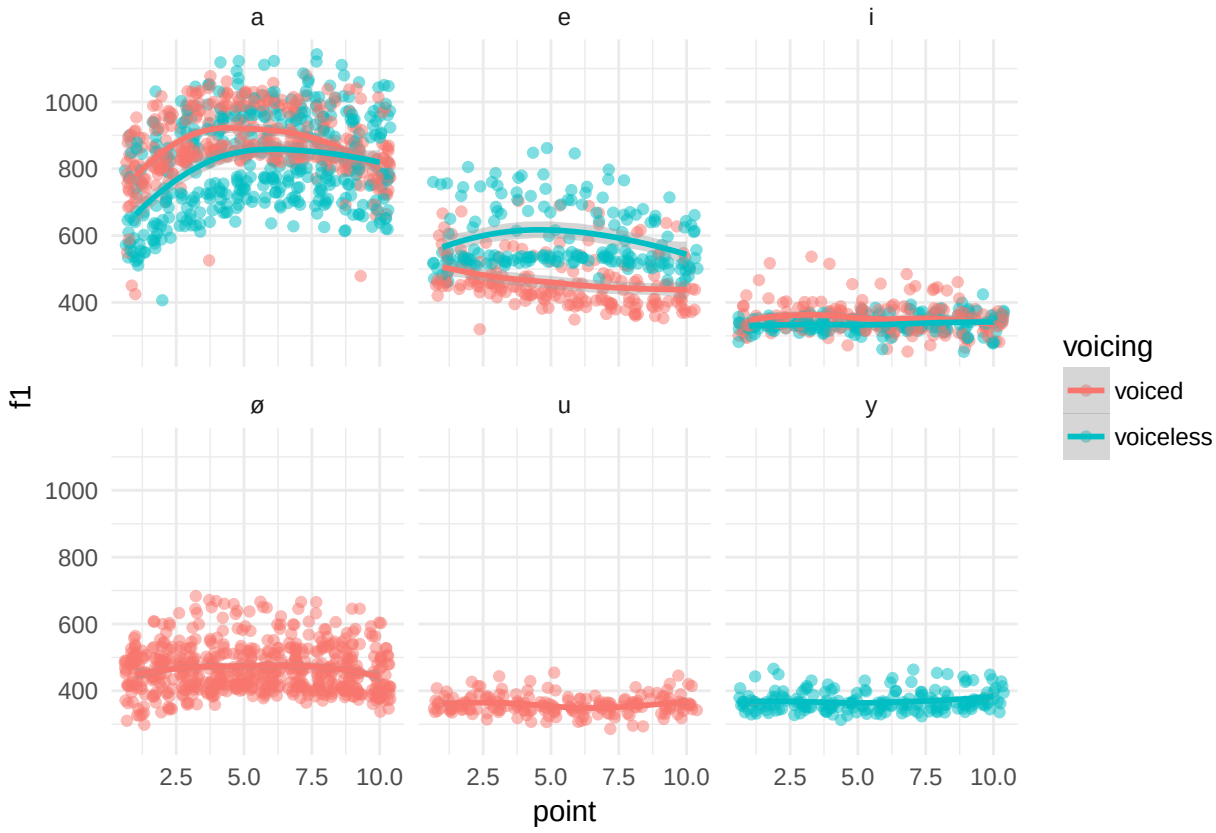
6 Vowel and consonant duration

```
durations %>%
    ggplot(aes(consonant_duration, vowel_duration, colour = voicing)) +
    geom_point() +
    geom_smooth(method = "lm") +
    facet_grid(position ~ .)
```

