

Phonetic Analysis and Tempo Adjustments for Improved Speech Recognition in Dysarthric Speakers

Feifei Xiong · Jon Barker · Heidi Christensen

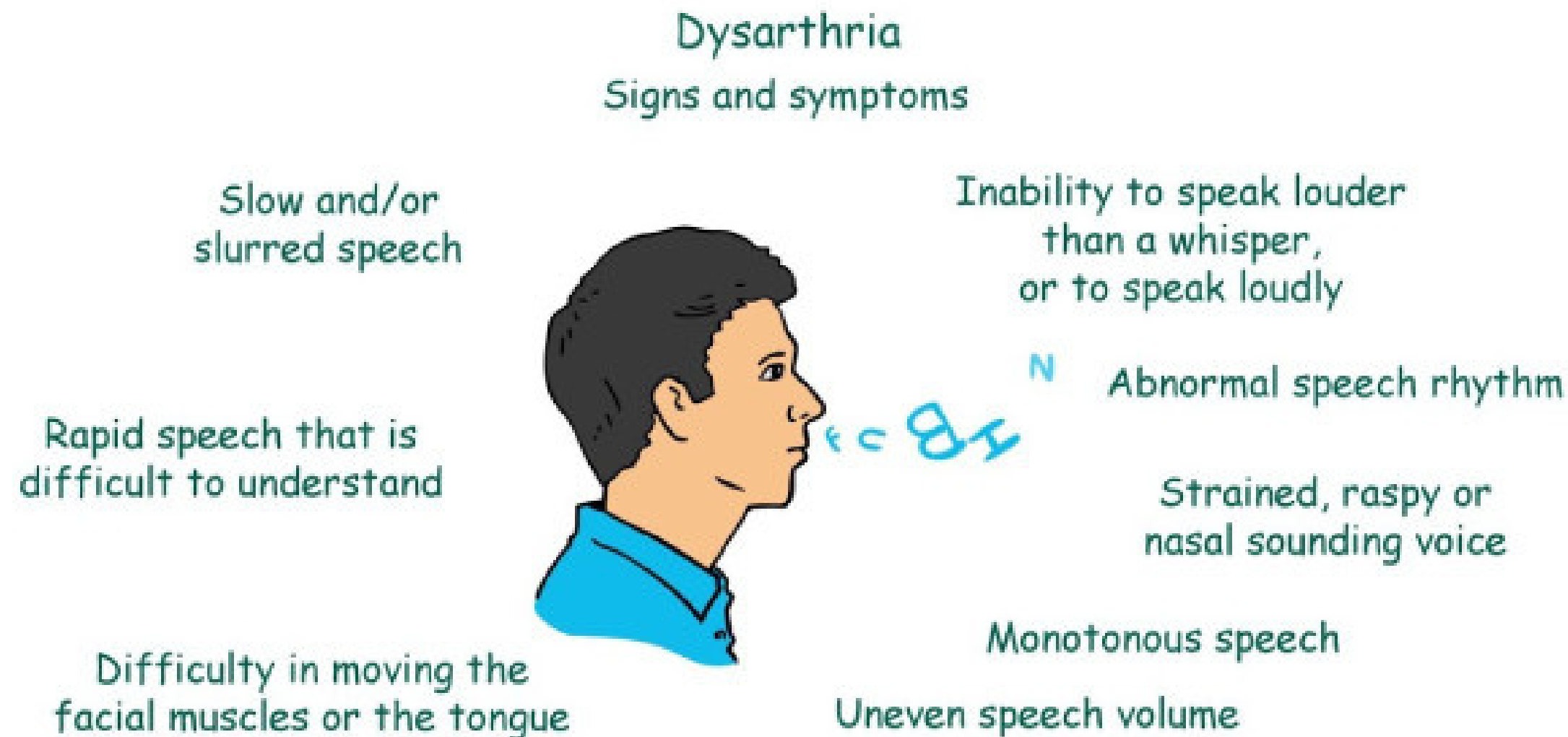
IEEE ICASSP-2019

Mid-Evaluation Presentation

Saketh Reddy Vemula - 2022114014

