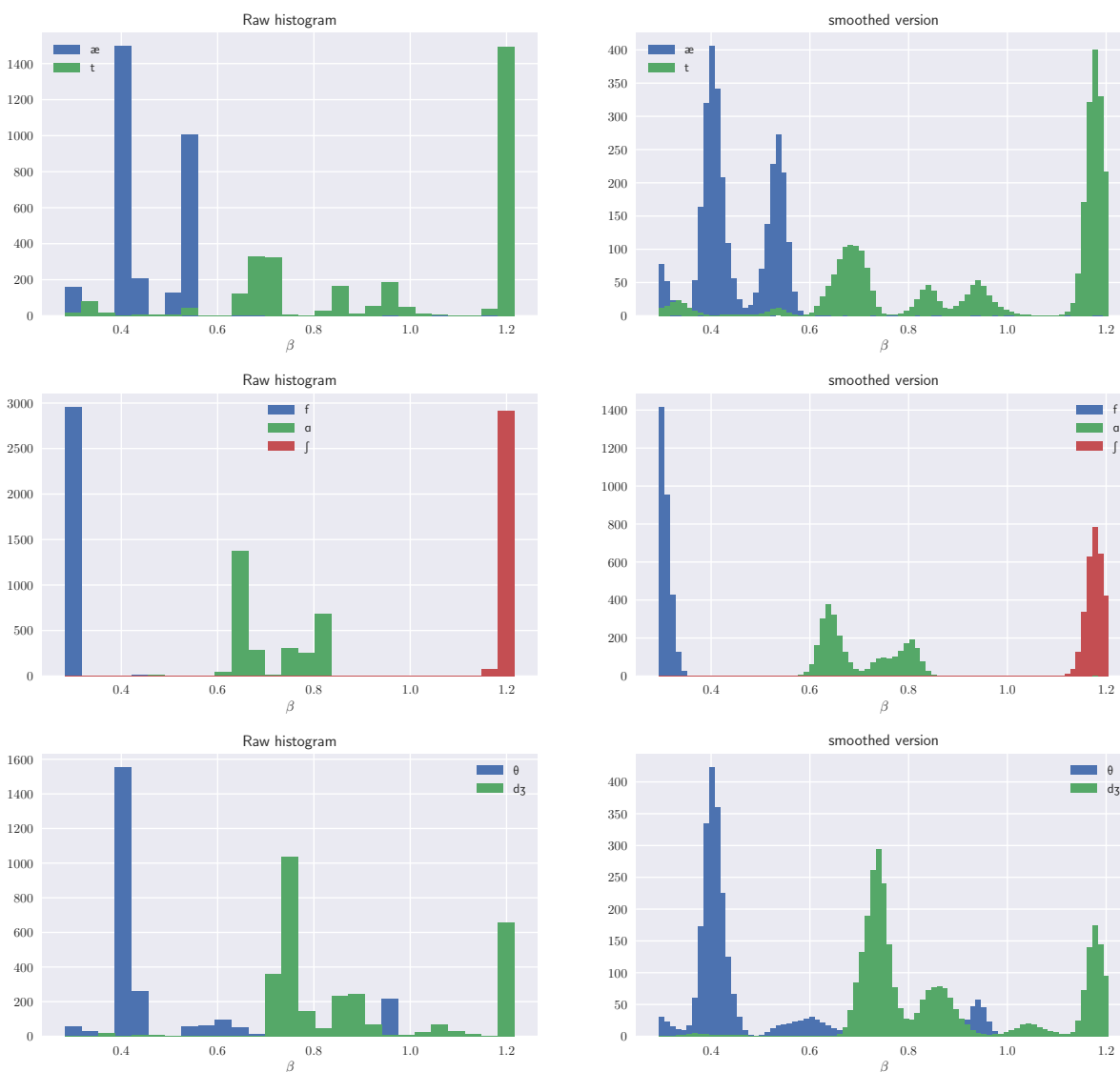


# SUPPLEMENTARY MATERIALS

## BOOTSTRAP DISTRIBUTIONS

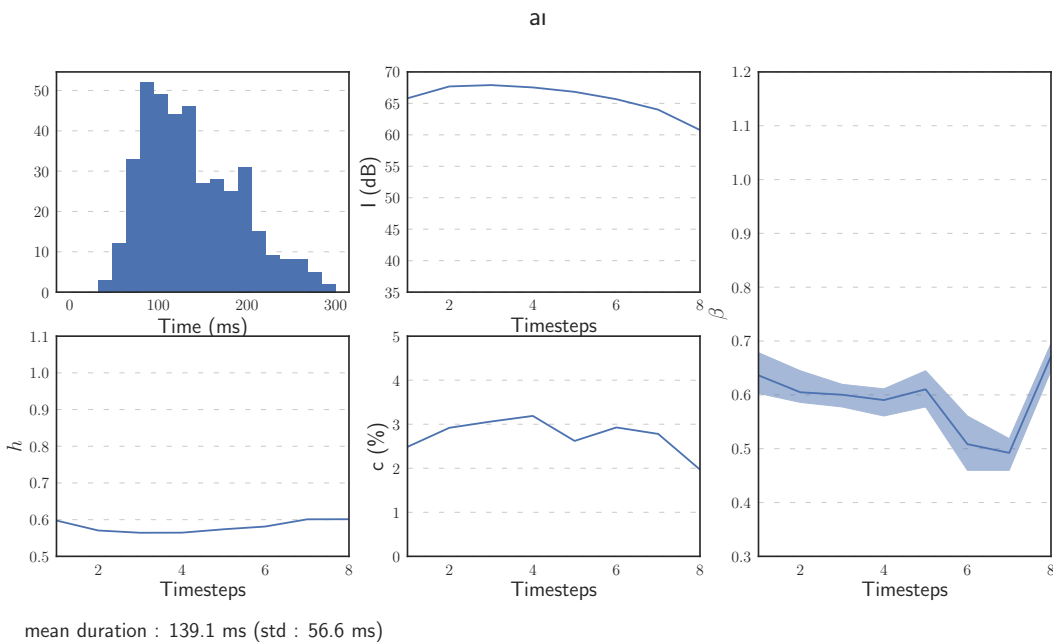
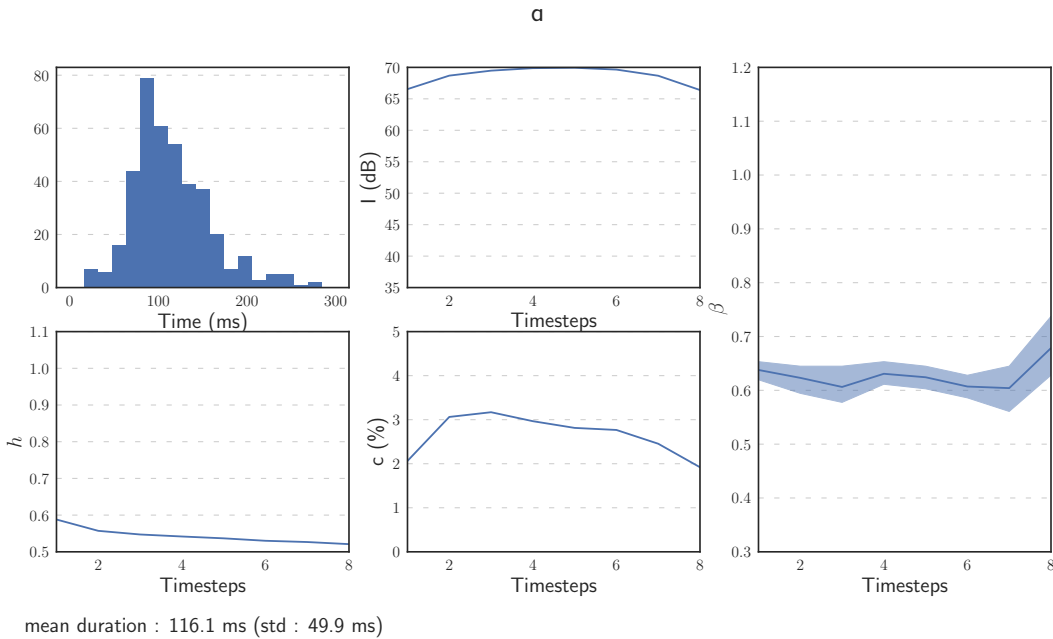