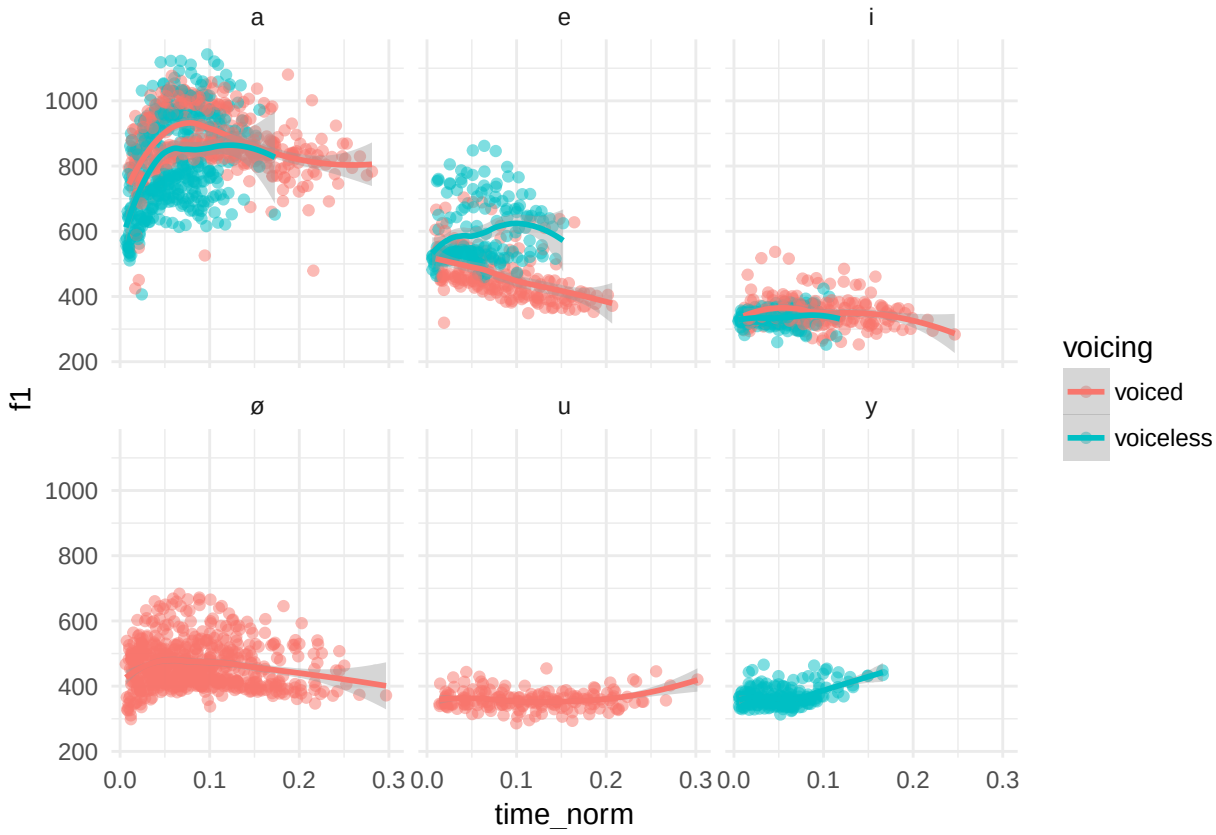


7 Acoustics

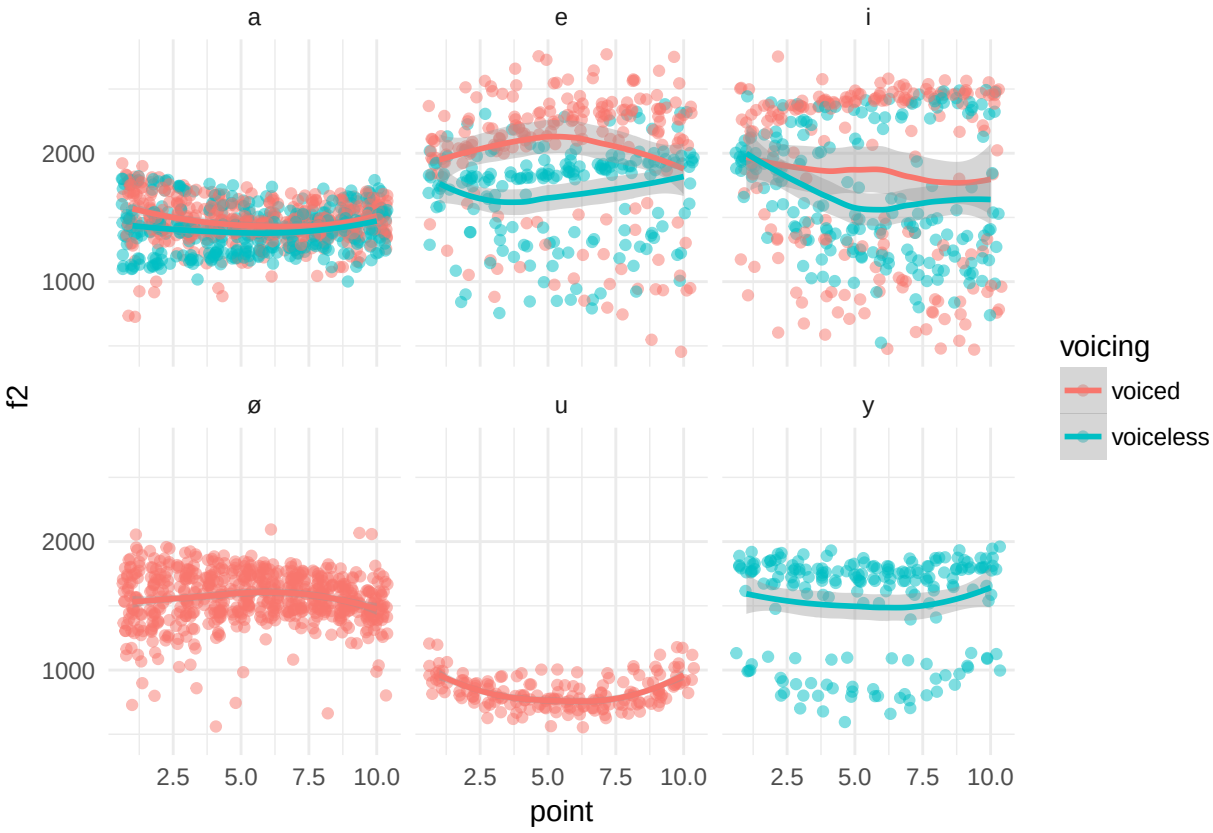
```
acoustics %>%
    ggplot(aes(point, f1, colour = voicing)) +
    geom_jitter(alpha = 0.5) +
    geom_smooth() +
    facet_wrap(~ vowel)
```




```
acoustics %>%
    ggplot(aes(time_norm, f1, colour = voicing)) +
    geom_point(alpha = 0.5) +
    geom_smooth() +
    facet_wrap(~ vowel)
```




```
acoustics %>%
    ggplot(aes(point, f2, colour = voicing)) +
    geom_jitter(alpha = 0.5) +
    geom_smooth() +
    facet_wrap(~ vowel)
```




```
acoustics %>%
    filter(pitch < 400) %>%
    ggplot(aes(point, pitch, colour = voicing)) +
    geom_jitter(alpha = 0.5) +
    geom_smooth() +
    facet_wrap(~ speaker)
```

