The effect of frequency and within-speaker variation in non-native speech perception

Marie-Catherine de Marneffe and Meghan Sumner

{mcdm, sumner}@stanford.edu • Linguistics Department • Stanford University

158th Meeting of the Acoustical Society of America, San Antonio, Texas, 26-30 October 2009

Primary questions

Does within-speaker variation in non-native speech help native speakers' perception?

Are listeners better at perceiving final stops when a constant non-English V/C ratio is used by the speaker, or does hearing variable ratios help?

Background & problem

➲ Research in speech perception and L2 learning has shown that exposure to multiple speakers helps adaptation to non-native speech [Barcroft & Sommers 2005, Bradlow & Bent 2008, Winters et al. 2005]
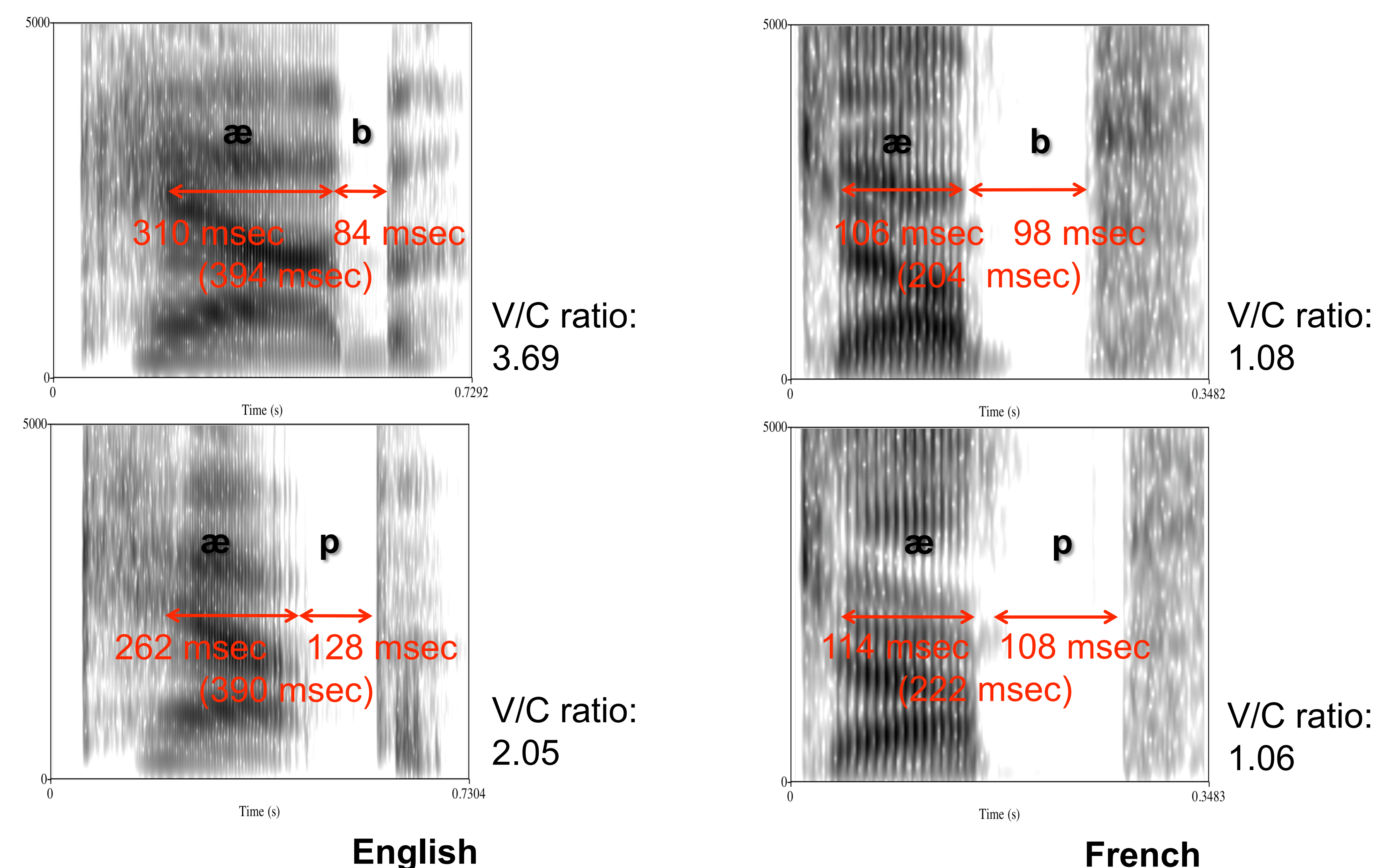
➲ **Question:**
Is variation within the speech of one single speaker also beneficial?

