

# Acoustics of fricatives in Nendaz Francoprovençal - LDA

Chong & Kasstan

Apr 2022

## Linear Discriminant Analysis

Model is run with 60-40 partition, with all parameters examined (Spectral Peak, all four spectral moments, duration and relative intensity, formants).

```
## [1] 150
# set seed for reproducible results
set.seed(123)

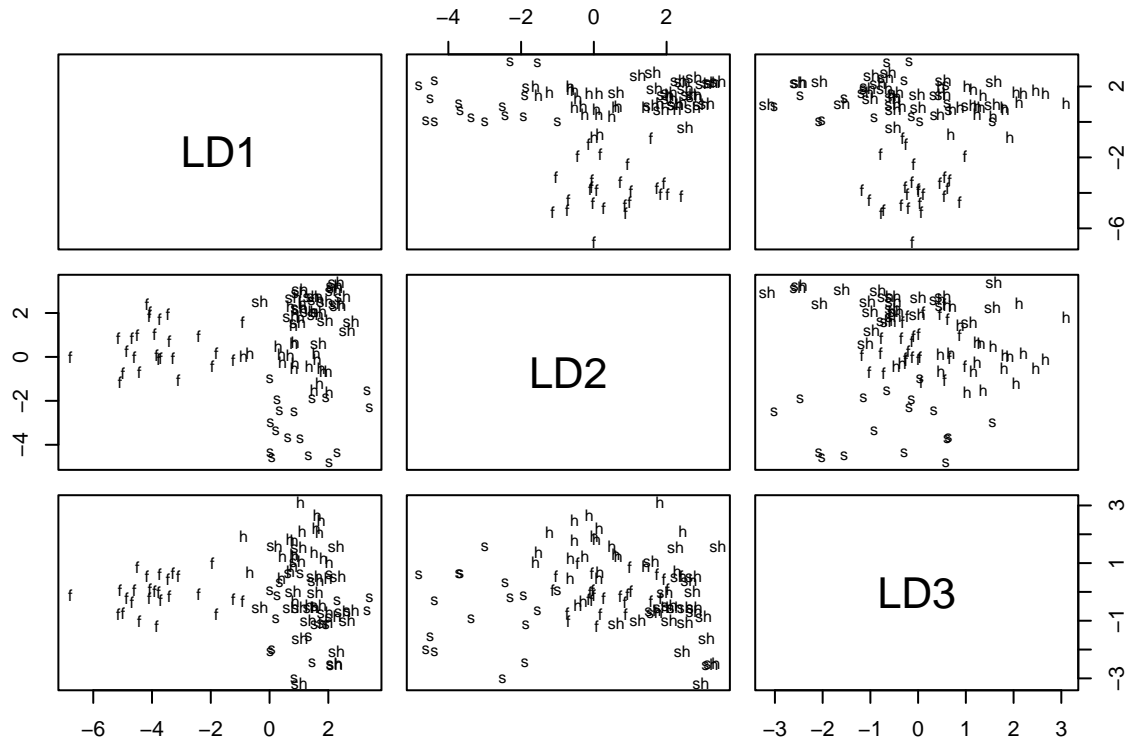
##creates a partition - split data 60% training, 40% test
training <- dat$seg %>%
  createDataPartition(p = 0.6, list = FALSE)

train <- dat[training, ]
test <- dat[-training, ]

preproc.param <- train %>%
  preProcess(method = c("center", "scale"))

train.trans <- preproc.param %>% predict(train)
test.trans <- preproc.param %>% predict(test)

## model includes spectral peak location, 4 spectral moments, AmpDiff and totalDur
model <- lda(seg ~ peak + cog + specSD + skew + kurt + totalDur + DiffAmp + f1 + f2 + f3, train.trans)
plot(model)
```