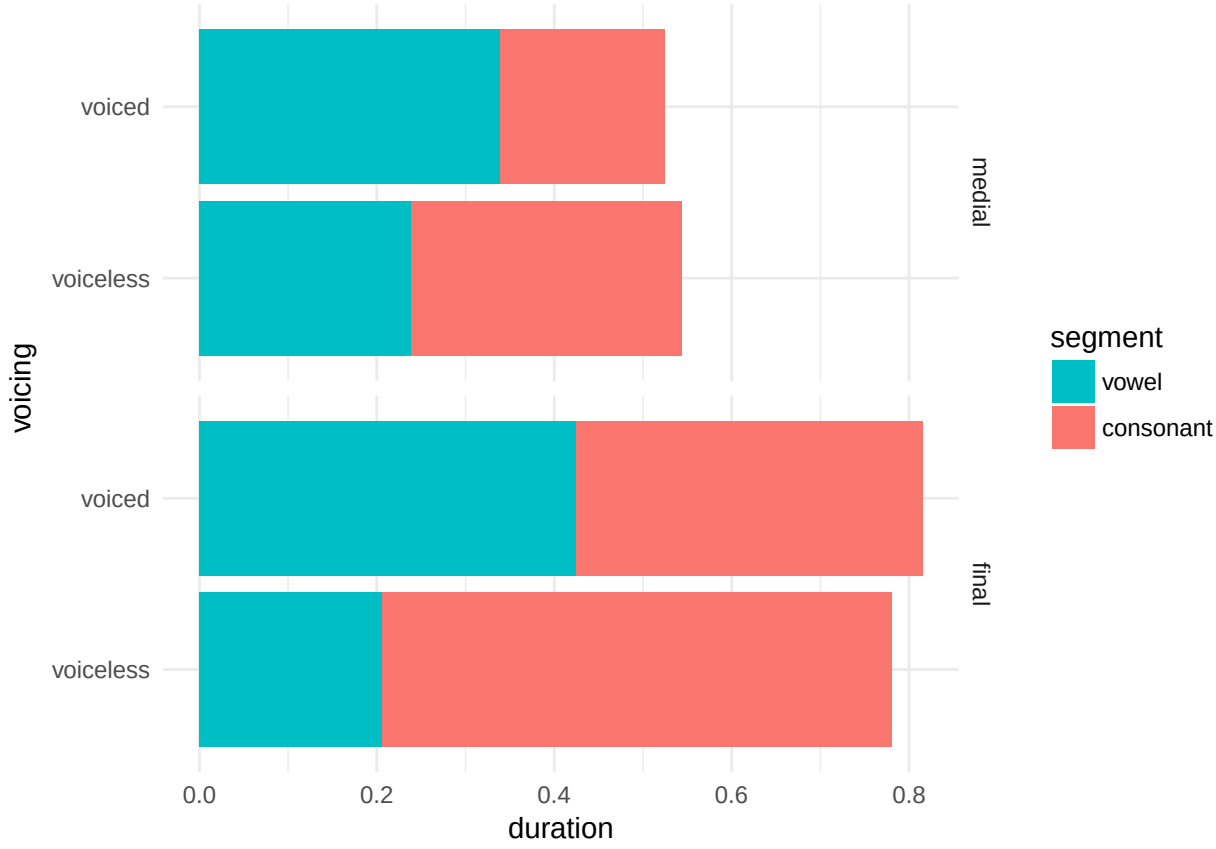
8 Proportions

```
proportions <- durations %>%
    group_by(speaker,voicing, position) %>%
    summarise(
        total =
            mean(vowel_duration, na.rm = TRUE) +
            mean(consonant_duration, na.rm = TRUE),
        vowel = mean(vowel_duration, na.rm = TRUE),
        consonant = mean(consonant_duration, na.rm = TRUE)
    ) %>%
    gather(segment, duration, vowel:consonant) %>%
    mutate(segment = factor(segment,
                            levels = c("consonant", "vowel")
                            )
           )
```

There is sentence final lengthening. The medial V-to-C ratio is not maintained in final position. The