➲ We study perception by native English speakers of French-accented English speech, and concentrate on final obstruent devoicing. We know that

(1) English speakers use V/C ratio as cue [Denes 1955, Port & Dalby 1982]

(2) V/C ratios for voiced and voiceless stops used in English and in French differ. The contrast on which English native speakers rely is therefore lost in French-accented English.

Fig. 1: English and French V/C ratios

310 msec 84 msec
(394 msec)
V/C ratio: 3.69

108 msec 98 msec
(204 msec)
V/C ratio: 1.08

262 msec 128 msec
(390 msec)
V/C ratio: 2.05

114 msec 108 msec
(222 msec)
V/C ratio: 1.06

English   **French**

➲ We manipulate V/C ratios, approximating real speech by using the ranges of V/C ratios appearing in natural speech (in a narrative reading task).

Experiment

Participants
48 native English speakers (Stanford University undergrads)

Stimuli
☛ Critical items:
20 pairs of pictures of voiced/voiceless words controlled for frequency (tag – tack, cab – cap)

☛ Voiced words from one native French speaker, late English learner

☛ 2 conditions: constant V/C ratios (average)
                  variable V/C ratios (short – average – long)

☛ Modified natural tokens: for each V+C ending, we use the ranges in the ratios found in the speech of 11 native French speakers
- the *average* ratio is the median of these ratios
- for the *short* and *long* ratios, we subtracted/added from the median the standard deviation divided by 2

☛ Fillers:
40 pairs which differ word-initially by one sound (lush – rush)
20 pairs which differ word-finally by one sound (moss – moth)

Procedure
☛ Familiarization phase for all items:
Subject sees a picture and the word associated with it

tag

☛ Click-on paradigm:
3 blocks of pairs of pictures + sound

Results

- Overall there is no effect of condition.

- Restricting the data to the best subjects (higher quartile), variable ratios hurt.

- In the case of the worst subjects, variability appears to help.

- For words ending in *g* (vs. *b* and *d*), there is a significant effect of condition (p-value < .05):

- for the best subjects, variable ratios hurt
- for the worst subjects, variable ratios help