model

```
## Call:
## lda(seg ~ peak + cog + specSD + skew + kurt + totalDur + DiffAmp +
##     f1 + f2 + f3, data = train.trans)
##
## Prior probabilities of groups:
##     f     h     s     sh
## 0.2795699 0.2580645 0.1827957 0.2795699
##
## Group means:
##      peak      cog    specSD      skew      kurt    totalDur
## f -0.11579265  0.32846847  1.2987300 -0.4933478 -0.69274600 -0.75910909
## h  0.04271517 -0.01256057 -0.1566675 -0.3181782 -0.35326123  0.58841722
## s  1.18888957  1.22380830 -0.4001138 -0.4775167  0.05240099 -0.05249021
## sh -0.70098761 -1.11705645 -0.8925009  1.0992732  0.98457111  0.25027525
##      DiffAmp      f1      f2      f3
## f -0.7311706  0.05032195 -0.8804192  0.1691945
## h -0.1737333  0.26378474  0.5714594  0.2882486
## s  0.5491667 -0.34230724  0.8047708  1.0235450
## sh 0.5324693 -0.06999929 -0.1732781 -1.1045111
##
## Coefficients of linear discriminants:
##      LD1      LD2      LD3
## peak -0.3327839 -0.05446638  0.23463443
## cog  0.5703216 -2.05495428 -1.56752820
## specSD -2.2445795  0.60247595 -0.03686148
## skew -0.1374280 -1.04655570 -1.10995205
## kurt -0.3737224  0.67971741 -0.37810702
## totalDur 0.7223574  0.26106502  0.48911694
## DiffAmp 0.2814586 -0.19056346 -0.29190023
```

```
## f1      0.3107870  0.12285354  0.33456680
## f2      0.5268944 -0.24923239  0.28972443
## f3     -0.4107780 -0.68551914  0.50536739
##
## Proportion of trace:
##   LD1   LD2   LD3
## 0.5725 0.3605 0.0670
```

```
predictions <- model %>% predict(test.trans)
mean(predictions$class == test.trans$seg)
```

```
## [1] 0.8421053
```

```
## contingency table
```

```
preds <- table(test.trans$seg, predictions$class)
preds
```

```
##
##      f  h  s sh
## f  13  2  1  0
## h   1 12  0  2
## s   0  2  8  0
## sh  0  1  0 15
```

```
diag(prop.table(preds, 1))
```

```
##      f      h      s      sh
## 0.8125 0.8000 0.8000 0.9375
```

```
## total percent correct
```

```
sum(diag(prop.table(preds)))
```

```
## [1] 0.8421053
```

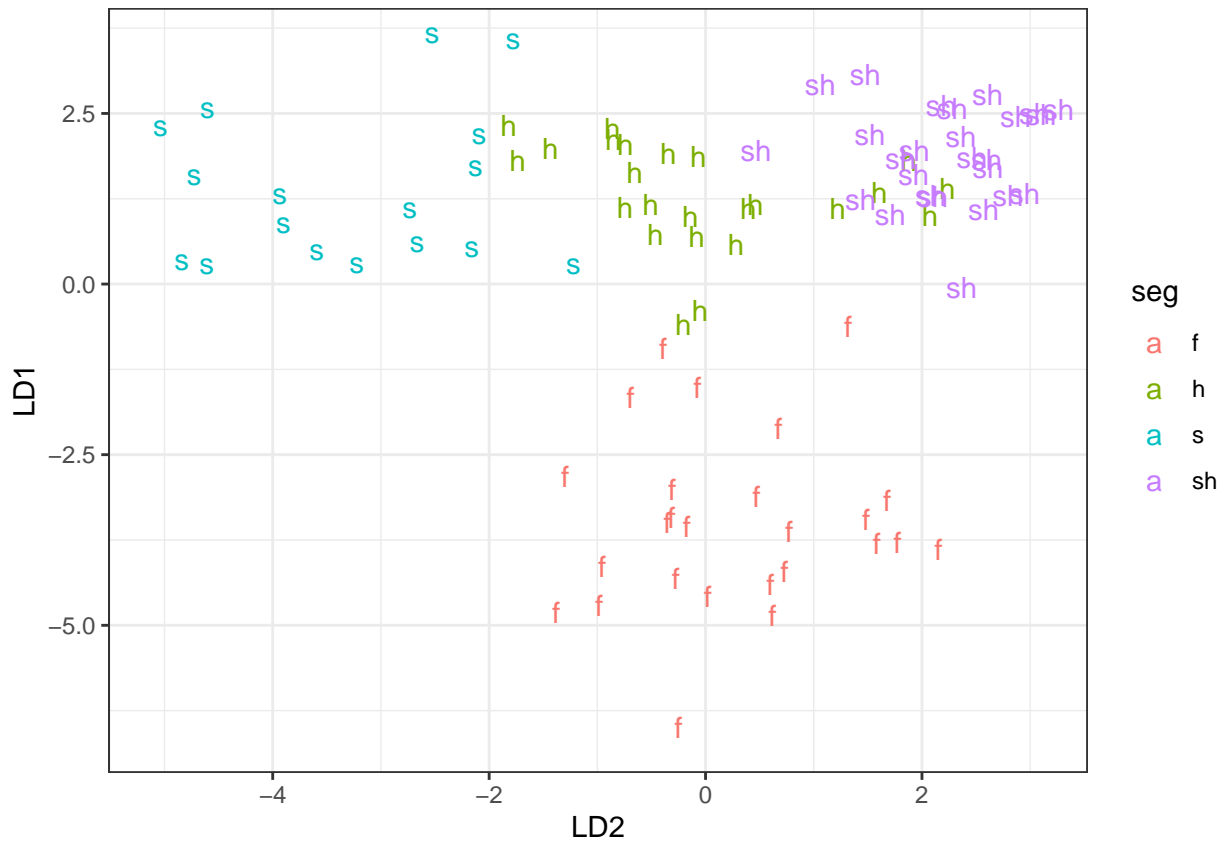
## Plotting pairwise dimensions

```
## plotting pairwise dimensions
```