Viswanath Vuppala - 2022101084

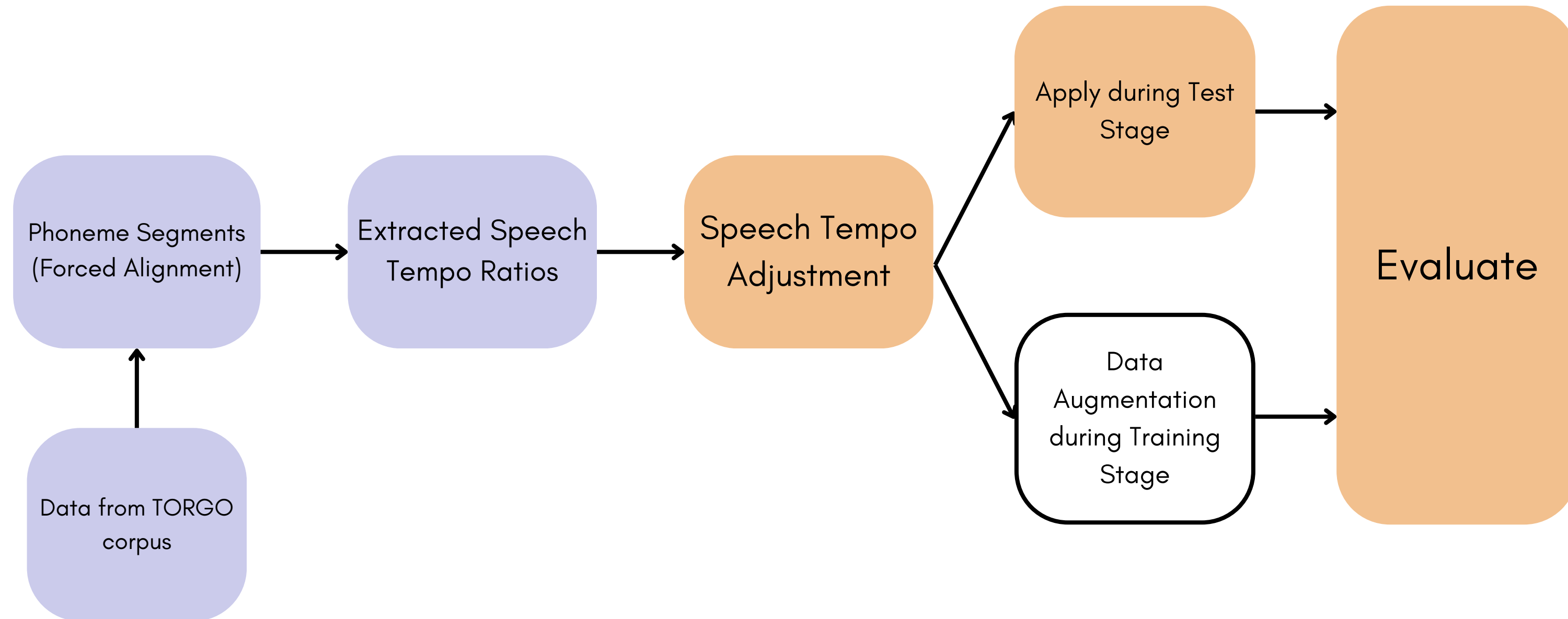
Introduction



- Increased respiration frequency
- Inadequate Pauses
- Breathy or hoarse voice
- Reduced speech
- Deviations in pitch and volume
- Mis-articulated Sounds

