

Thomas DRUGMAN – University of Mons, Belgium

John KANE, Christer GOBL – Trinity College Dublin

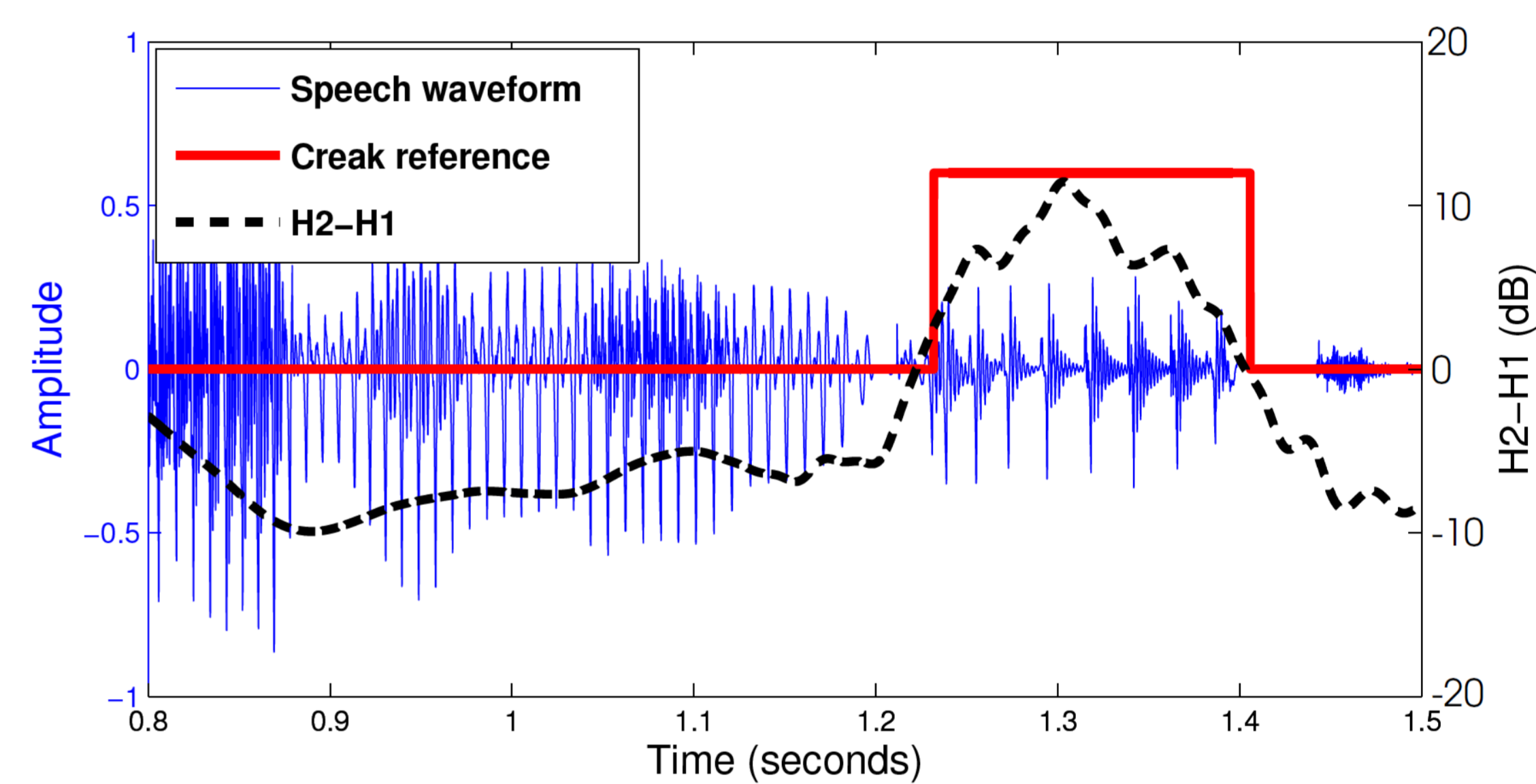
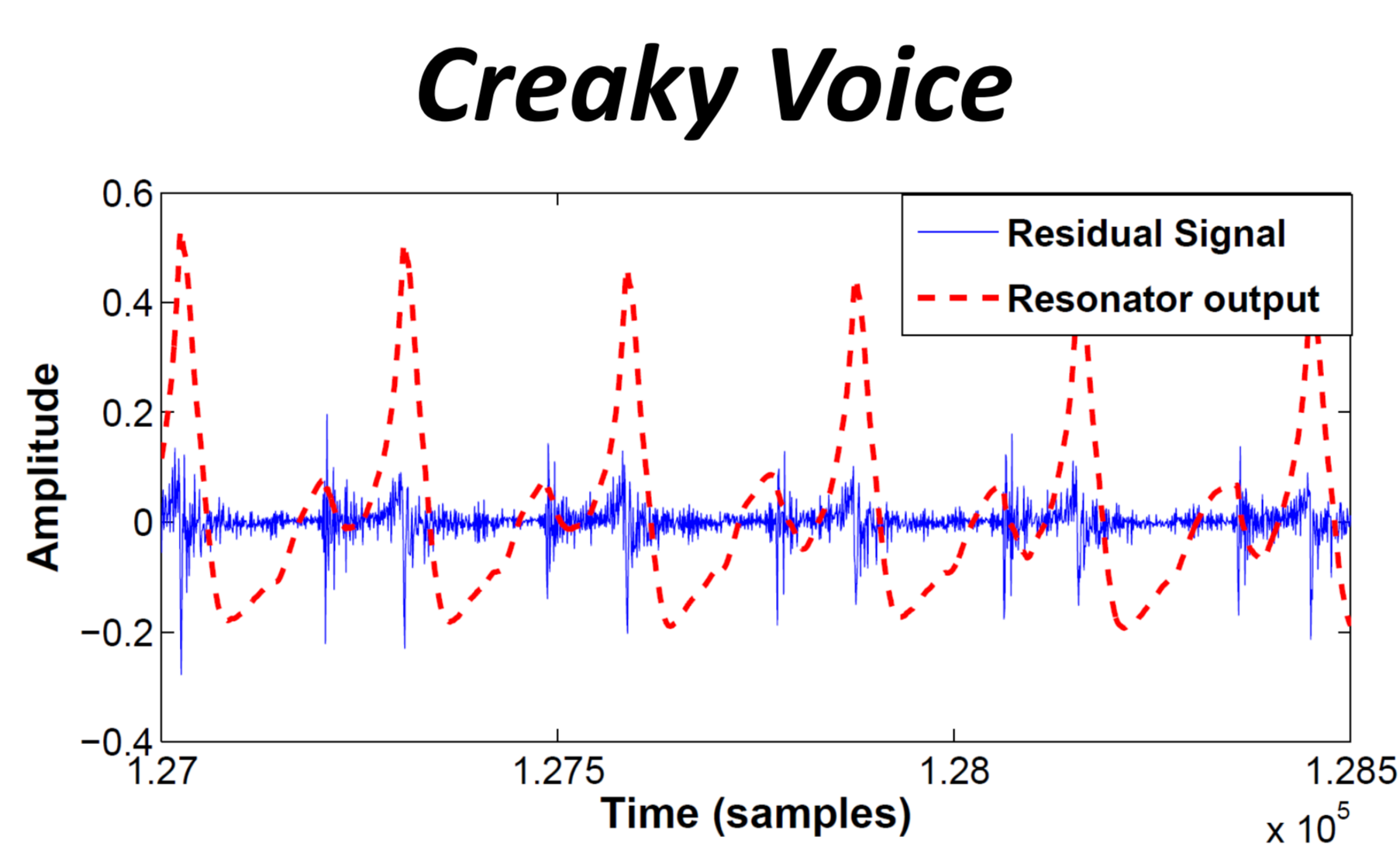