**Figure:** Bootstrap distributions for phonemes æ,t (up), f, α, ʃ (middle), θ, and dʒ (bottom). The procedure is the so-called smoothed percentile bootstrap.  $\beta^*$  is computed after resampling with repetition of 400 slices from the TIMIT database (this is done 3000 times). This results in a first histogram (left) which is smoothed with a Gaussian filter ( $\sigma = 0.015$ ) to obtain the final bootstrap distribution (right). The mean value and the bootstrap confidence intervals are retrieved from the distribution (ex: 70% CI is chosen so that 15% of the samples are excluded left and right).

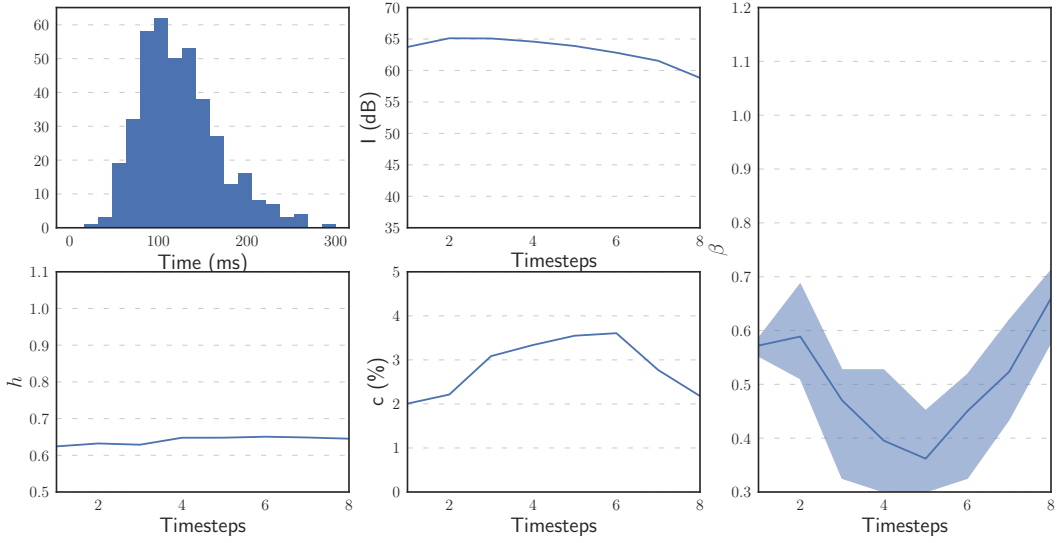
## TEMPORAL EVOLUTIONS OF $\beta$

**Top left:** Duration histograms for the occurrences retrieved from the TIMIT database (absolute time), **top center:** intensity  $I$  (dB) over time, **bottom left:** entropy score  $h$  over time, **bottom center:** contrast  $c$  over time, **right:** exponent  $\beta$  over time.

### Vowels and Diphtongs



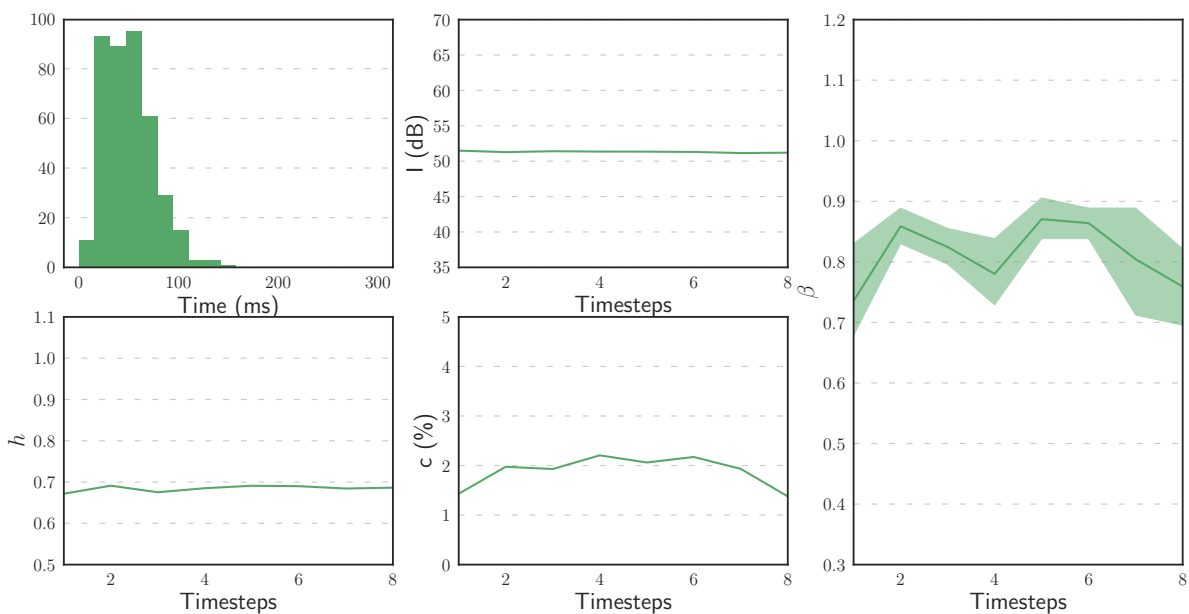
ei



mean duration : 126.7 ms (std : 50.8 ms)