Fig. 2: Overall accuracies per condition

constant ratios
variable ratios

all subjects   best subjects   worst subjects
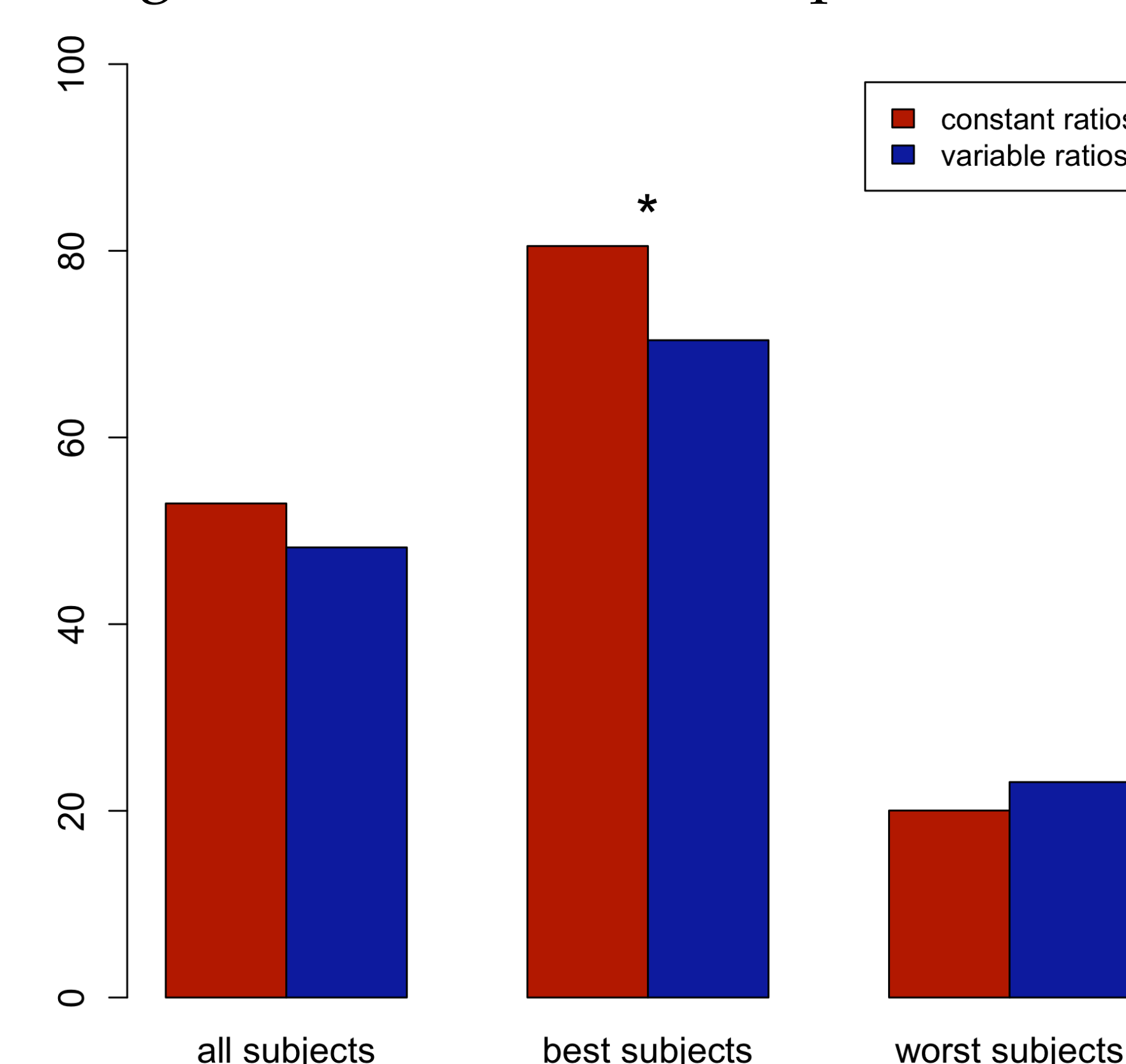
Frequency & variability

In a logistic regression predicting accuracy on all the subjects (with subjects and items as random effects),

➲ We find significant main effects of (p-value = .00):

- frequency ratio (voiced word frequency/voiceless word frequency) bigger ratios lead to better accuracy
- frequency of the words clicked on by the subject

➲ The condition is not a significant factor

➲ But there is a significant interaction between the frequency ratio and the condition (p-value < .05):

Variability starts to lead to higher accuracy when the frequency ratio between the voiced and voiceless words becomes bigger.

