Study of the voicing contrast in English affricates
due Tue 12/1

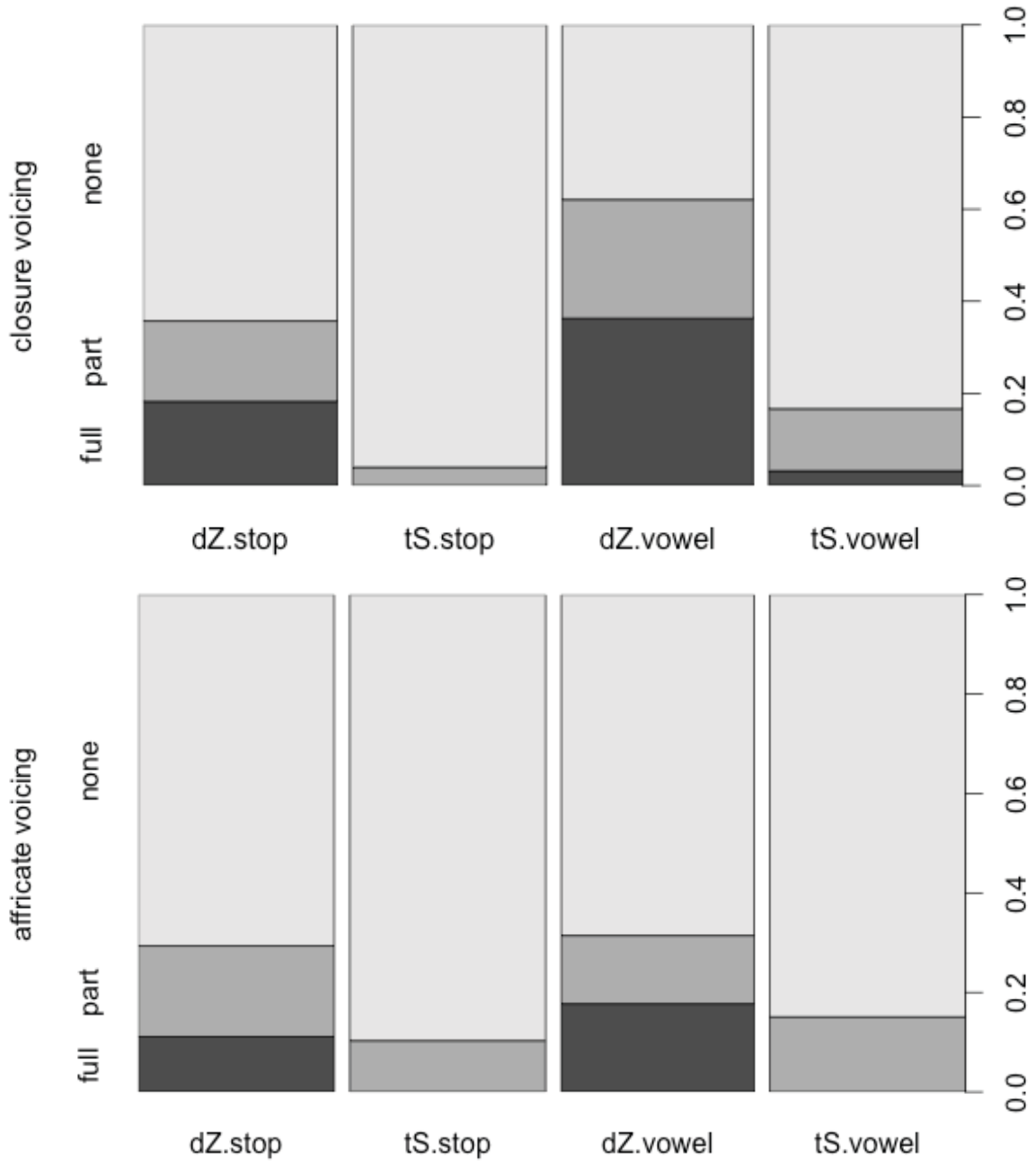
Assignment: Write up our study of voicing contrasts in English affricates. The full results are available on the class website in an excel spreadsheet (let me know if you prefer some other format).

Voicing in affricates

We are studying the realization of the contrast between the affricates /tʃ/ and /dʒ/ because a preliminary investigation indicated that /dʒ/, although usually regarded as a voiced affricate, can be voiceless, and seems to be quite consistently voiceless in certain contexts, e.g. in utterance initial position or after a voiceless consonant. It is well established that underlyingly voiced stops /b, d, g/ are usually realized as voiceless in these contexts, and that the contrast with voiceless /p, t, k/ is maintained by aspirating the voiceless stops [p^h, t^h, k^h].