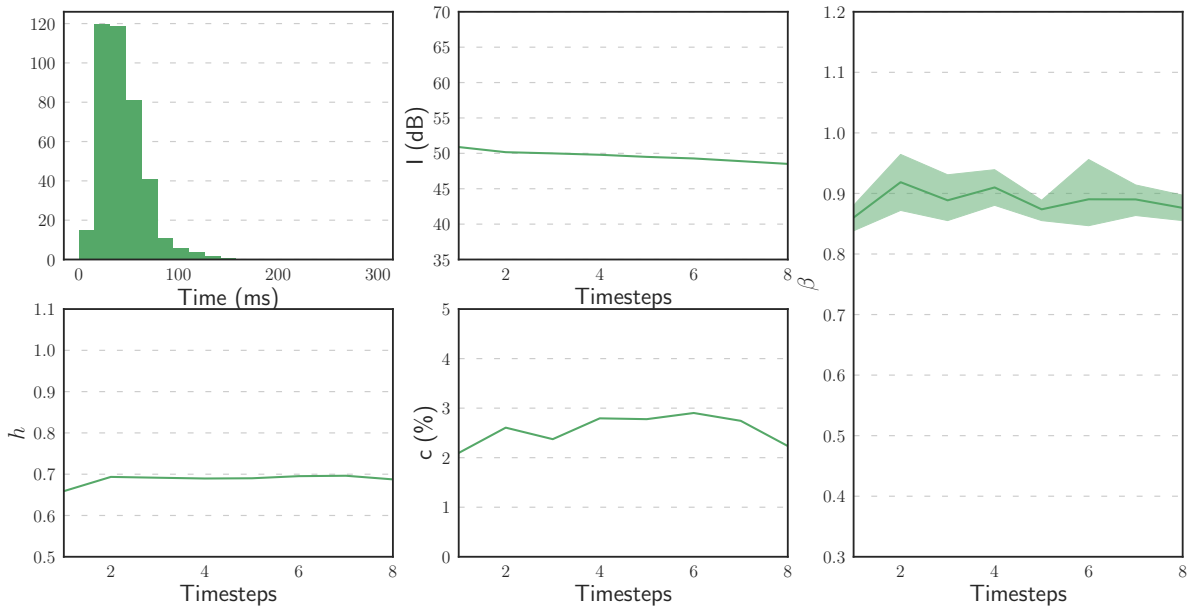
# Nasals

m



mean duration : 50.9 ms (std : 24.6 ms)

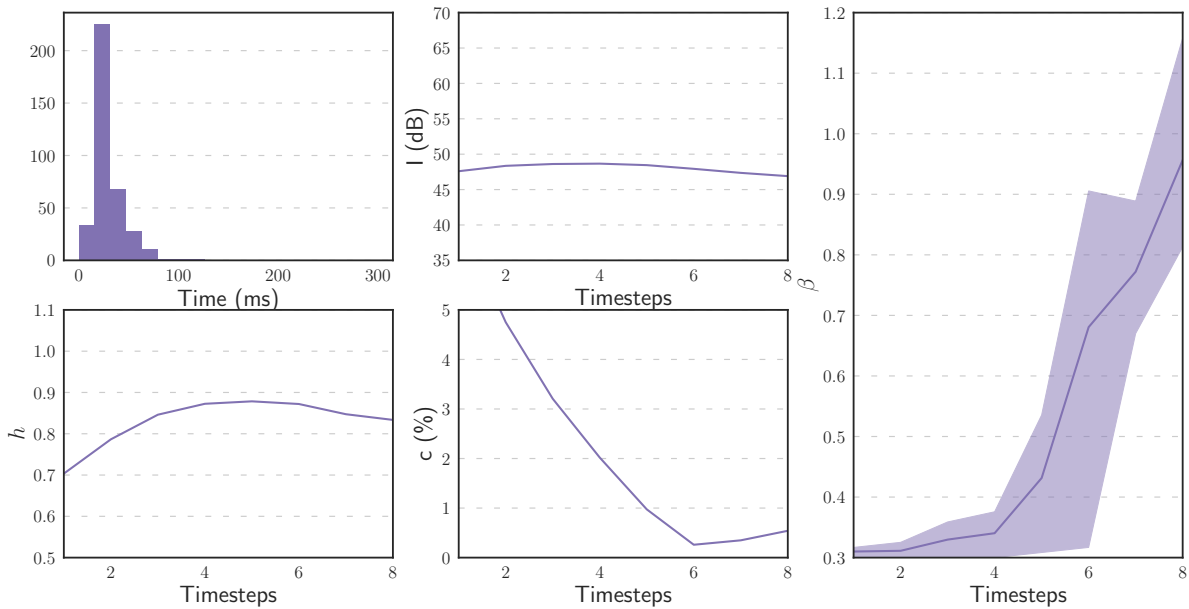
n



mean duration : 44.0 ms (std : 22.2 ms)