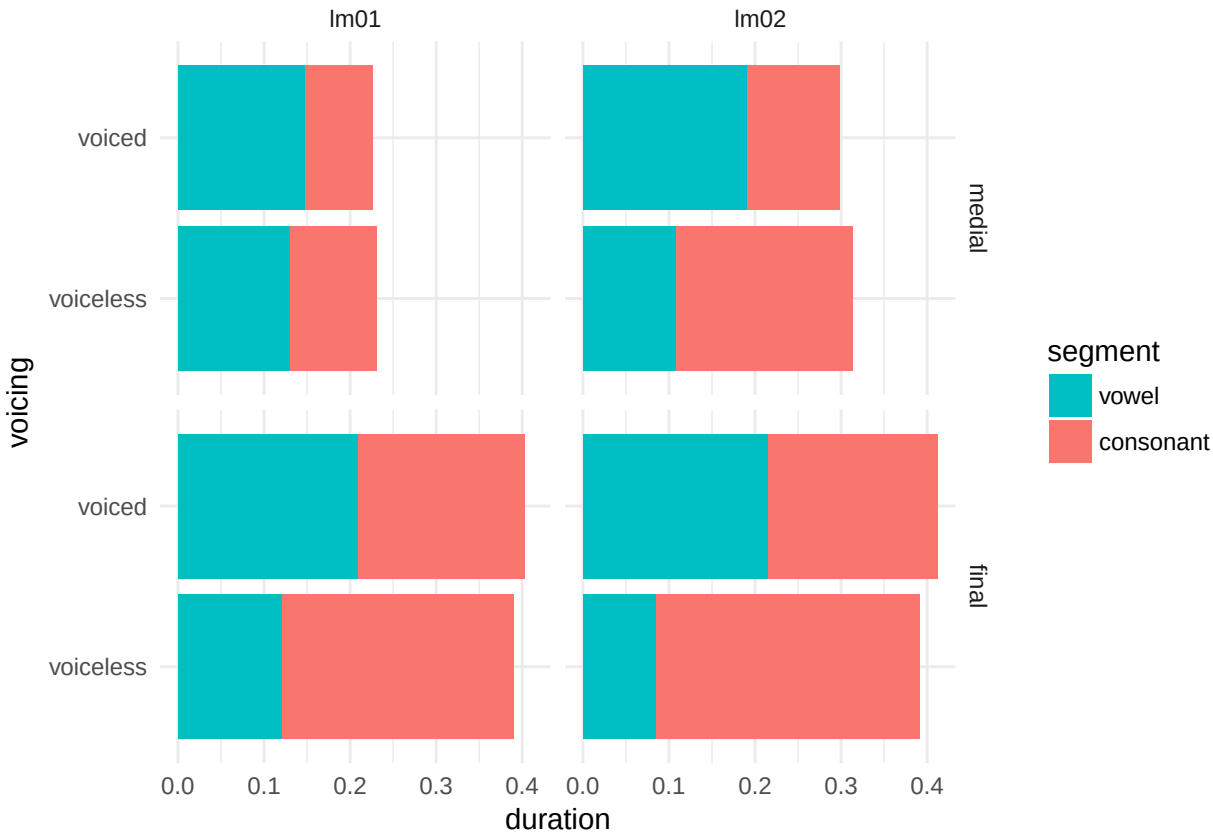
magnitude of the voicing effect increases in sentence final position. The plot averages accross speakers.

```
ggplot(proportions, aes(voicing, duration, fill = segment)) +
    geom_bar(stat = "identity") +
    coord_flip() +
    facet_grid(position ~ .) +
    guides(fill = guide_legend(reverse = TRUE))
```



Plot with individual speakers.

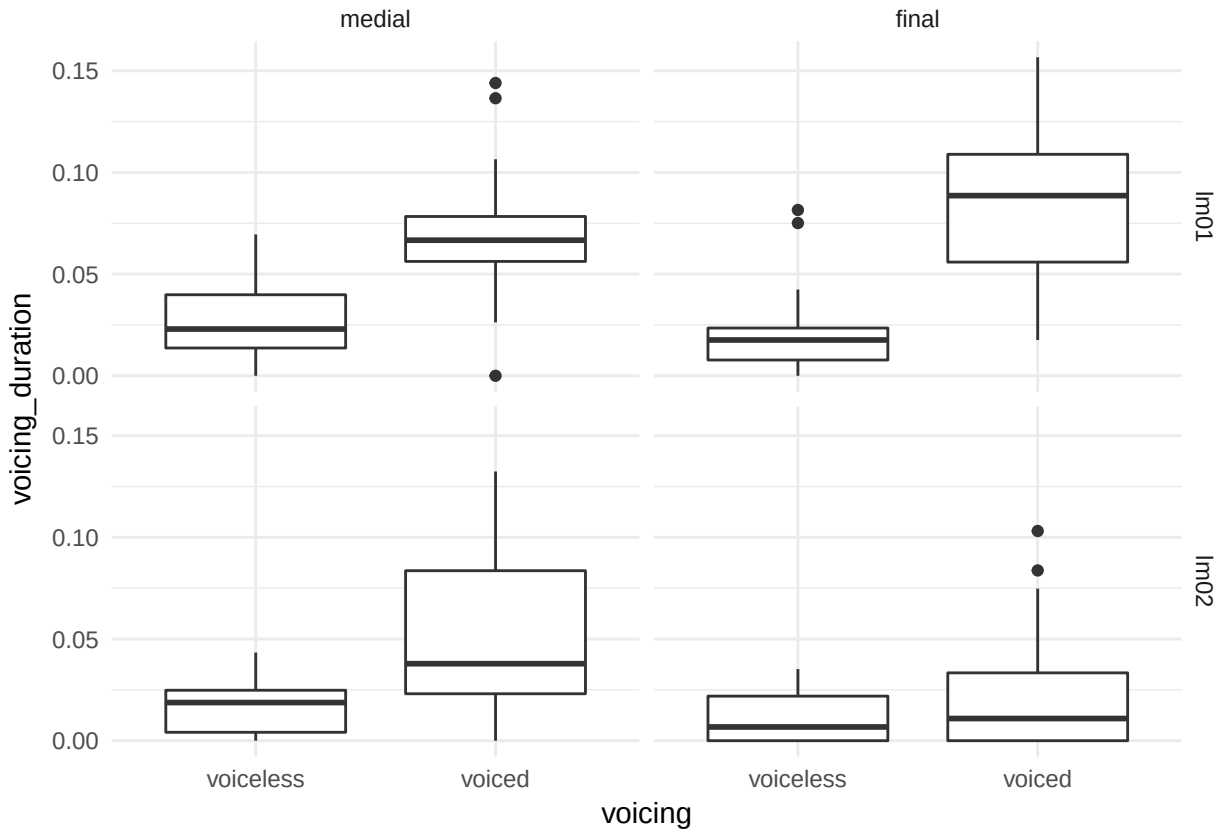
```
ggplot(proportions, aes(voicing, duration, fill = segment)) +
    geom_bar(stat = "identity") +
    coord_flip() +
    facet_grid(position ~ speaker) +
    guides(fill = guide_legend(reverse = TRUE))
```