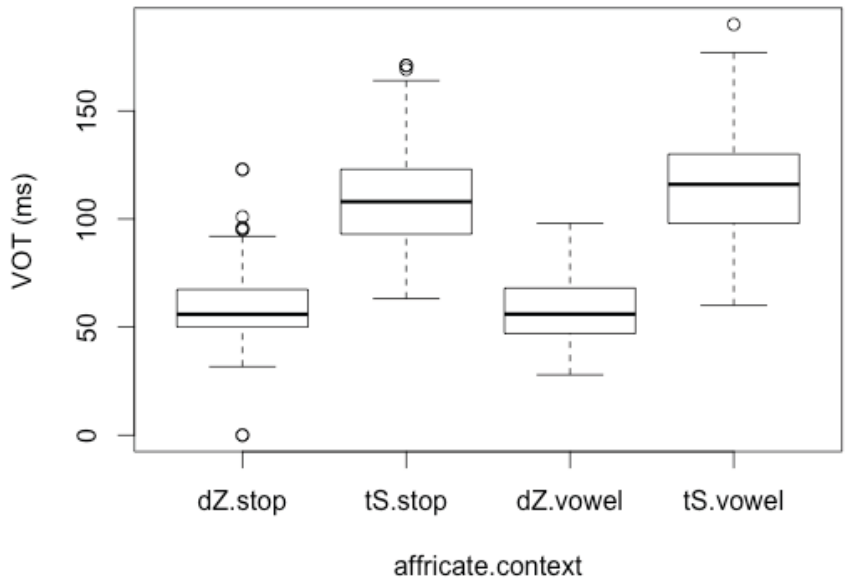
We looked at affricates in two environments: after a voiceless stop and after a vowel. The expectation was that voiced affricate /dʒ/ would be devoiced after a voiceless stop, but that we might see voiced realizations between vowels. Did this happen?

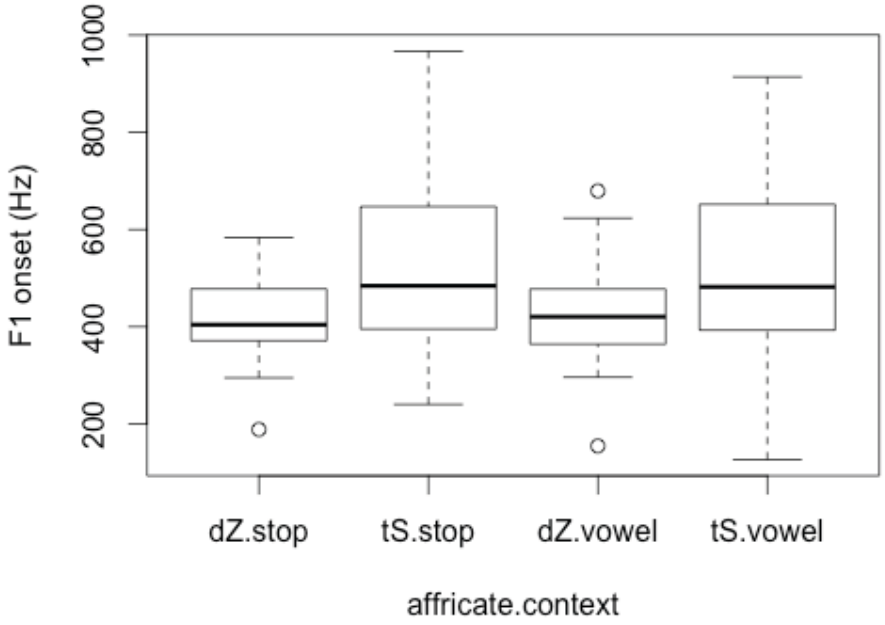
The following figures show the proportion of affricates produced with full vs. partial vs. no voicing during the stop closure and the frication portions of the affricates. /tʃ/ is labeled 'tS' and /dʒ/ is labeled 'dZ', 'stop' indicates a preceding stop and 'vowel' indicates a preceding vowel. These data are a bit rough – I had to guess what people meant by some of the labels used, but the basic pattern seems clear: the stop phase of /dʒ/ is realized with more voicing after vowels than after stops, but is still often completely voiceless. Voicing of the frication of /dʒ/ does not appear to be affected by context. In most cases frication is completely voiceless, and the voiced frication is only observed in a few subjects. So we were not generally successful in eliciting fully voiced affricates – it seems to be common to devoice /dʒ/ even between vowels, at least with word-initial /dʒ/.