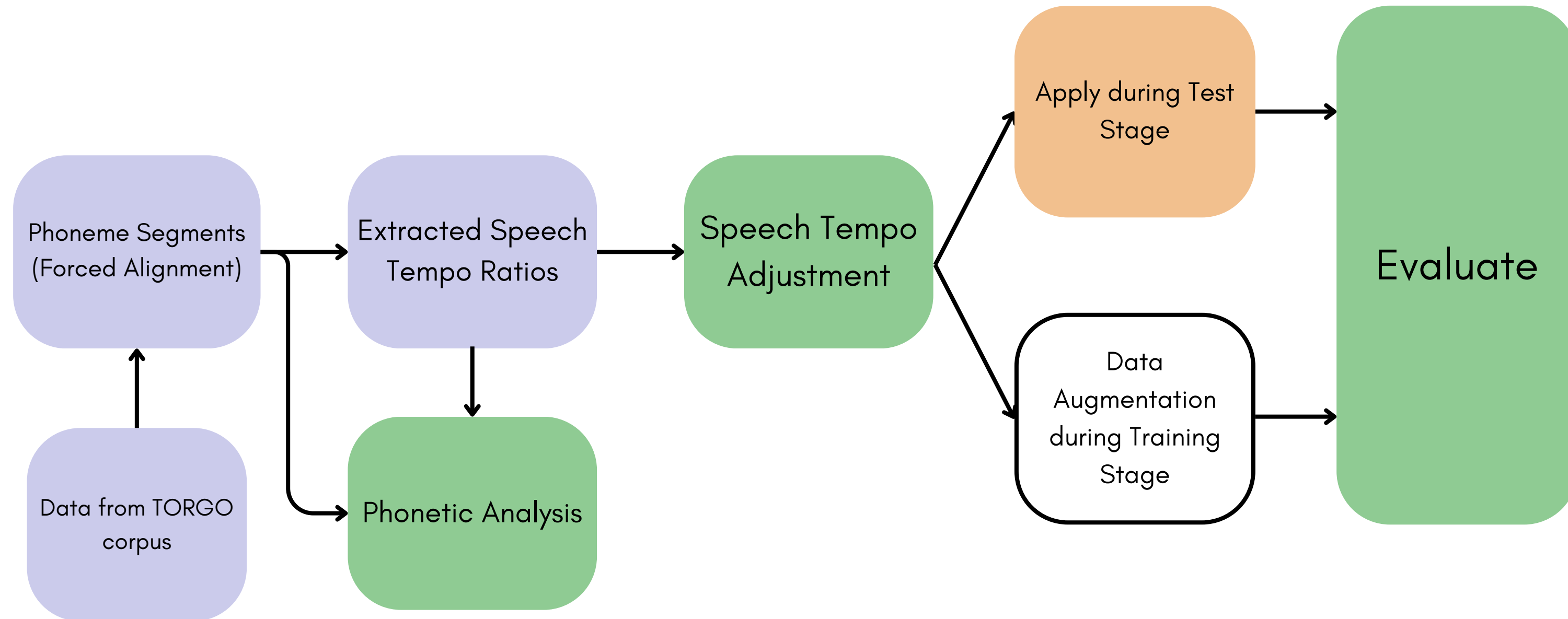
Earlier Conclusions

Presented Speech tempo adjustment for improved performance of ASR trained on typical speech

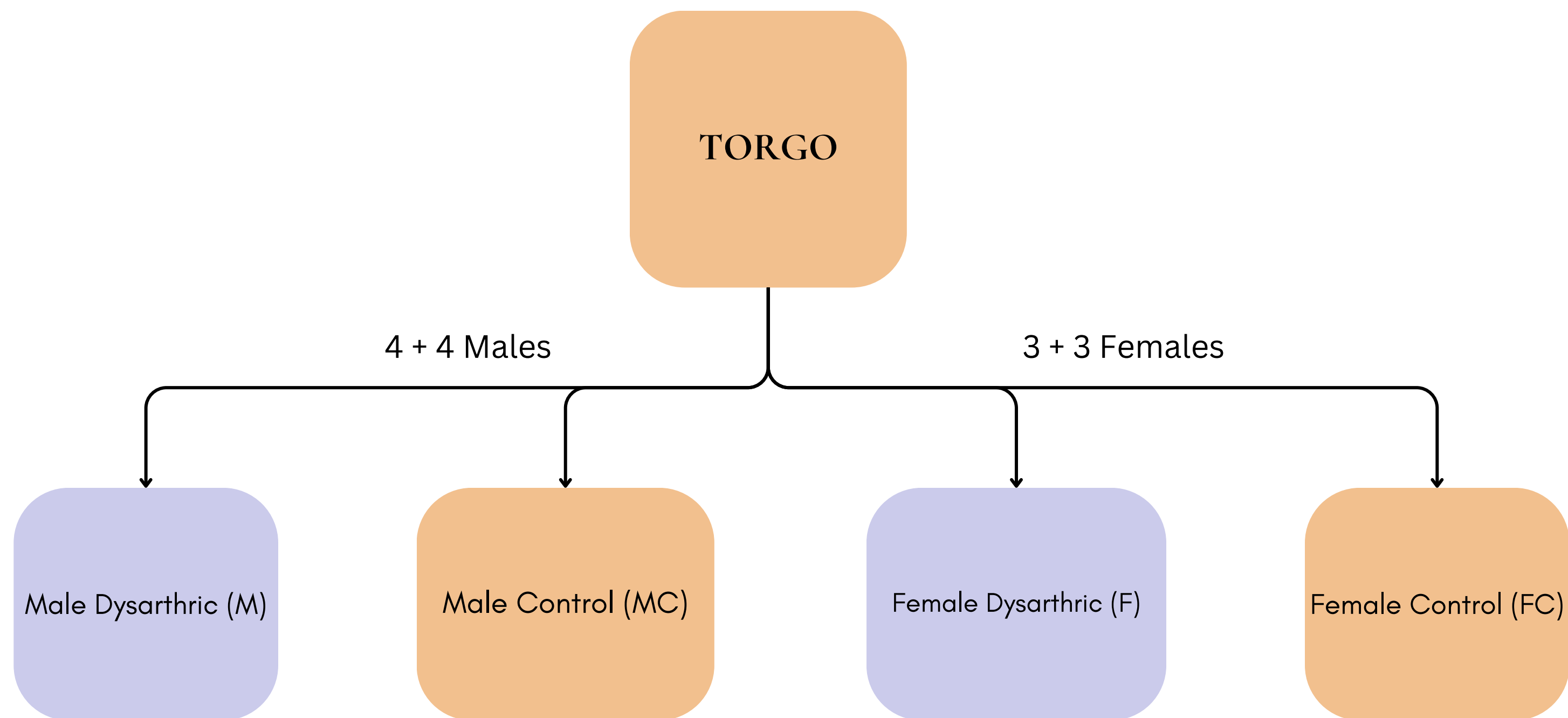


Objectives

- Analyze dysarthric speech at the phonetic level, specifically tempo.
- Implement phoneme-based tempo adjustments to assess their impact on automatic speech recognition (ASR) performance.
- Evaluate the effectiveness of phoneme-based tempo modifications in improving ASR accuracy for dysarthric speakers.