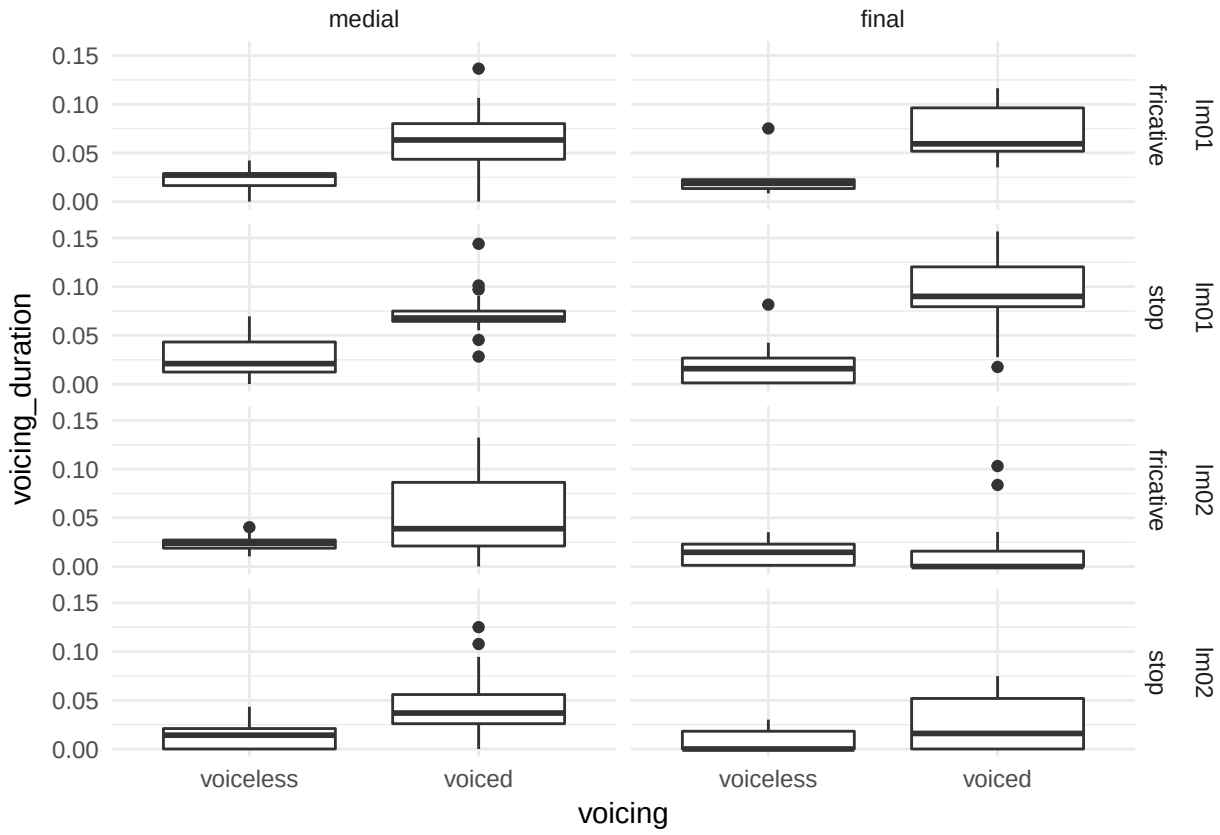
9 Voicing durations

There is some voicing bleed in voiceless consonants in both speakers. Voicing duration in voiced consonants varies. (Boxplots used here, although voicing can be 0). LM02 devoiced more in sentence-final position compared to LM01.

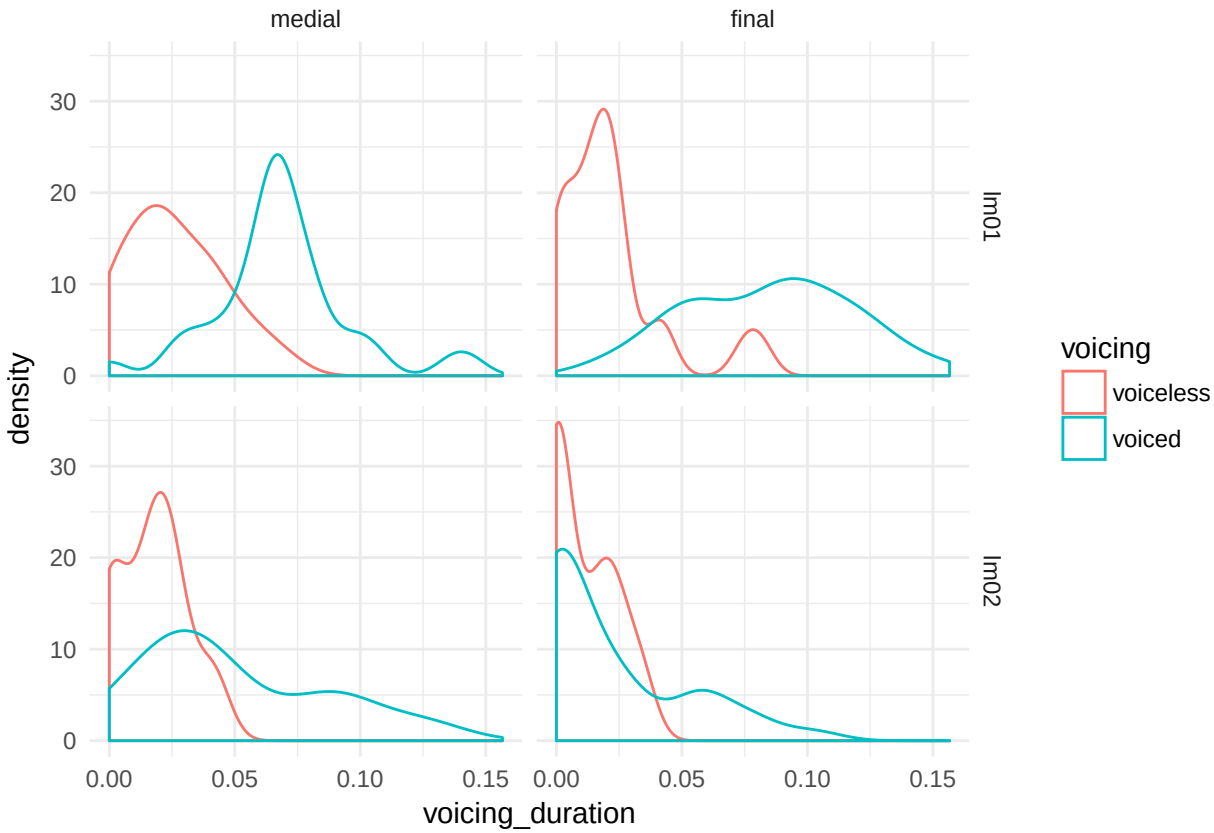
```
voicing %>%
    ggplot(aes(voicing, voicing_duration)) +
    geom_boxplot() +
    facet_grid(speaker ~ position)
```




```
voicing %>%
    ggplot(aes(voicing, voicing_duration)) +
    geom_boxplot() +
    facet_grid(speaker + manner ~ position)
```