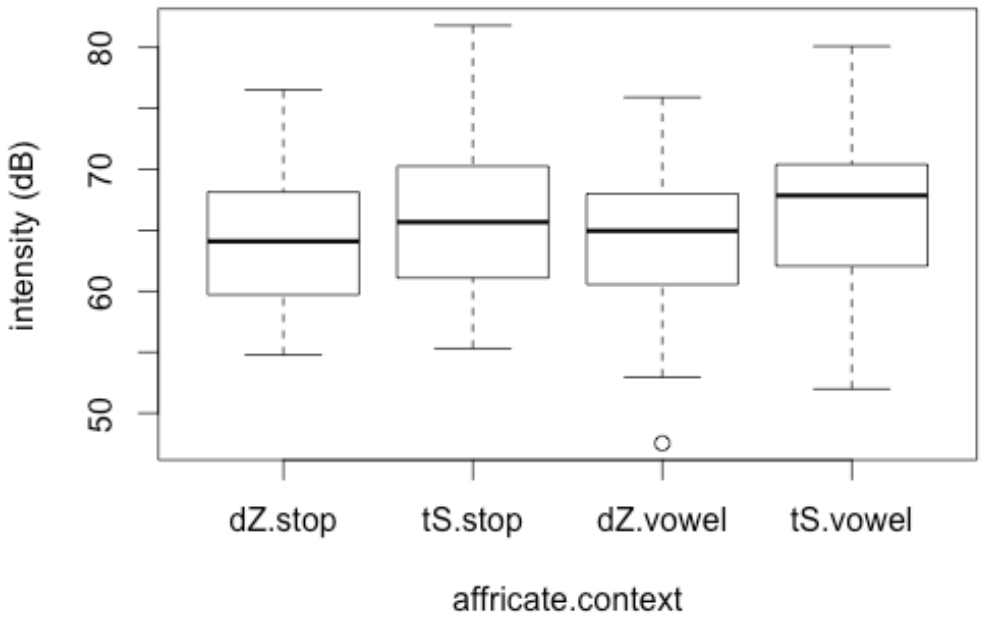


Plots of results

The following 'box and whisker' plots summarize the data visually. The bottom of each box is the 25th percentile of the group of data, the middle line is the 50th percentile (median), and the top of the box is the 75th percentile. The 'whiskers' show the range of the data, up to a maximum of 1.5 times the interquartile range above and below the median. Any points outside that range are plotted individually ('outliers').


VOT (ms)
0
50
100
150
dZ.stop
tS.stop
dZ.vowel
tS.vowel
affricate.context


F1 onset (Hz)
200
400
600
800
1000
dZ.stop
tS.stop
dZ.vowel
tS.vowel
affricate.context



Means and standard deviations for each measurement by affricate and context:

VOT:

context	affricate	mean VOT (ms)	VOT s.d.
stop	dʒ	58	18
vowel	dʒ	58	15
stop	tʃ	109	22
vowel	tʃ	116	22

Peak frication intensity:

context	affricate	mean intensity (dB)	s.d.
stop	dʒ	64.1	5.4
vowel	dʒ	64.3	5.6
stop	tʃ	66.3	6.1
vowel	tʃ	67.2	5.8

F1 onset:

context	affricate	mean F1 onset (Hz)	s.d.
stop	dʒ	421	71
vowel	dʒ	425	81
stop	tʃ	535	180
vowel	tʃ	527	170

Statistical analysis:
Linear mixed effects models were fitted to each measure. The fixed effects were affricate (tʃ vs. dʒ), context (preceding stop vs. preceding vowel) and the interaction between them. The models included random effects by subject corresponding to all fixed effects (these allow for speaker-specific variation in the coefficients of these factors), and a

random intercept for each 'rhyme' – a factor that groups together minimal pairs, so e.g. 'cheer' and 'jeer' both have rhyme 'eer'. This factor is intended to account for any effect of the rest of the word on VOT, intensity and F1 onset. E.g. F1 onset is expected to vary as a function of vowel height. This factor is treated as a random effect because the words that we examined are just a sample of the full range of words beginning with affricates.

The models were fitted using the *lmer* function from the *lme4 R* package (Bates et al 2011).