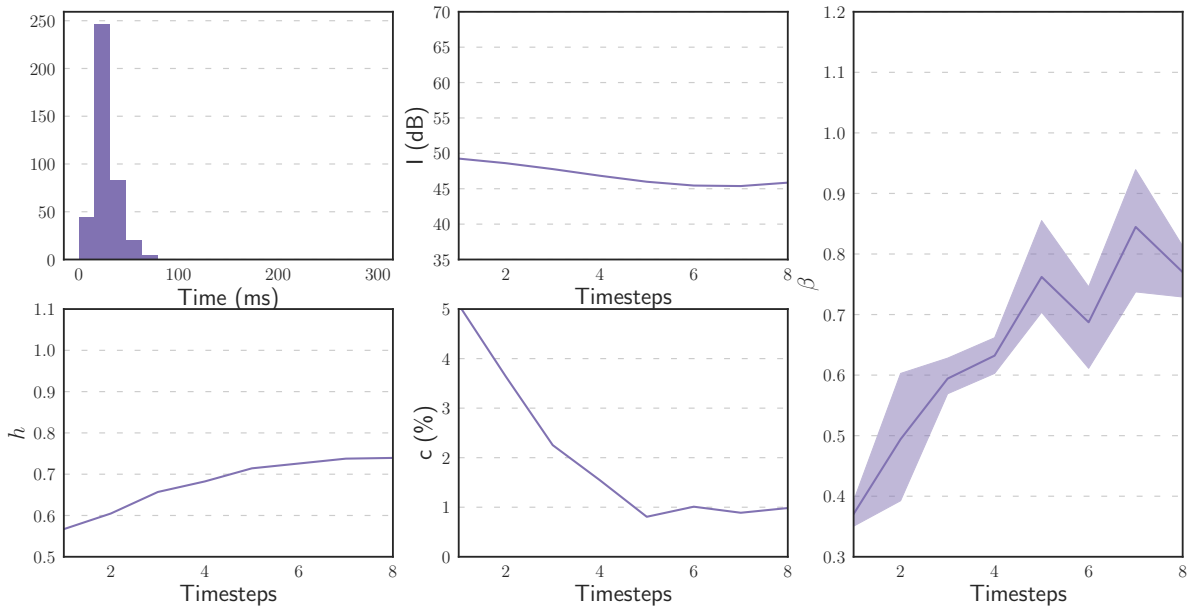
# Stops

d



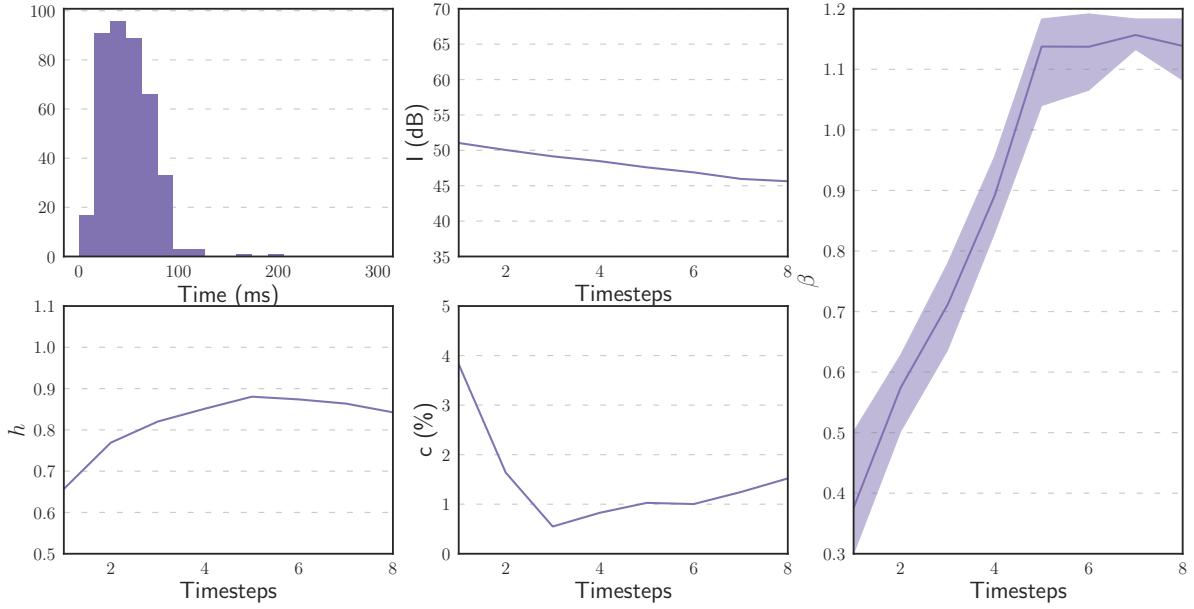
mean duration : 28.4 ms (std : 14.9 ms)

g



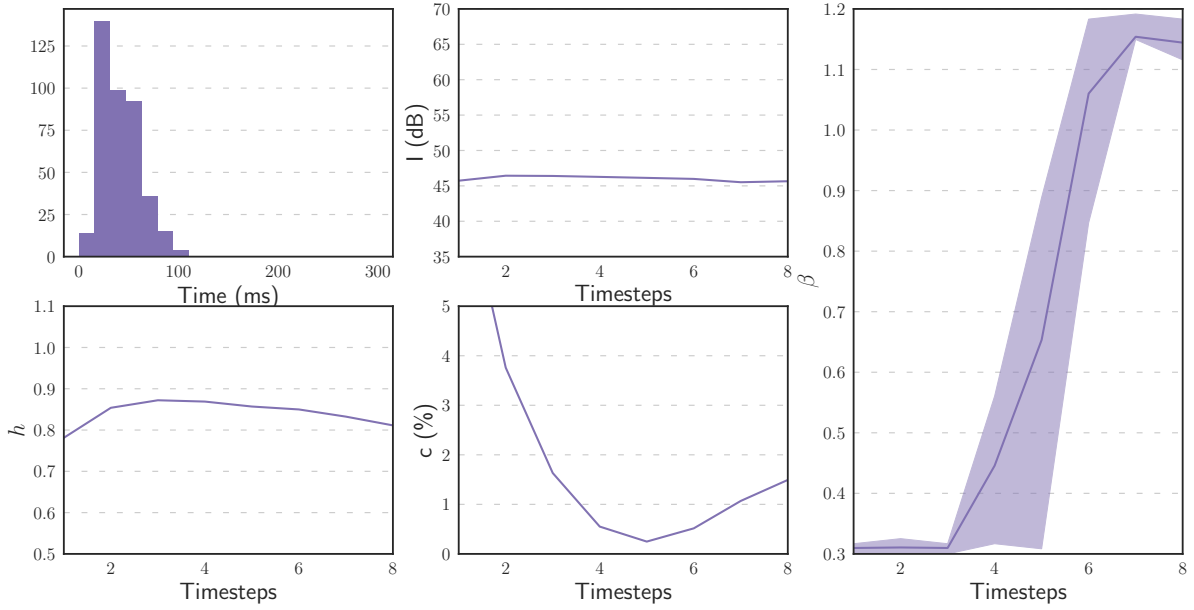
mean duration : 26.5 ms (std : 11.1 ms)