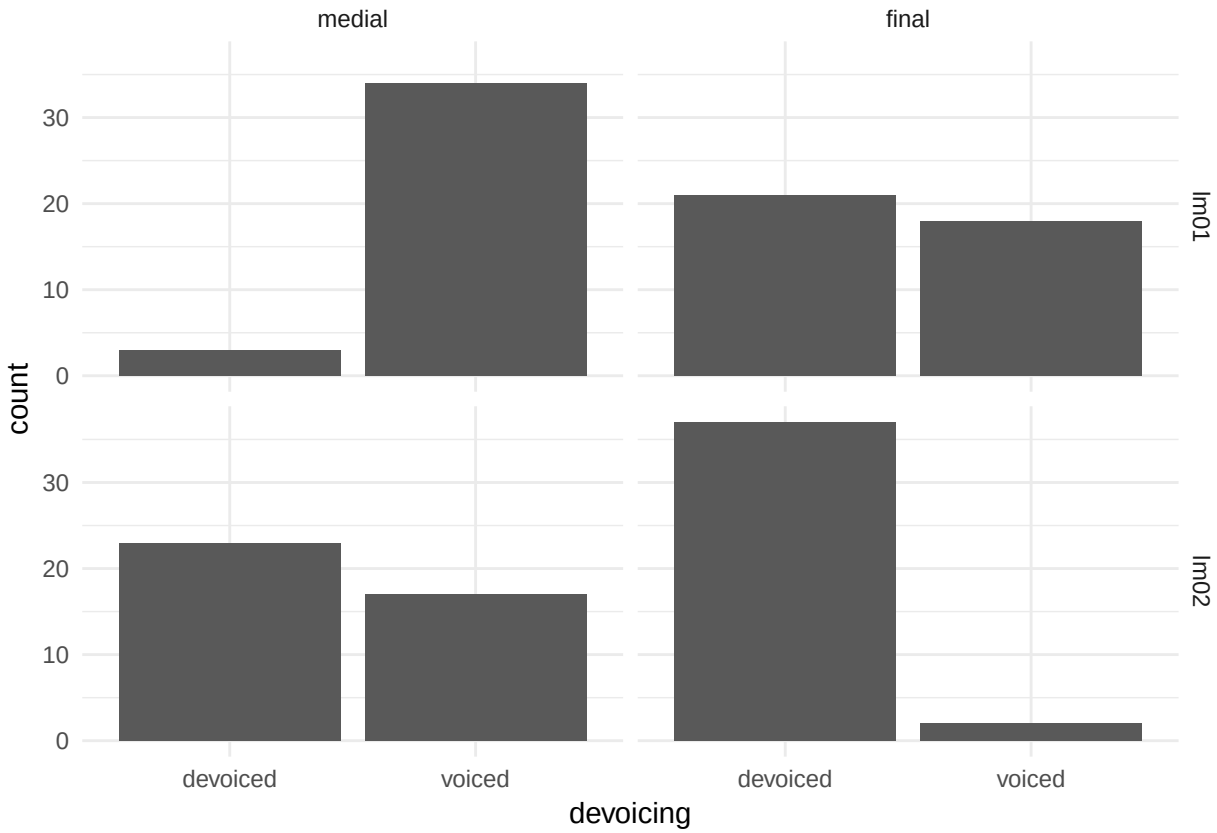



```
voicing %>%
    ggplot(aes(voicing_duration,colour = voicing)) +
    geom_density() +
    facet_grid(speaker ~ position)
```