Dataset Overview



TORGO Database: Acoustic and articulatory speech from speakers with dysarthria

Phoneme Grouping

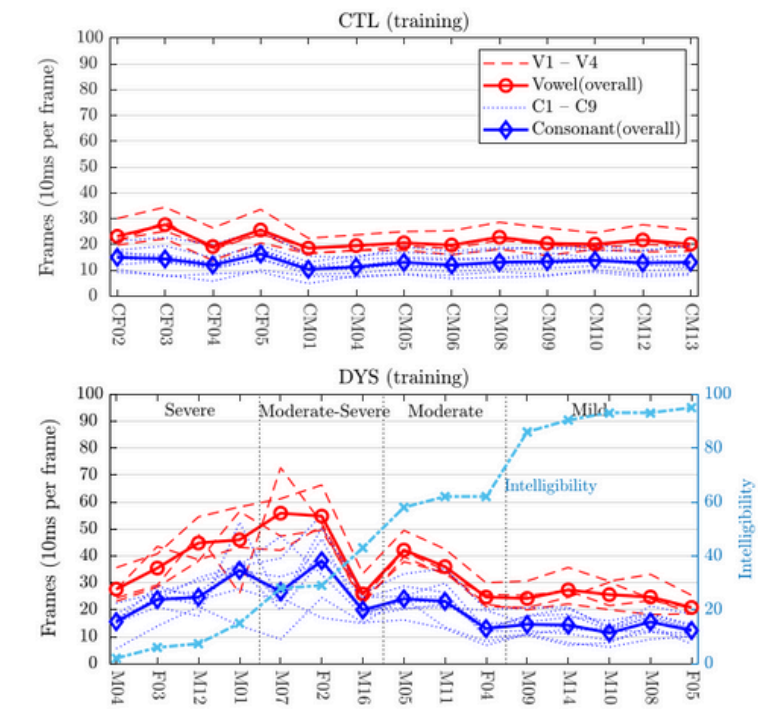
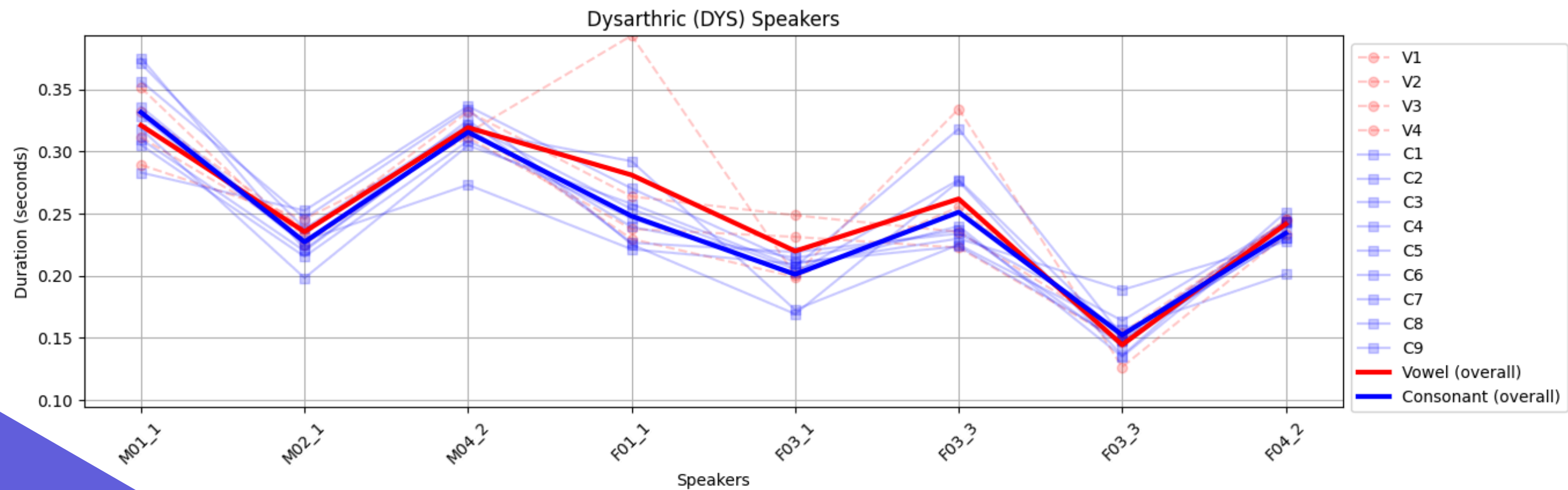
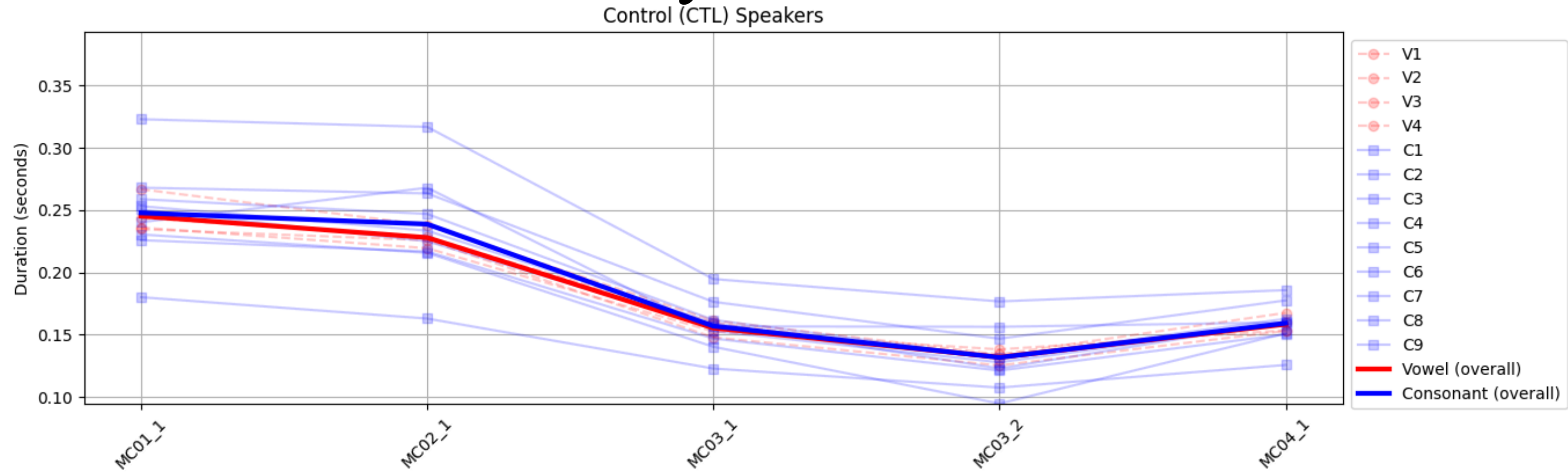
Vowels	V1: Short Vowels	AH AO AX EH IH UH
	V2: Medium Vowels	AE
	V3: Long Vowels	AA ER IY UW
	V4: Diphtongs	AW AY EY OW OY