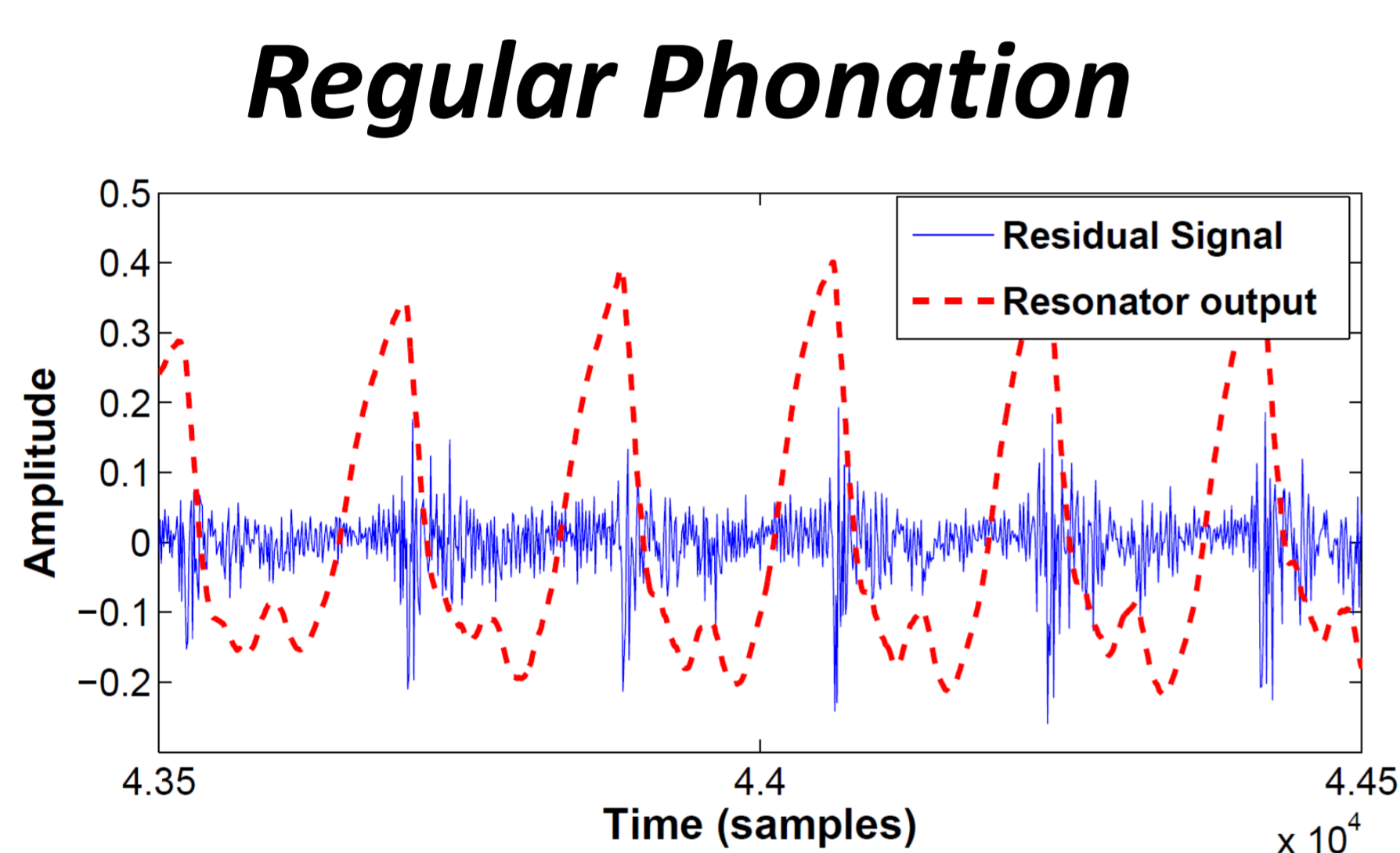
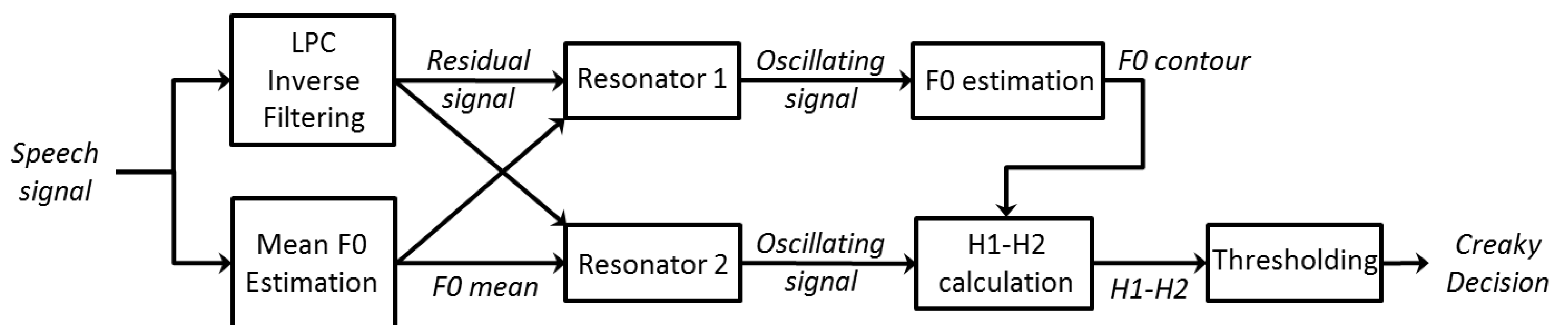
Problem Positioning