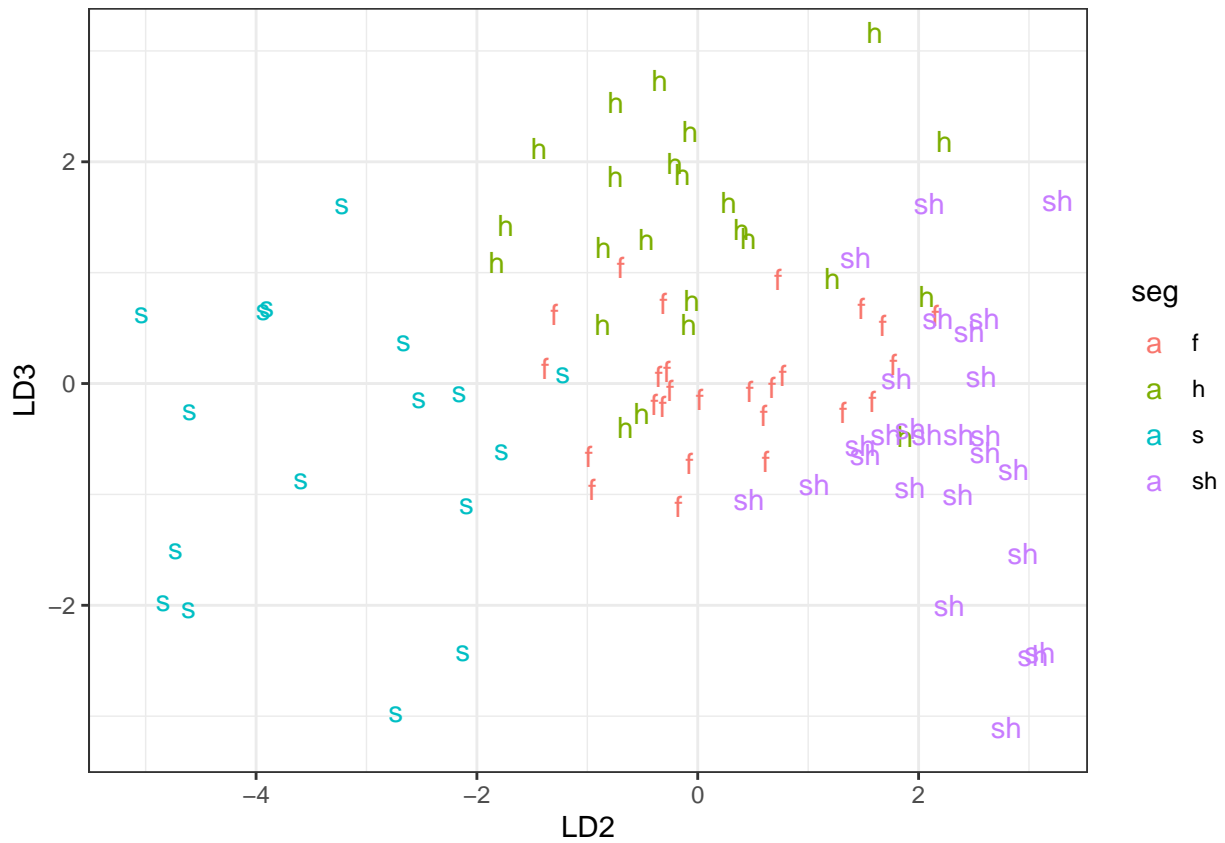
```
lda.data <- cbind(train.trans, predict(model)$x)
```

```
## LD1 vs. LD2
```