Consonants	C1: Glides	L R W Y
	C2: Unvoiced Stops	K P T
	C3: Voiced Stops	B D G
	C4: Nasals	M N NG
	C5: Unvoiced Fricatives	F S SH TH
	C6: Voiced Fricatives	DH V Z ZH
	C7: Unvoiced Affricatives	CH
	C8: Voiced Affricatives	JH
	C9: Aspirates	HH
	Silence	bcl cl dcl gcl kcl nol nolo pcl tl
	Inconsistencies	sill lab2 a ahl ahn aor ee et ff h nh o rly

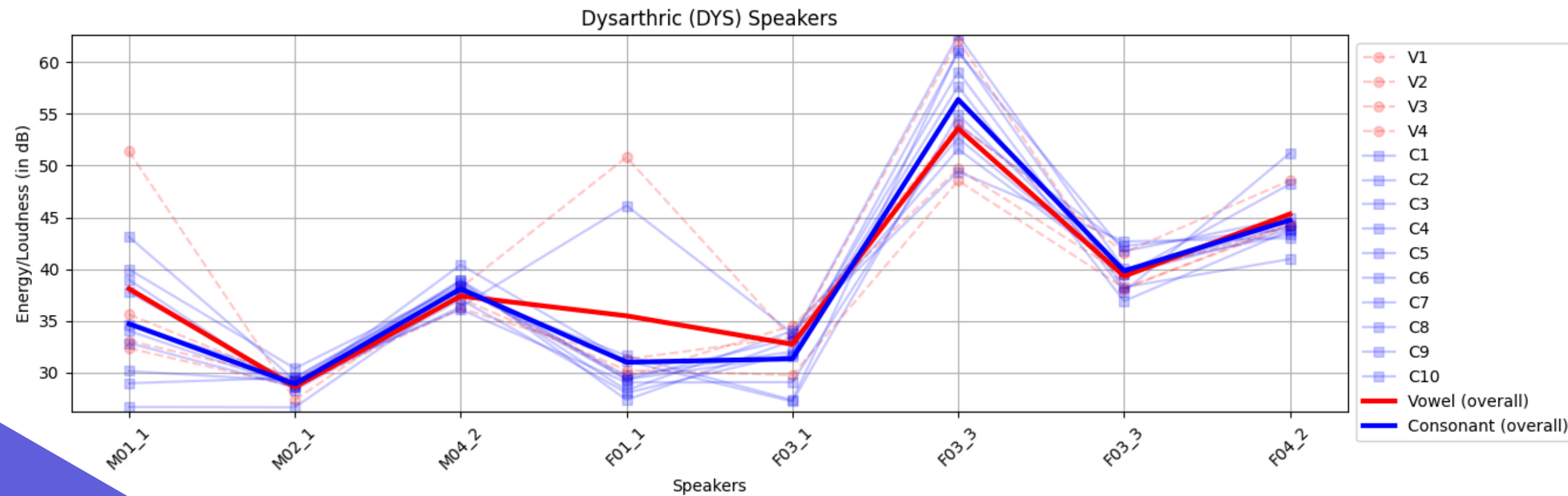
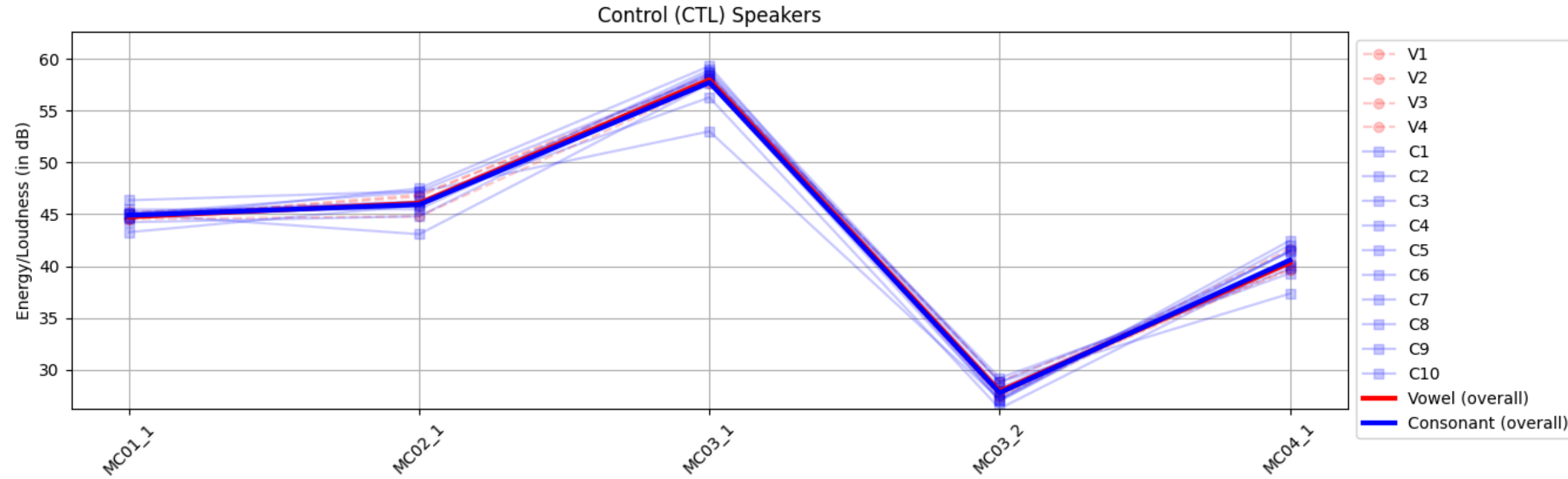
The image features a white background with decorative blue geometric shapes in the corners. In the top-right corner, there is a large blue triangle pointing downwards and to the left, with a smaller, darker blue triangle nested within its right side. In the bottom-left corner, there is a large blue triangle pointing upwards and to the right, with a smaller, darker blue triangle nested within its left side.