k



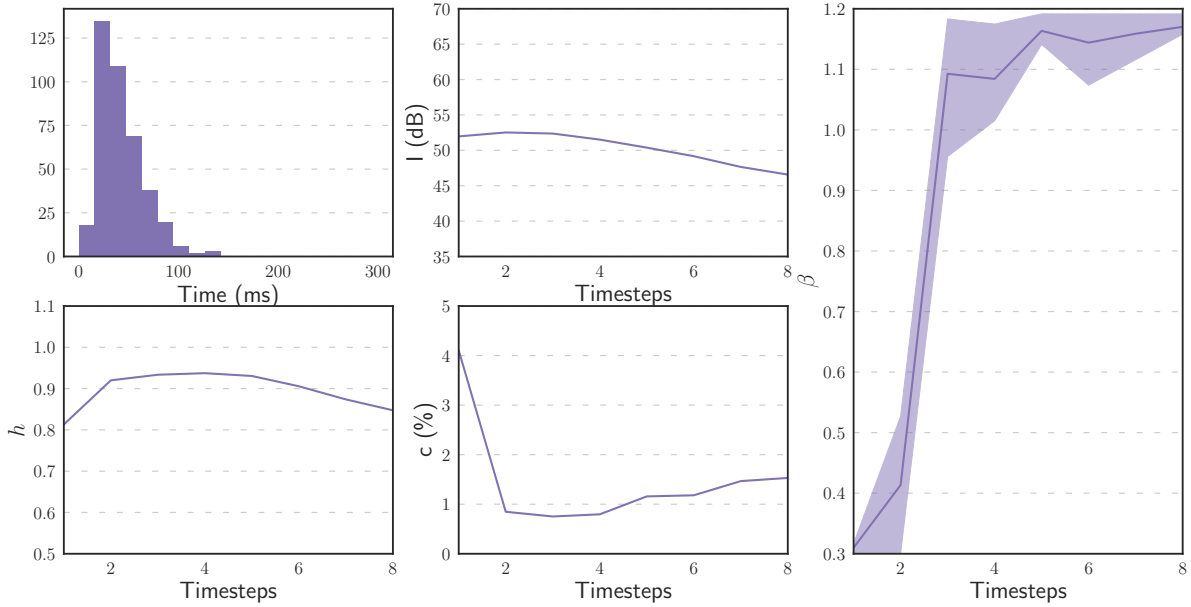
mean duration : 48.9 ms (std : 23.8 ms)

p



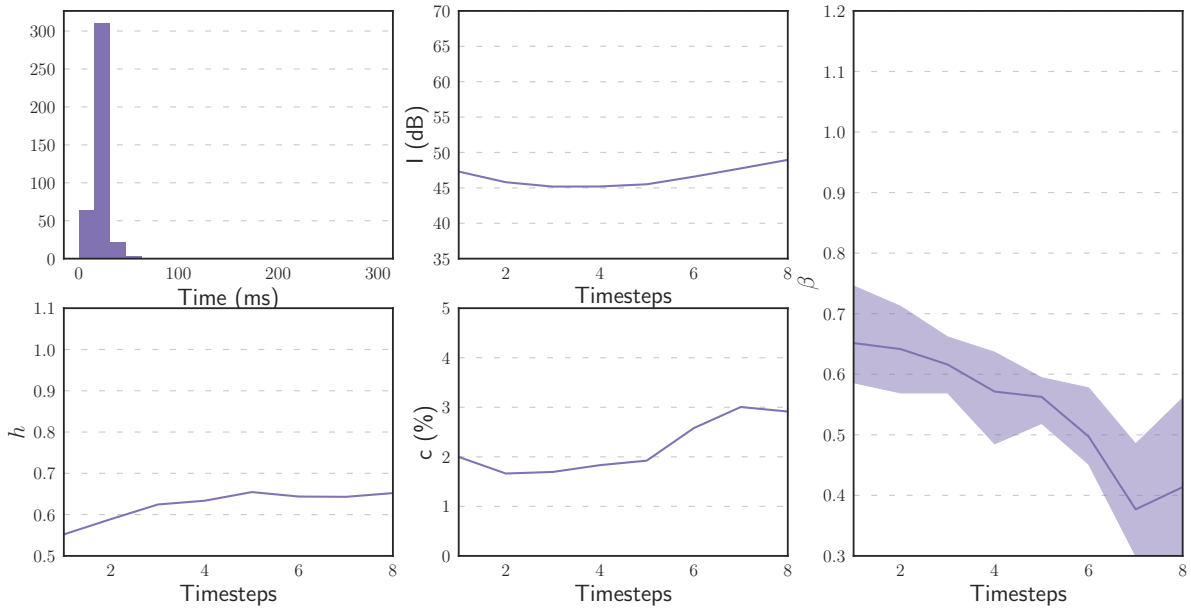
mean duration : 41.9 ms (std : 19.4 ms)

t



mean duration : 42.9 ms (std : 22.3 ms)

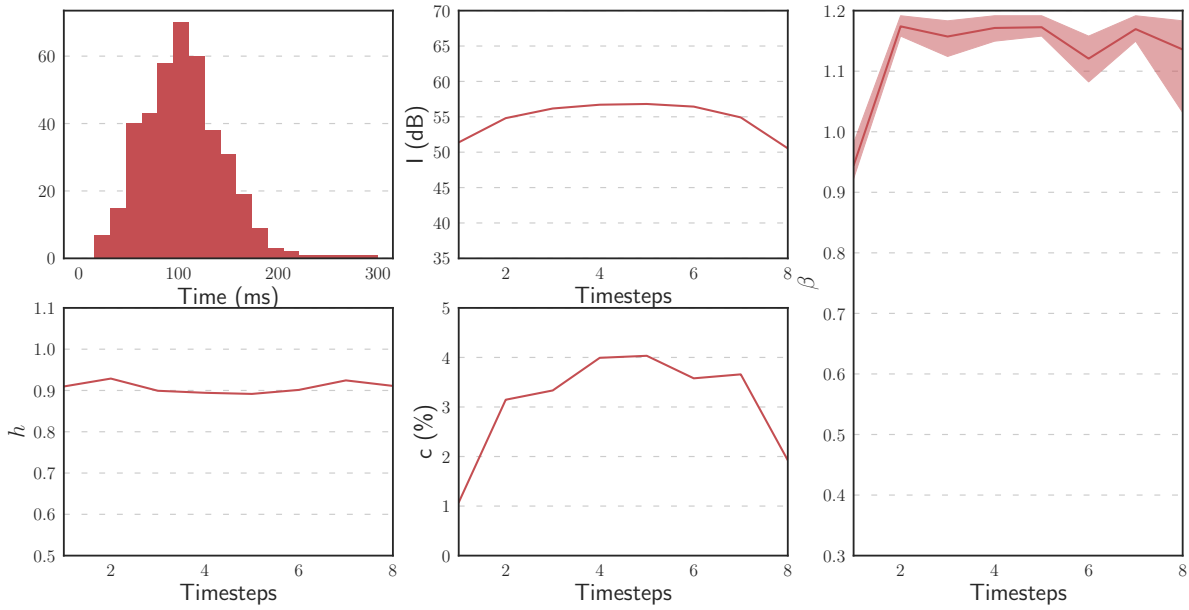
r



mean duration : 21.1 ms (std : 6.6 ms)

