

An Adaptive Sampling Procedure for Speech Perception Experiments

Interspeech 2006 ICSLP – Paper 1147

Geoffrey Stewart Morrison, University of Alberta

Problem

- In synthetic-speech perception experiments, as the number of dimensions increases, the number of stimuli increases exponentially.
- A relatively large number of responses per stimulus is desirable for building an accurate model of a listener's speech perception.
- Listeners can quickly become fatigued.

Proposed Solution

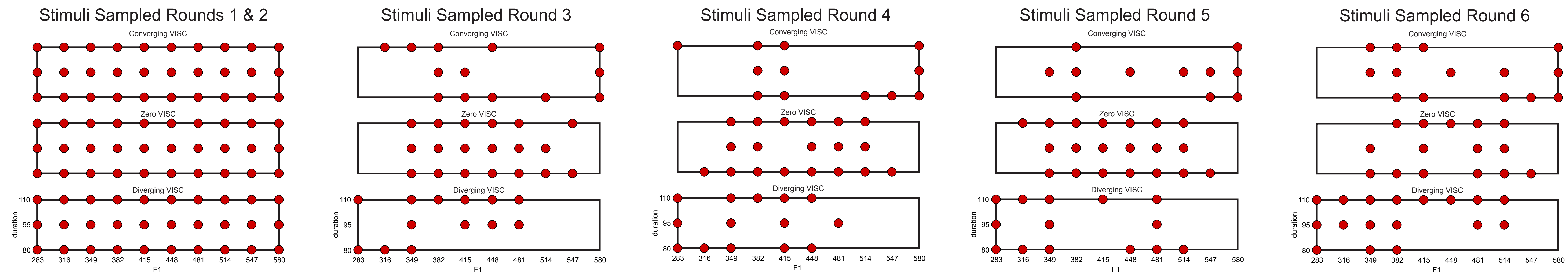
- Stimuli far from any boundaries will have an expected probability of close to 1 for one category and 0 for all other categories.
- This will be so irrespective of the number of times that they are identified by the listener.
- Once these stimuli have been identified, there is no further benefit from resampling them.
- Stimuli near boundaries will have expected probabilities not close to 0 or 1.
- These stimuli must be sampled numerous times in order to represent the probabilities with reasonable resolution.

Illustrative Example

- Categories: English /i/ /ɪ/ /e/ /ɛ/
- Dimensions:
 - covarying F1–F2 (10 levels)
 - vowel inherent spectral change (3 levels)
 - vowel duration (3 levels)
- Total number of stimuli: 90
- Number of trials completable in 30 minutes: 360 = 4 responses per stimulus

Aim

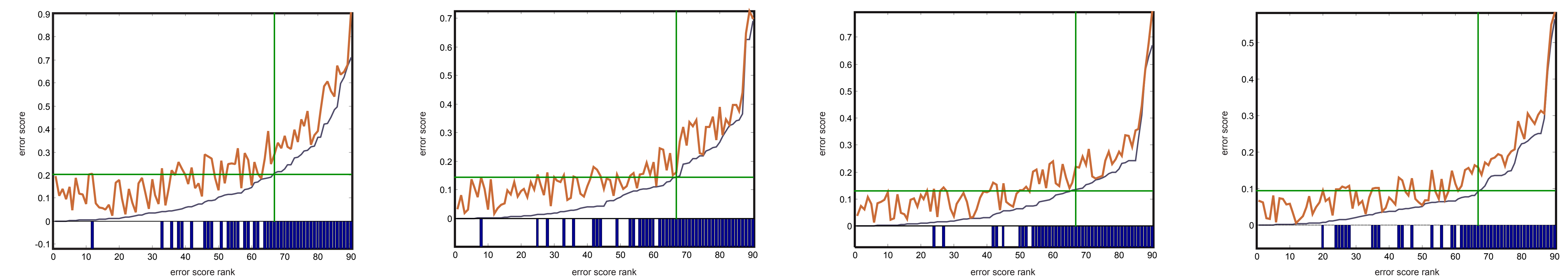
– Obtain a sampling resolution comparable to 6 responses per stimulus (580 responses) in a time of 30 minutes.