Google Sheet

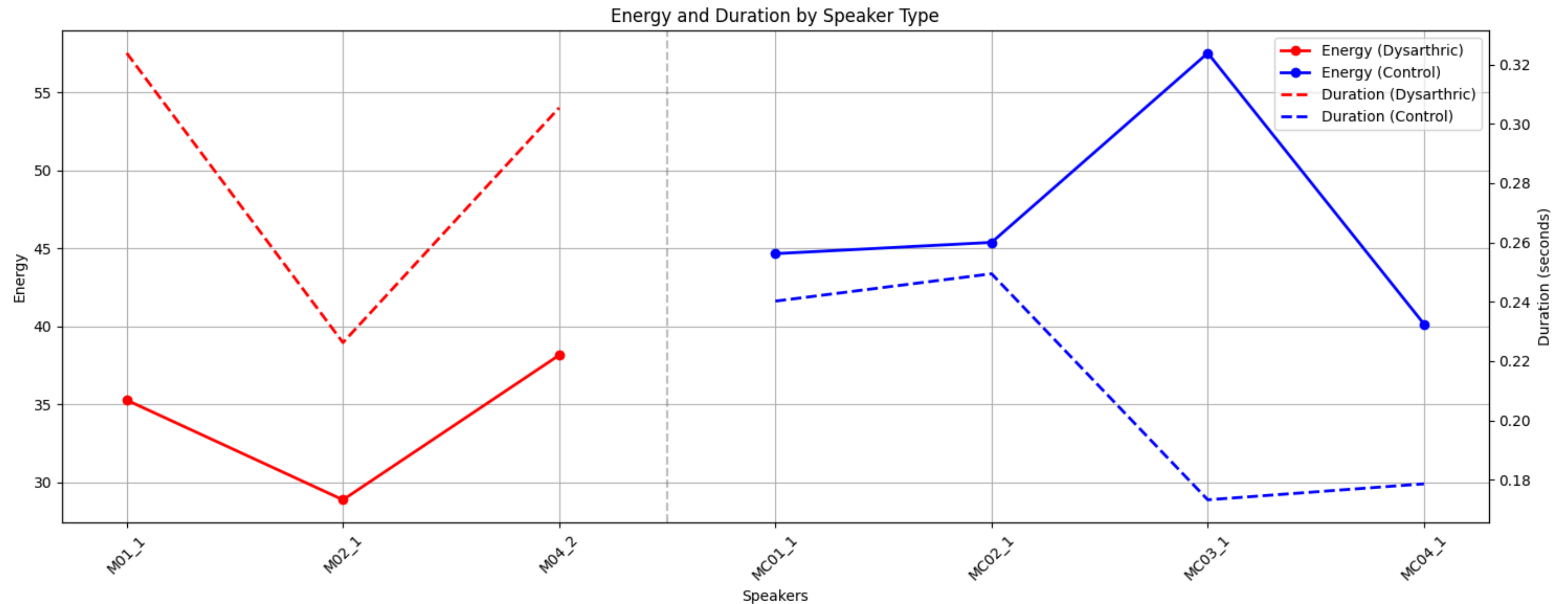
Phoneme duration Analysis



Phoneme Energy (Loudness) Analysis



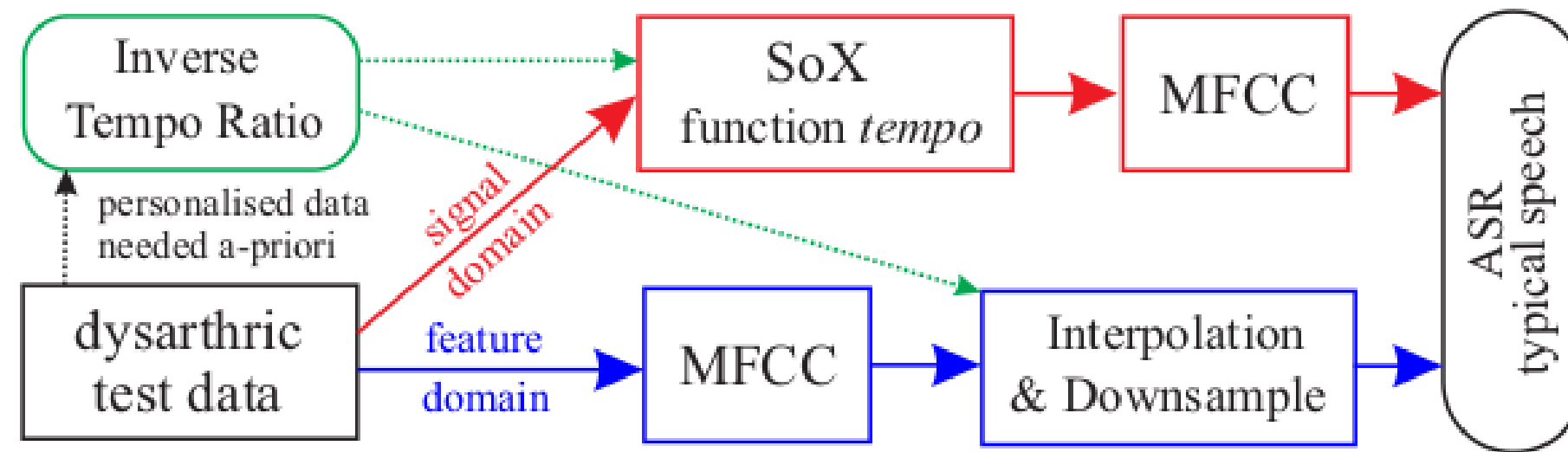