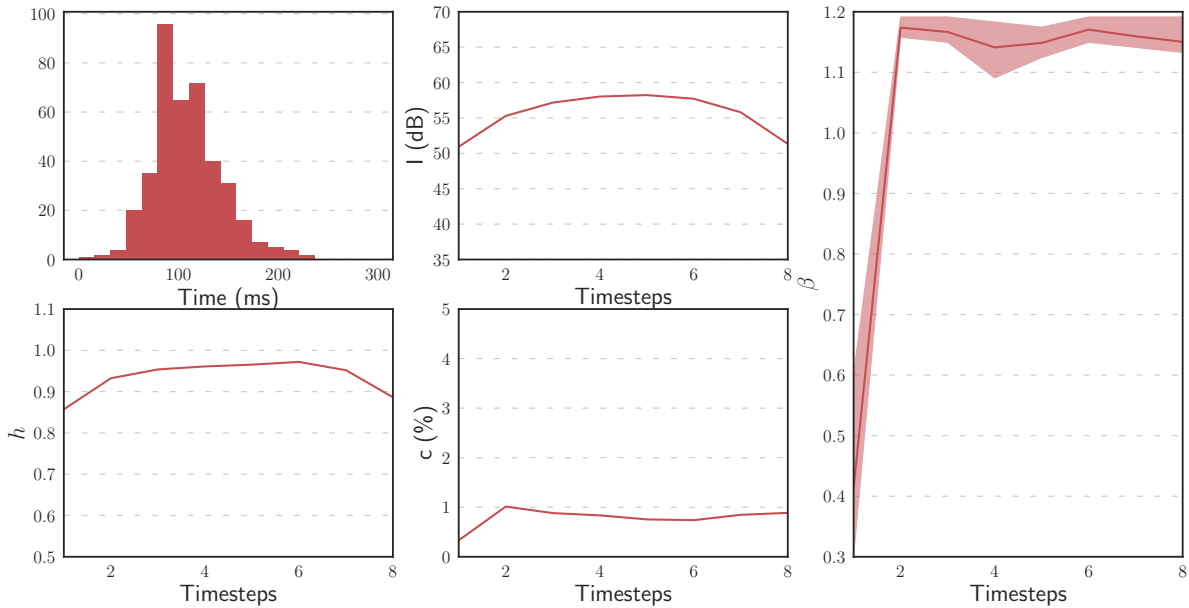
# Fricatives

s



mean duration : 105.8 ms (std : 41.2 ms)

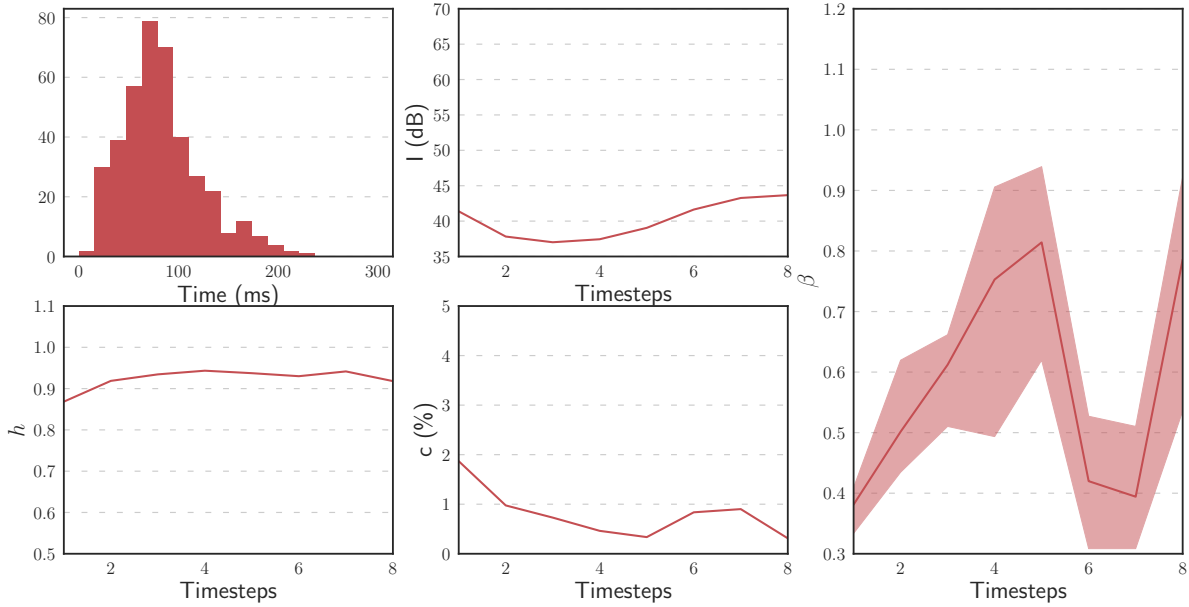
ʃ



mean duration : 108.8 ms (std : 34.1 ms)