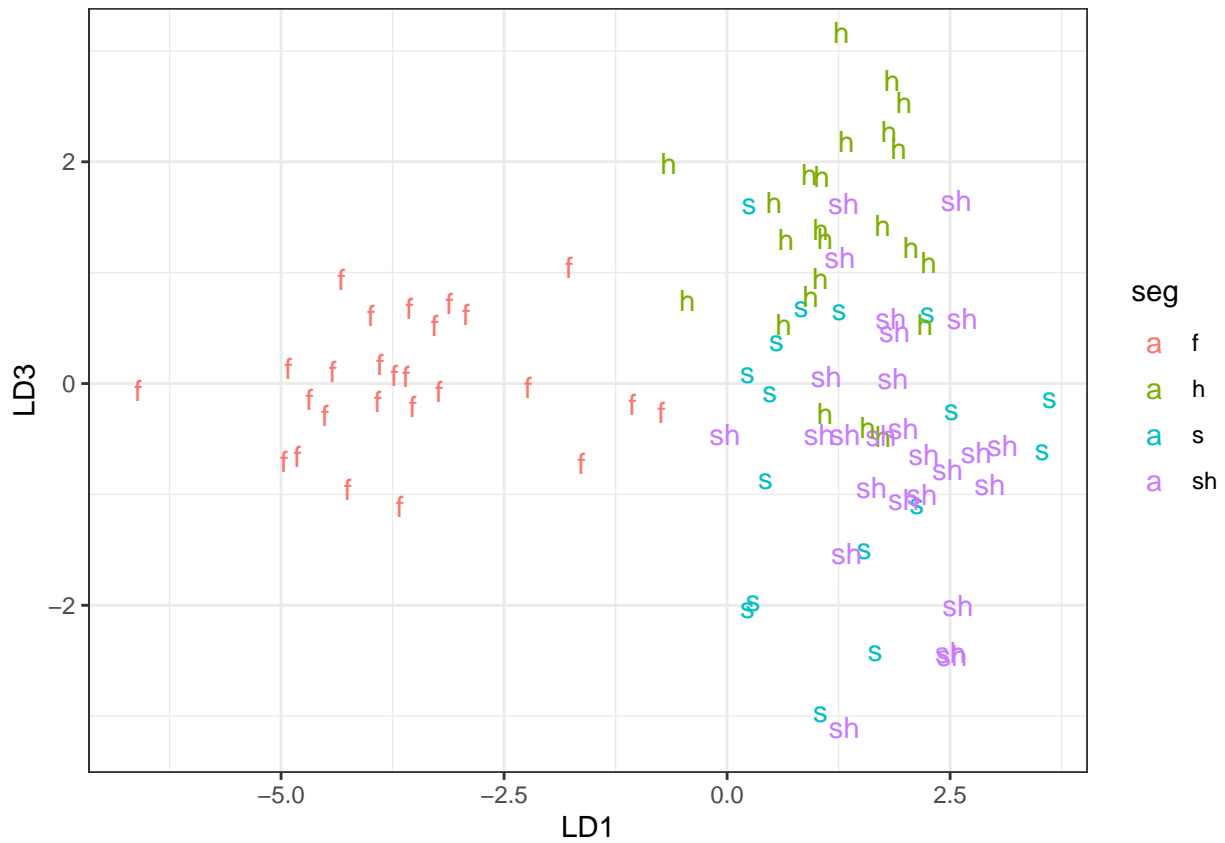
```
ggplot(lda.data, aes(LD2, LD1)) +
  # geom_point() +
  geom_text(aes(label=seg, color=seg), hjust=0, vjust=0) +
  theme_bw()
```



```
## LD2 vs. LD3
ggplot(lda.data, aes(LD2, LD3)) +
  # geom_point() +
  geom_text(aes(label=seg, color=seg), hjust=0, vjust=0) +
  theme_bw()
```