Fig. 3: Accuracy (all subjects) given frequency ratio (voiced/voiceless word)

constant ratios
variable ratios
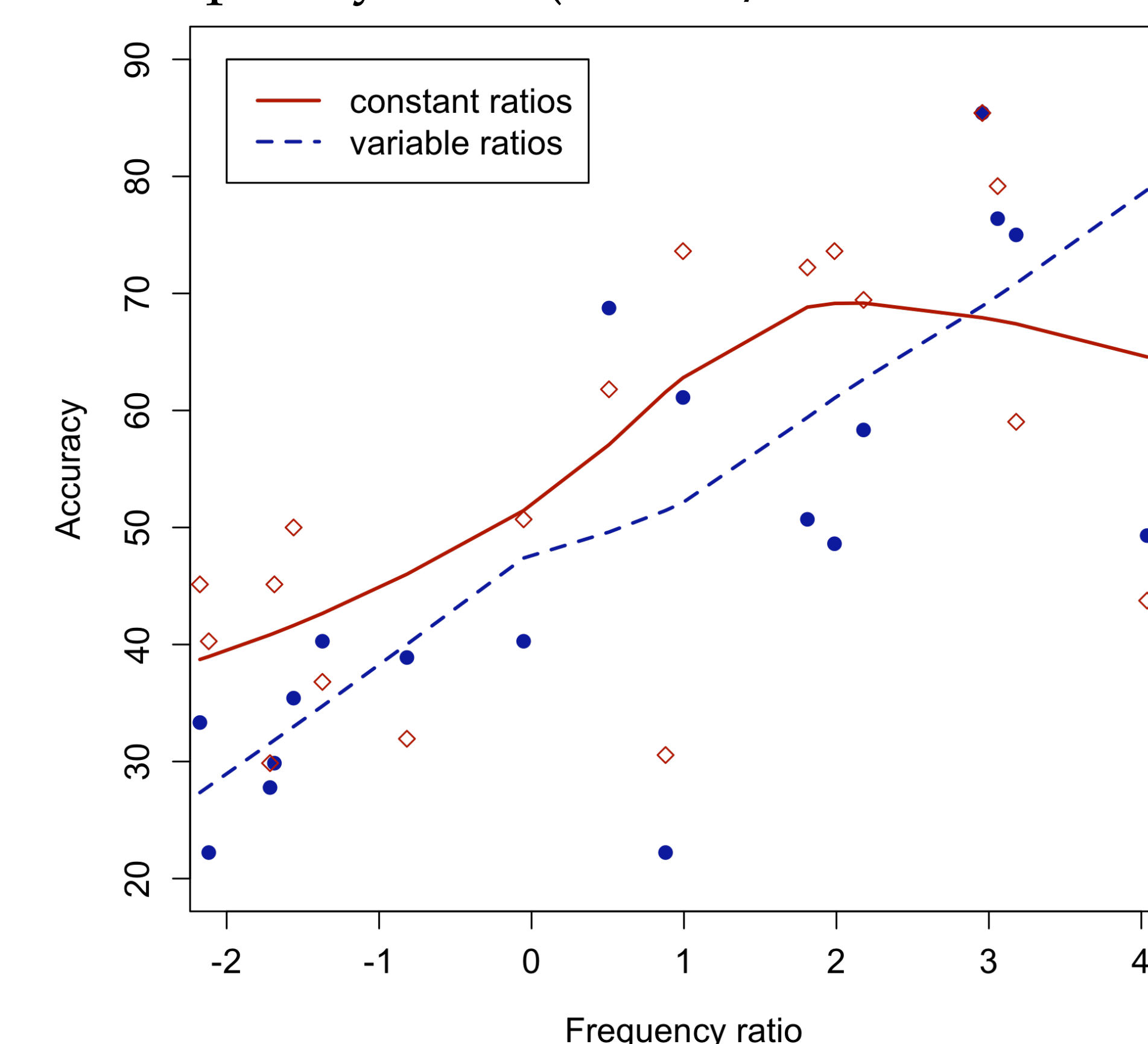
2 items (cod, cad) have been excluded from all the analyses because their accuracies were extremely low.

Conclusion & follow-up work

➲ Frequency of the words have a clear impact on perception: more frequent words are better perceived.

➲ Frequency intertwined with variable V/C ratios also helps perception.

➲ **Work in progress**
There is no clear indication yet that variation in non-native speech per se has an impact on perception. However for the worst subjects there is a trend indicating that variation within-speaker might play a role. The current measures might not be sensitive enough.

We are currently running the same experiment using a mouse-tracking paradigm.