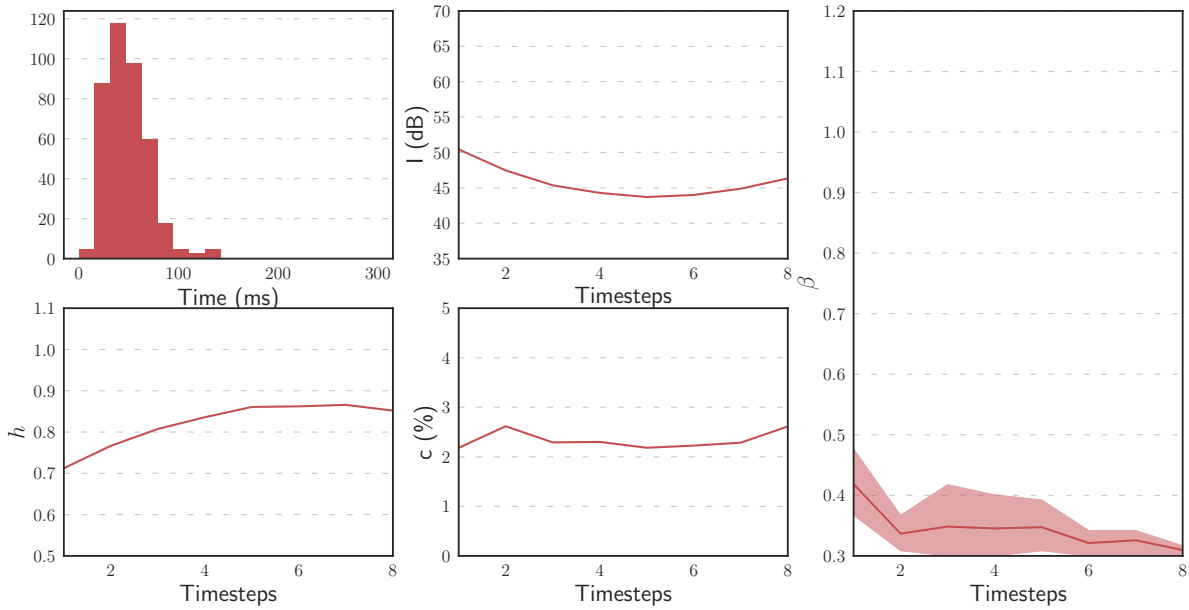


$\theta$



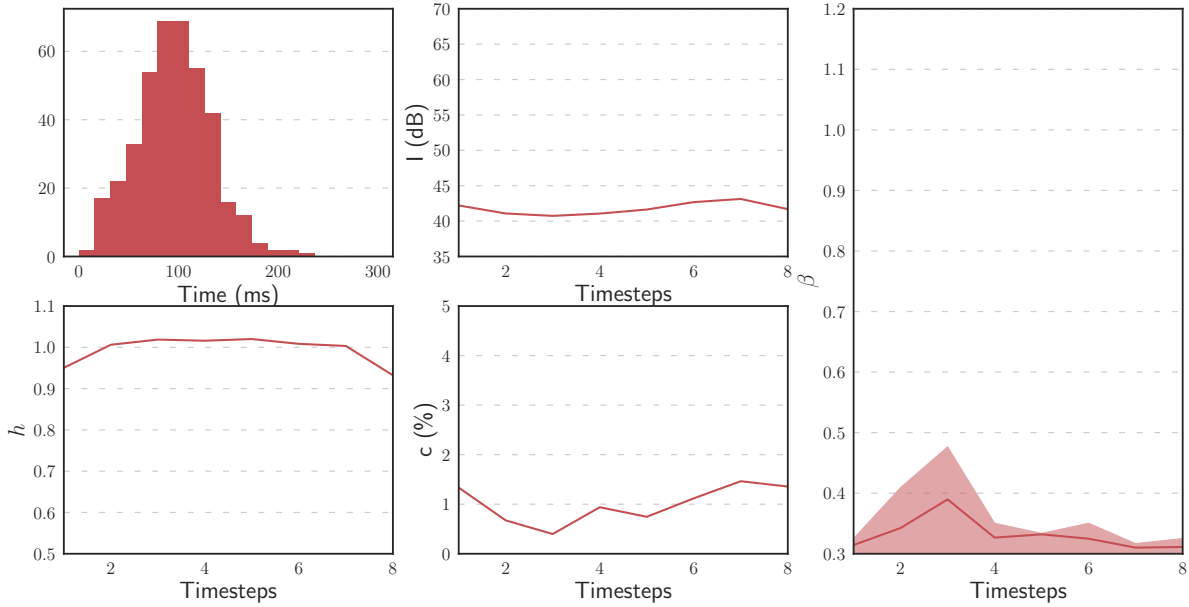
mean duration : 83.4 ms (std : 39.7 ms)

$\nu$



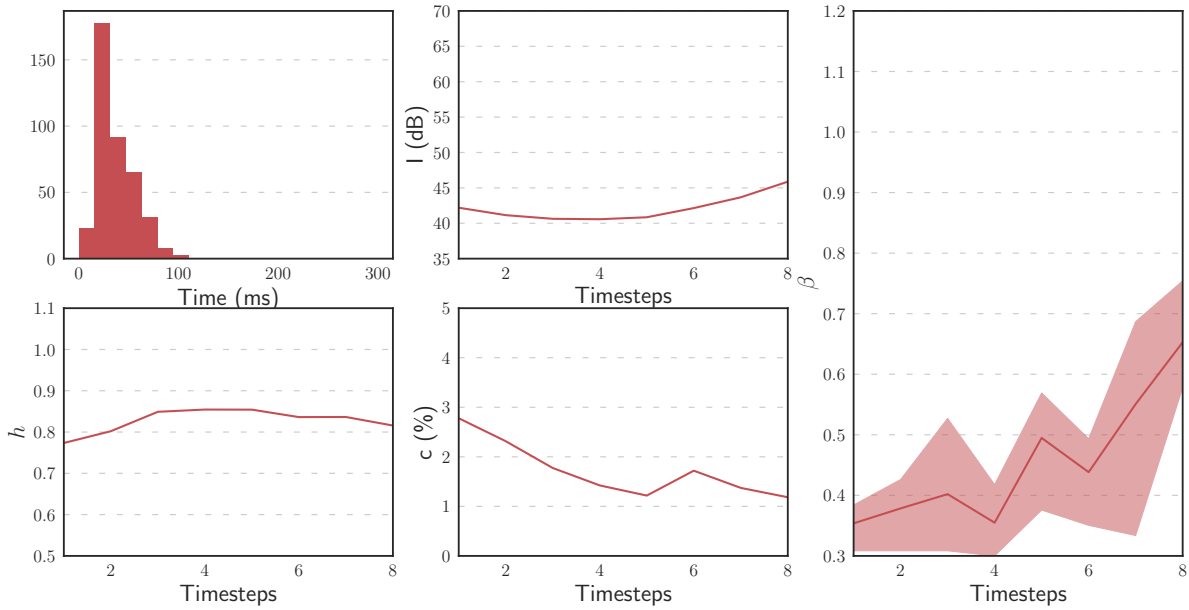
mean duration : 49.1 ms (std : 21.9 ms)

f



mean duration : 95.8 ms (std : 37.3 ms)