Overall Average Energy and duration



Future Work: Speech Tempo Adjustment for ASR

Test Stage

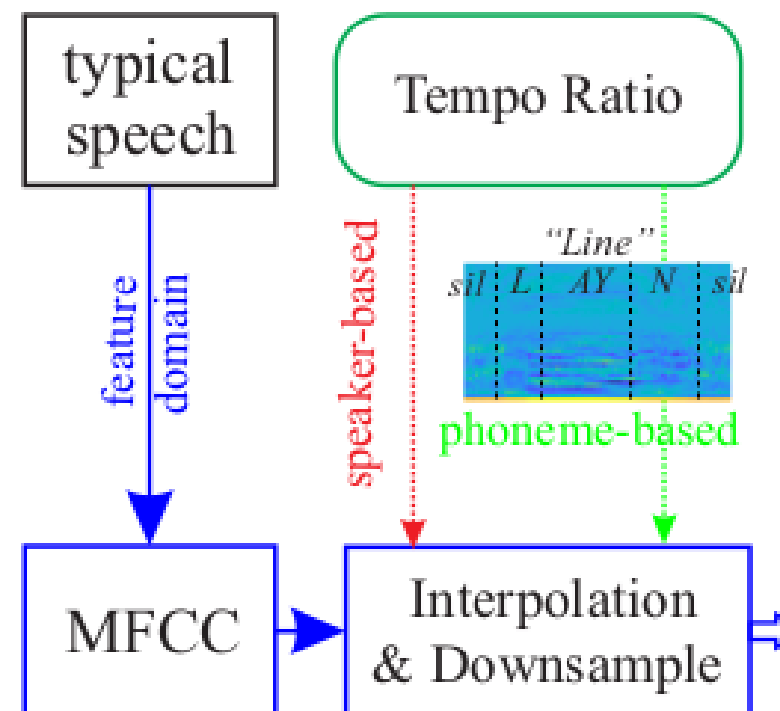
$$\overline{\mathcal{R}_{d \leftarrow c}}$$