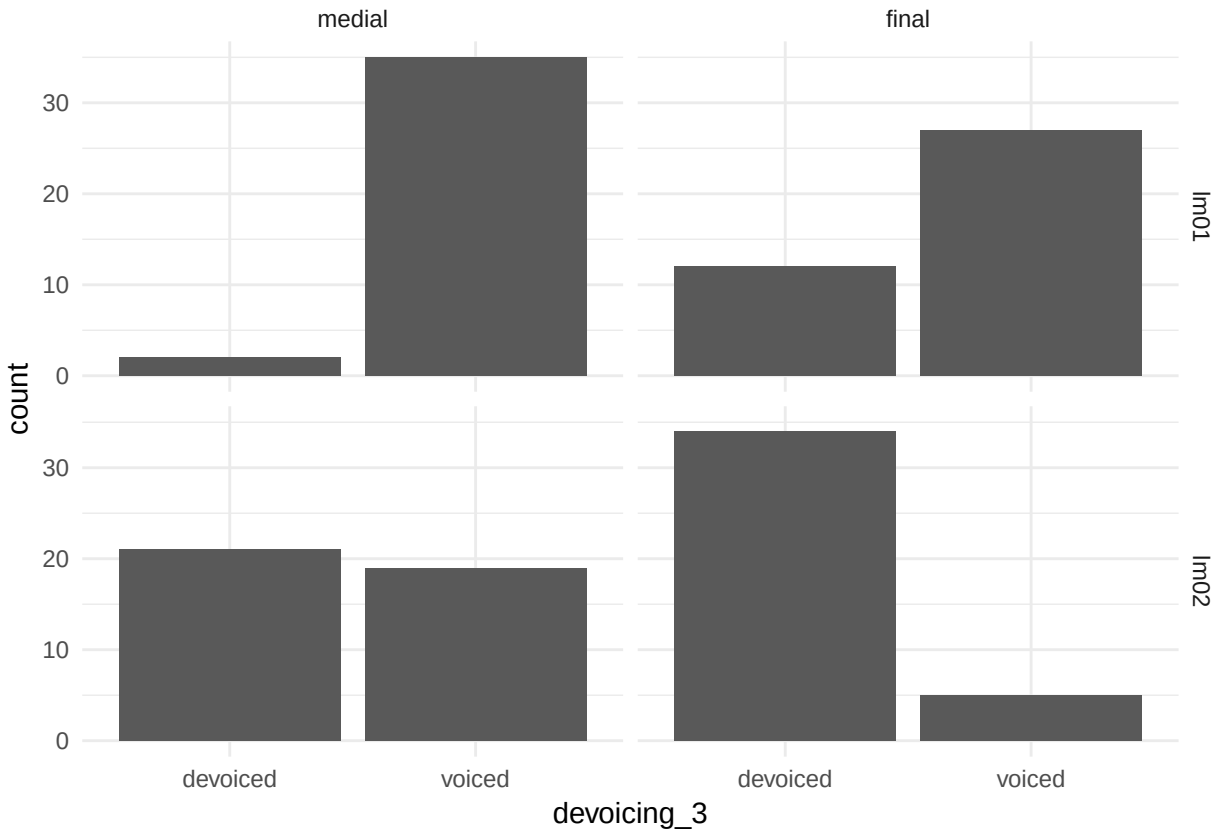
The following plots the number of voiced consonants that are `voiced` (voicing duration $> 50\%$ of consonant duration), or `devoiced` (voicing duration $< 50\%$ of consonant duration). Two patterns emerge. (1) LM01 more consistently employes voiced consonants in sentence-medial position, while there's a 50:50 chance at sentence-final position (but cf. below). (2) LM02 has about 50:50 chance at sentence medial position, but strongly favours devoicing at sentence=final position.

```
voicing %>%
    filter(voicing == "voiced") %>%
    ggplot(aes(devoicing)) +
    geom_bar() +
    facet_grid(speaker ~ position)
```



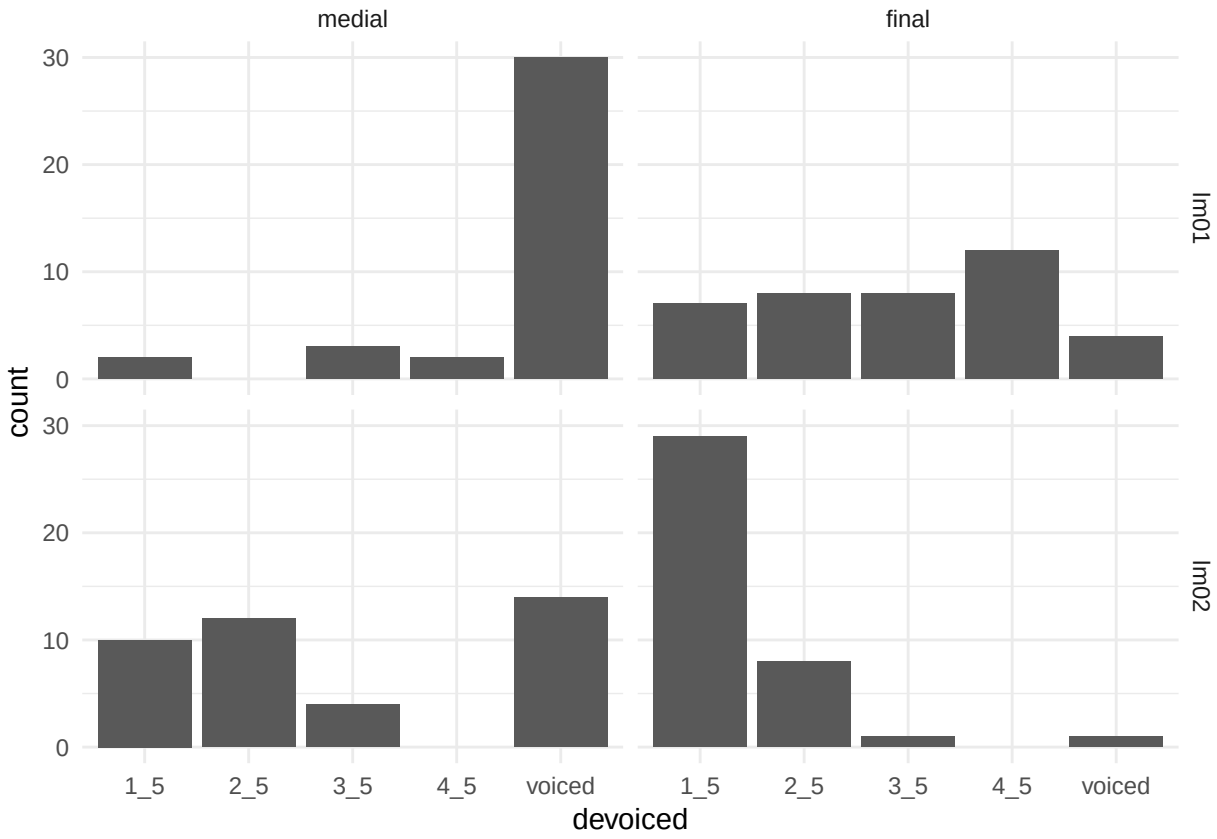
Things change a bit if the cut off is at 1/3 of the consonant duration (rather than 1/2). LM01 favours consonants with more than 1/3 voicing in sentence-final.

```
voicing %>%
    filter(voicing == "voiced") %>%
    ggplot(aes(devoicing_3)) +
    geom_bar() +
    facet_grid(speaker ~ position)
```