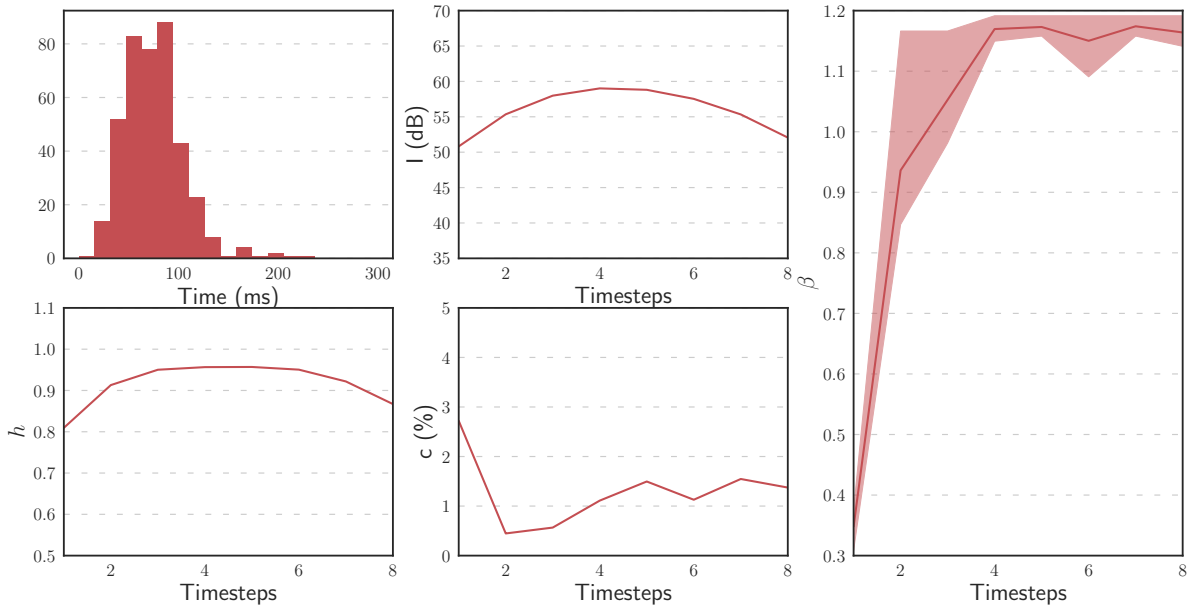
$\delta$



mean duration : 36.4 ms (std : 18.4 ms)

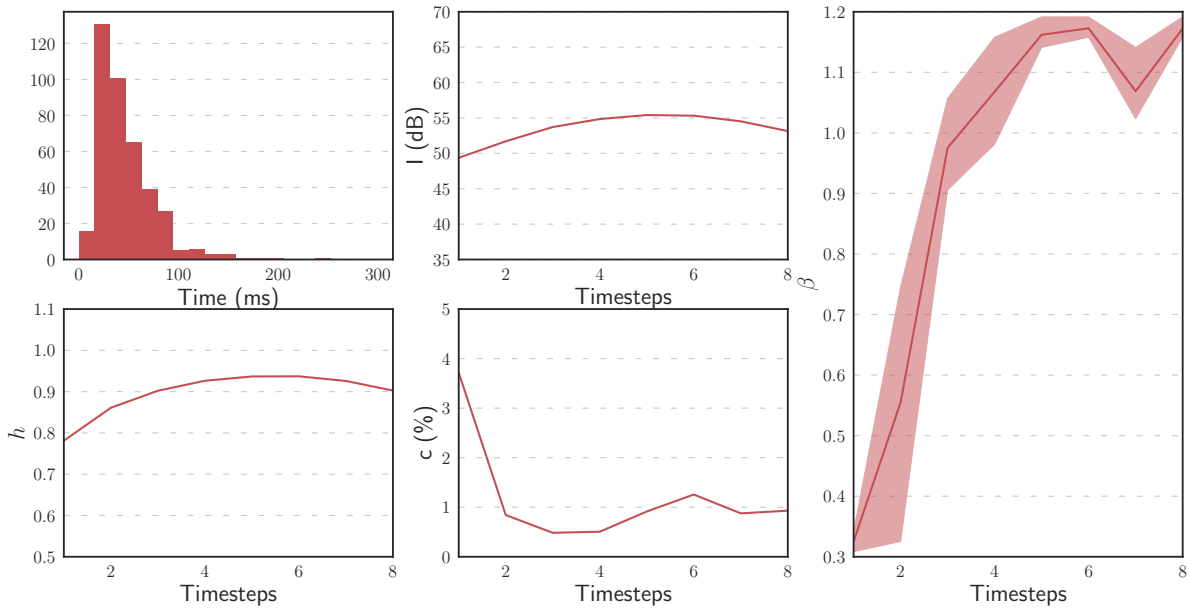
# Affricates

tʃ



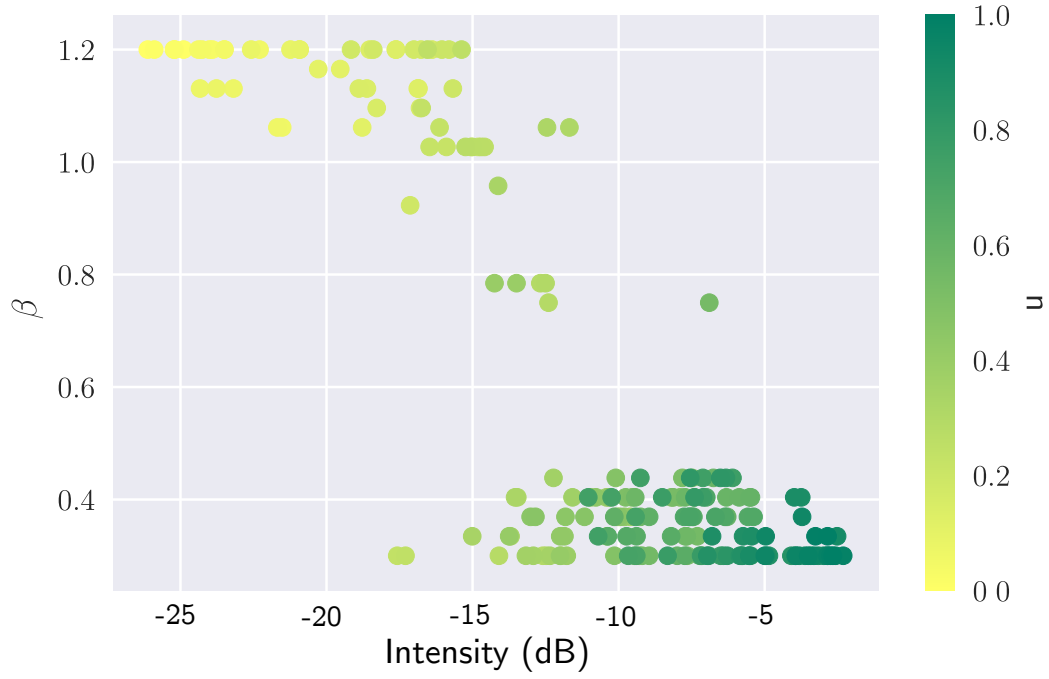
mean duration : 75.1 ms (std : 30.4 ms)

dʒ



mean duration : 47.0 ms (std : 29.7 ms)

$\beta$  W.R.T INTENSITY FOR SYNTHETIC VOWEL-LIKE SOUNDS



**Figure:** Scatter plot of simulated samples on the  $(I, \beta)$  plane: exponent  $\beta$  against intensity  $I$  in dB (ref:max). Each point is a sample of Simulation 2 on synthesized vowels. The control parameter  $u$  determines the aperture of the cylindrical waveguide from  $r = 0.2\text{cm}$  ( $u = 0$ ) to  $r = 1.3\text{cm}$  ( $u = 1$ ).