GOAL: Automatically detect the segments of creaky voice from the speech signal

CREAKY VOICE: - 'a rough quality with the sensation of additional impulses'
- also called glottal fry, vocal fry, or laryngealization

APPLICABILITY: speech synthesis, speaker identification, ASR, sociological studies, etc.

Resonator-based Creaky Voice Detection (RCVD)

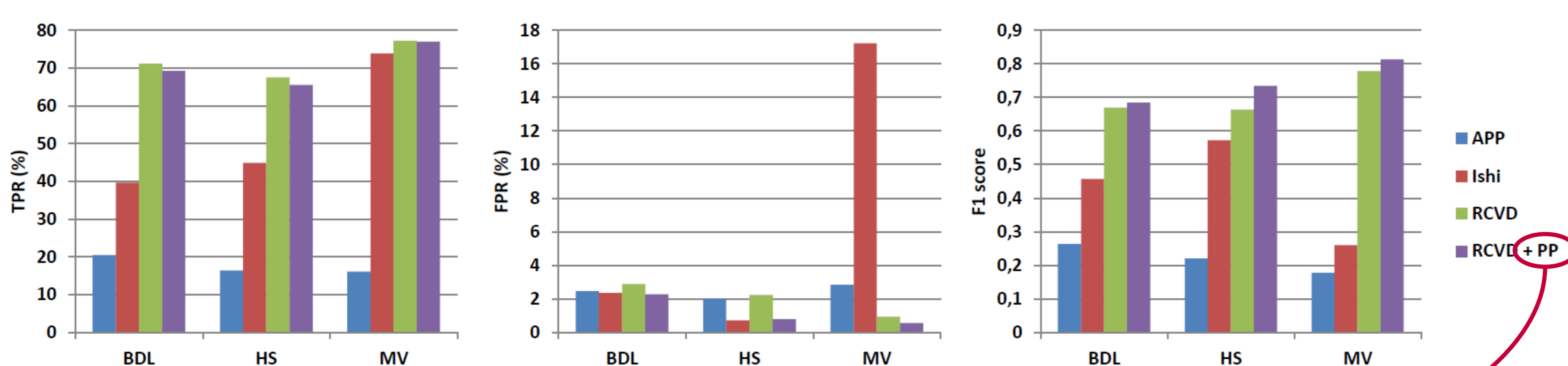


Experimental Evaluation

EXISTING METHODS: - Extension of the Aperiodicity, Periodicity and Pitch (APP) detector [1]
- Ishi's method for creaky voice detection [2]

DATABASES: - With creaky voice: ARCTIC BDL (US male), HS (Finnish female), MV (Finnish male)
- Without creaky production: ARCTIC AWB (Scottish male), SLT (US female)

METRICS: - At the frame level: TPR, FPR and F1 score (single measure based on precision and recall)
- At the event level: number of misses, false alarms and hits (events correctly detected)



Post-process to merge too close detections and remove too short ones

Database	Metric	APP	Ishi	RCVD	RCVD + PP
BDL	Misses	22	8	3	3
	FAs	215	93	37	27
	Hits	47	61	66	66
HS	Misses	59	12	12	15
	FAs	301	38	89	37
	Hits	46	93	93	90
MV	Misses	44	2	7	7
	FAs	353	920	49	14
	Hits	31	73	68	68
AWB	FAs	371	76	33	12
SLT	FAs	122	6	19	4

[1] Vishnubhotla, S., Espy-Wilson, C., "Automatic detection of irregular phonation in continuous speech", Proceedings of Interspeech, Pittsburgh, pp. 949-952, 2006.

[2] Ishi, C., Sakakibara, K., Ishiguro, H., and Hagita, N., "A method for automatic detection of vocal fry", IEEE Transactions on Audio, Speech and Language Processing, 16 (1), pp. 47-56, 2008