The summary of the model for VOT is shown below

```
Linear mixed model fit by REML ['lmerMod']
Formula: VOT_ms ~ affricate * context + (affricate * context | subject)
+      (1 | rhyme)
   Data: data

REML criterion at convergence: 4149.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.6219 -0.6142 -0.0755  0.5111  5.2508

Random effects:
 Groups   Name                      Variance Std.Dev. Corr
 subject  (Intercept)               130.79   11.436
          affricatetS                91.36    9.558   -0.14
          contextvowel               12.99    3.605   -0.01  0.04
          affricatetS:contextvowel   11.94    3.455    0.33  0.85  0.29
 rhyme    (Intercept)                22.97    4.792
 Residual                           205.26   14.327
Number of obs: 502, groups:  subject, 9; rhyme, 7

Fixed effects:
                          Estimate Std. Error t value
(Intercept)               58.35866    4.40938  13.235
affricatetS               50.58084    3.66190  13.813
contextvowel              -0.07297    2.17479  -0.034
affricatetS:contextvowel   6.81468    2.80548   2.429

Correlation of Fixed Effects:
            (Intr) affrcS cntxtv
affricatetS -0.209
contextvowl -0.176  0.226
affrcttS:cn  0.249 -0.012 -0.474
```

The list of fixed effects shows the estimates of the coefficients for each factor, the standard errors of these estimates, and the *t* value of the coefficient (estimate/standard error). To a reasonable approximation, we can say that coefficients with a *t*-value greater than 2 (or less than -2) are significantly different from 0 with $p < 0.05$, since the

probability of a *t*-value with magnitude greater than 2 is less than 0.05 for 6 degrees of freedom or greater, and we have enough observations that the degrees of freedom should be substantially greater than 6.

In interpreting the fixed effects, bear in mind that the baseline category for `affricate` is /dʒ/, and the baseline category for context is after a stop, so the intercept corresponds to mean VOT for /dʒ/ after a stop, and the remaining factors specify deviations from this baseline. So:

- the coefficient of `affricate` is the difference in VOT between /tʃ/ and /dʒ/
- the coefficient of `context` is the difference in VOT between /dʒ/ after a vowel and /dʒ/ after a stop.
- the coefficient of `affricatetS:contextvowel` is the difference between VOT of /tʃ/ after a vowel and /dʒ/ after a vowel.

A significant interaction between `affricate` and `context` (as is the case here for VOT) means that the difference between the affricates varies significantly across the contexts.

Model for frication intensity:

```
Linear mixed model fit by REML ['lmerMod']
Formula: intensity ~ affricate * context + (affricate * context |
subject) +
    (1 | rhyme)
   Data: data

REML criterion at convergence: 2261.6

Scaled residuals:
    Min      1Q  Median      3Q     Max
-5.3712 -0.5817  0.0170  0.5482  3.3219

Random effects:
 Groups   Name                      Variance Std.Dev. Corr
 subject  (Intercept)               26.0868  5.1075
          affricatetS                3.7417  1.9344    0.38
          contextvowel               1.0358  1.0178    0.15 -0.05
          affricatetS:contextvowel   0.4492  0.6703   -0.58 -0.80 -0.53
 rhyme    (Intercept)                0.1742  0.4174
 Residual                            4.4883  2.1186
Number of obs: 502, groups:  subject, 9; rhyme, 7

Fixed effects:
                          Estimate Std. Error t value
(Intercept)                64.1259     1.7202   37.28
affricatetS                 2.1439     0.6978    3.07
contextvowel                0.3257     0.4324    0.75
affricatetS:contextvowel    0.5576     0.4393    1.27
```

Model for F1 onset:

```
Linear mixed model fit by REML ['lmerMod']
Formula: F1 ~ affricate * context + (affricate + context | subject) +
(1 | rhyme)
   Data: data

REML criterion at convergence: 5831.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.4155 -0.5859  0.0177  0.5726  4.3263

Random effects:
 Groups   Name         Variance Std.Dev. Corr
 subject  (Intercept)  2481.69  49.817
          affricatetS  4651.26  68.200    0.93      Seg So
          contextvowel   24.95   4.995    0.00 -0.36
 rhyme    (Intercept)  5998.58  77.451
 Residual              6027.77  77.639
Number of obs: 502, groups:  subject, 9; rhyme, 7

Fixed effects:
                         Estimate Std. Error t value
(Intercept)               420.964     34.359  12.252
affricatetS               114.056     24.748   4.609
contextvowel                2.181      9.963   0.219
affricatetS:contextvowel   -9.870     13.862  -0.712
```

Reference:

Bates, Douglas; Martin Maechler and Ben Bolker (2011). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-39. http://CRAN.R-project.org/package=lme4

24.915 / 24.963 Quantum Optical Communication

Fall 2015