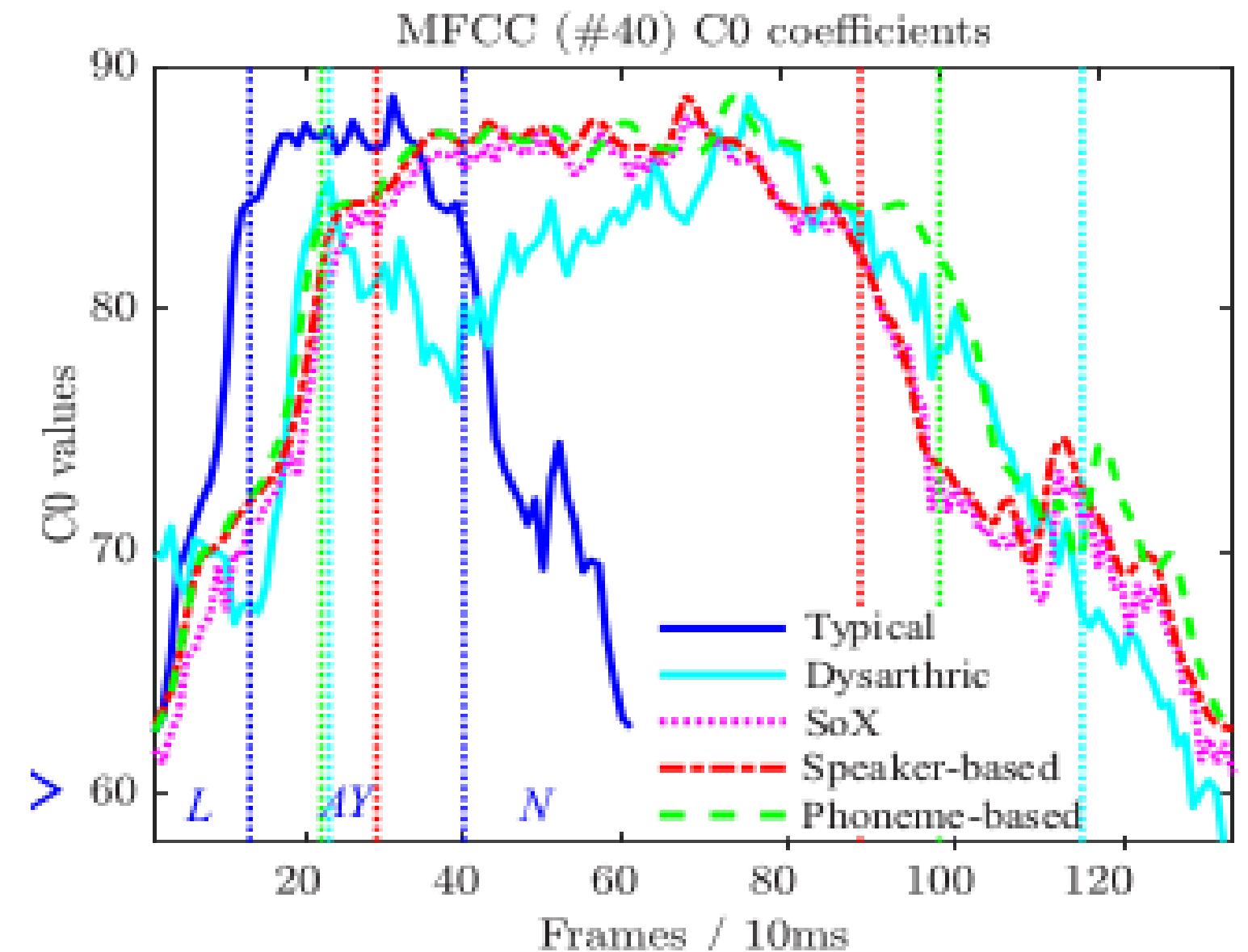
Note:

Phoneme-based tempo ratios are not possible in Test Stage due to lack of alignment knowledge in dysarthric test data

Training Stage



$$\mathcal{R}_{d \leftarrow c}(p) = \frac{T_d(p)}{T_c(p)}$$





Thank You