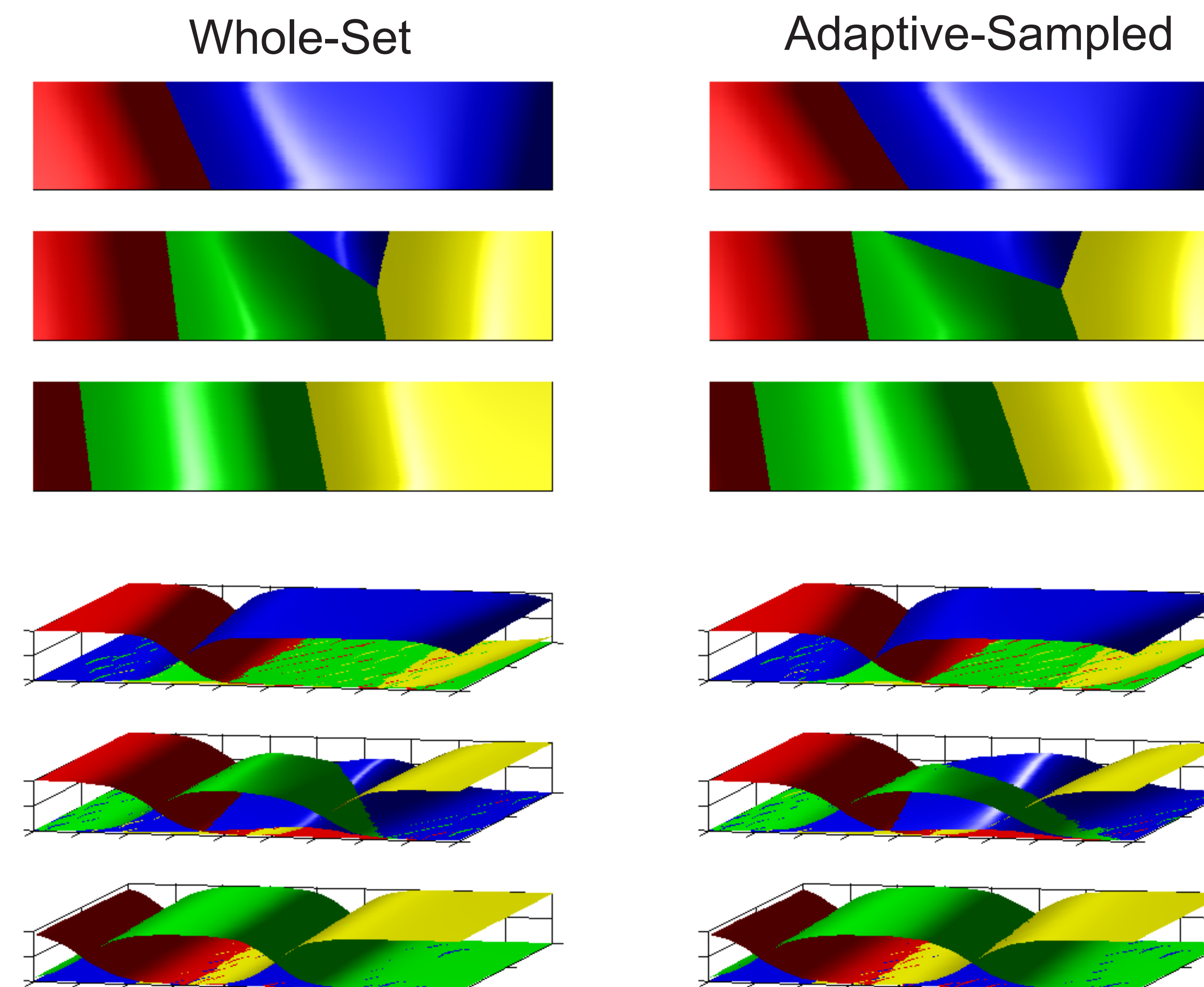
Procedure

1. All stimuli are presented twice in random order, the listener identifies the stimuli (180 responses).
2. A logistic regression model is fitted to the response data, and the predicted probabilities of each category for each stimulus are calculated.
3. The error between the observed and predicted probabilities for each stimulus is calculated.
4. Half the stimuli are resampled (45 responses).
 - 4.1 stochastic noise is added to the error scores
 - 4.2 the stimuli with the highest error+noise scores are resampled
5. Steps 2 through 4 are repeated three more times.

Total number of responses: 360
 Minimum number of responses for a stimulus: 2
 Maximum number of responses for a stimulus: 6



Probability Plots



/i/ /e/ /ɪ/ /ɛ/

Simulations

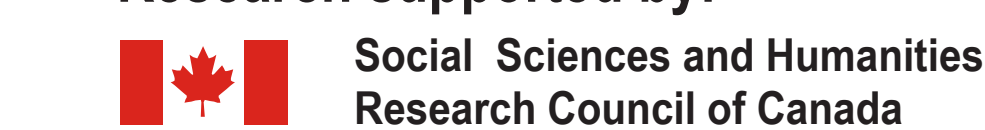
- one listener identified each stimulus 6 times
- logistic regression model fitted, a posteriori probabilities used as parameter values in a multinomial sample generator
- 100 samples of 6 responses per stimulus generated
- “whole-set” and “adaptive-sampled” models fitted

Results:

- differences in coefficient values either not significant or small in magnitude

Error or Coefficient	Sampling Procedure		Difference			
	Whole-Set Mean	Adaptive Mean	Mean (sd)	%	t(99)	p
SAEP	6.813	6.793	-0.020 (0.426)	-0.3	-0.471	.6386
i	34.113	32.923	-1.190 (1.349)	-3.5	-8.823	.0000 **
ɪ	7.141	7.074	-0.068 (0.906)	-0.9	-0.748	.4562
e	-8.147	-8.181	-0.034 (0.792)	+0.4	-0.426	.6708
i×F1	-0.077	-0.074	0.003 (0.003)	-3.6	10.047	.0000 **
ɪ×F1	-0.007	-0.007	0.000 (0.002)	-2.0	0.850	.3975
e×F1	0.012	0.012	0.000 (0.001)	-1.8	-1.742	.0846
i×ΔF1	-2.028	-1.939	0.089 (0.164)	-4.4	5.461	.0000 **
ɪ×ΔF1	1.510	1.486	-0.025 (0.137)	-1.6	-1.802	.0746
e×ΔF1	-3.445	-3.384	0.061 (0.189)	-1.8	3.217	.0018 **
i×dur	-0.037	-0.036	0.001 (0.006)	-3.3	2.020	.0461 *
ɪ×dur	-0.020	-0.021	-0.001 (0.004)	+3.3	-1.612	.1101
e×dur	0.044	0.045	0.001 (0.005)	+1.2	0.999	.3202

Research supported by:



Conseil de recherches en science humaines du Canada