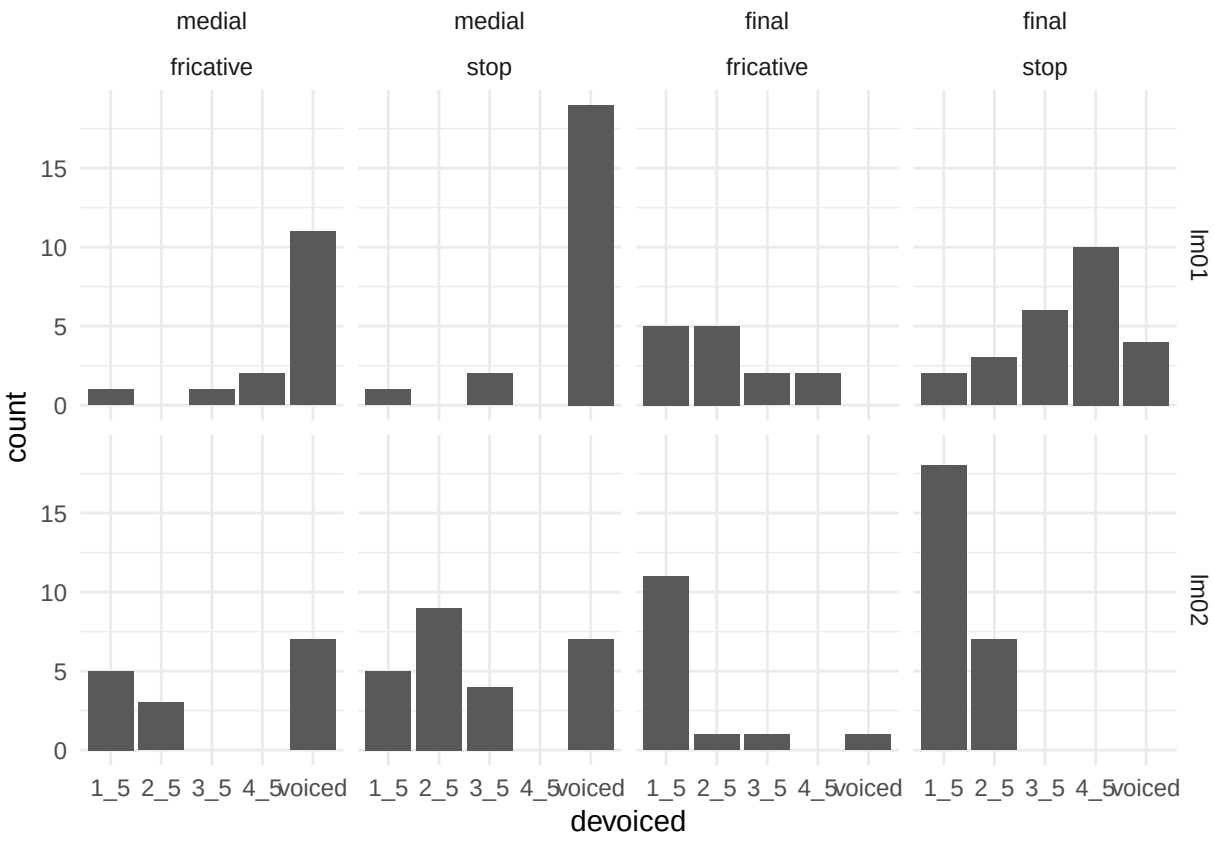


This is probably the most insightful plot. Here bins are created at each fifth of consonant duration: 1_5 means voicing duration $< 1/5$ of consonant duration, 2_5 means $< 2/5$ and $> 1/5$, and so on... The idiosyncratic patterns of sentence final devoicing is interesting: LM01 doesn't really favours one bin over the other. For LM02 devoiced consonants in sentence final position with less than 1/5 of consonant duration seem to be favoured.

```
voicing %>%
    filter(voicing == "voiced") %>%
    ggplot(aes(devoiced)) +
    geom_bar() +
    facet_grid(speaker ~ position)
```




```
voicing %>%
    filter(voicing == "voiced") %>%
    ggplot(aes(devoiced)) +
    geom_bar() +
    facet_grid(speaker ~ position + manner)
```


medial
fricative
medial
stop
final
fricative
final
stop
lm01
lm02
count
devoiced
0
5
10
15
1_5
2_5
3_5
4_5
voiced