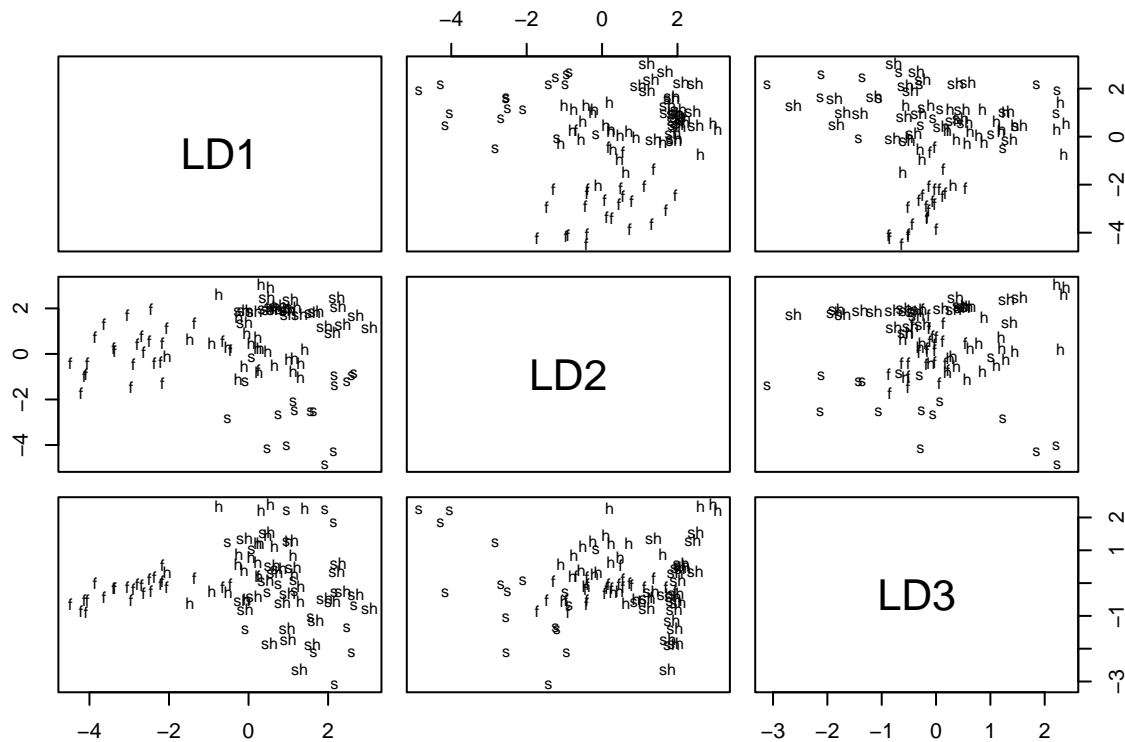
```
## LD1 vs. LD3
ggplot(lda.data, aes(LD1, LD3)) +
  # geom_point() +
  geom_text(aes(label=seg, color=seg), hjust=0, vjust=0) +
  theme_bw()
```



### Running model with only main parameters from Model 1 above

Model with only CoG, Spec Variance and Skewness.

```
model12 <- lda(seg ~ cog + specSD + skew, train.trans)
plot(model12)
```



```
model2
```

```
## Call:
## lda(seg ~ cog + specSD + skew, data = train.trans)
##
## Prior probabilities of groups:
##      f      h      s      sh
## 0.2795699 0.2580645 0.1827957 0.2795699
##
## Group means:
##      cog      specSD      skew
## f  0.32846847  1.2987300 -0.4933478
## h -0.01256057 -0.1566675 -0.3181782
## s  1.22380830 -0.4001138 -0.4775167
## sh -1.11705645 -0.8925009  1.0992732
##
## Coefficients of linear discriminants:
##      LD1      LD2      LD3
## cog  0.55520973 -2.1891935 -1.1902915
## specSD -2.10135418  0.2638420 -0.6052934
## skew  0.04727865 -0.5242007 -2.0079742
##
## Proportion of trace:
##      LD1      LD2      LD3
## 0.5645 0.3950 0.0405
```

```
predictions2 <- model2 %>% predict(test.trans)
mean(predictions2$class == test.trans$seg)
```

```
## [1] 0.8070175
```

```
## contingency table
preds2 <- table(test.trans$seg, predictions2$class)
preds2
```

```
##
##      f  h  s sh
## f  13  2  1  0
## h   2 11  0  2
## s   0  2  8  0
## sh  0  1  1 14
```

```
diag(prop.table(preds2, 1))
```

```
##      f      h      s      sh
## 0.8125000 0.7333333 0.8000000 0.8750000
```

```
# total percent correct
sum(diag(prop.table(preds2)))
```

```
## [1] 0.8070175
```