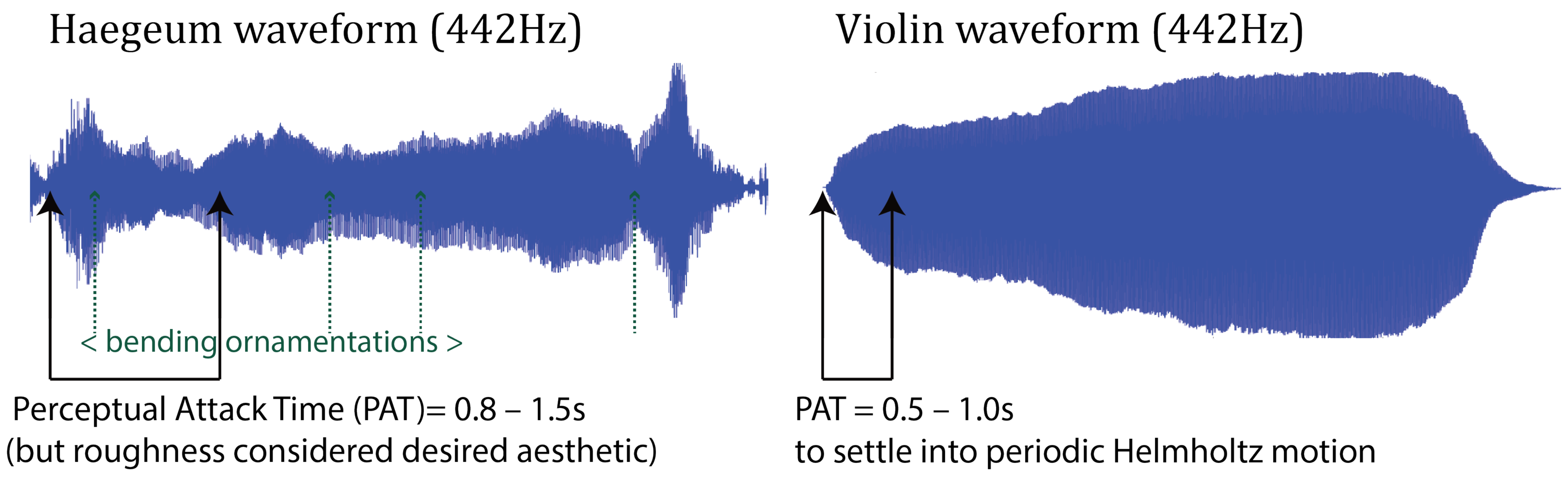
Computational Onset Detection for the Haegeum:
Addressing Cultural and Acoustic Complexity

links to paper + contact info

Michaella Moon, Dale Carnegie, Jim Murphy
Victoria University of Wellington
School of Engineering and Computer Science
New Zealand

The haegeum, a two-stringed bowed instrument central to Korean traditional music (Gugak), presents unique challenges for computational onset detection due to its flexible intonation, ornamented phrasing, and prolonged tonal stabilization. We conducted a study comparing Western musicians with those trained in Gugak, revealing cultural differences in how perceptual onsets are identified. These insights informed the development of a tolerance window and an evaluation dataset specifically tailored to the haegeum's timbre and the aesthetics of Gugak, to create a pitch detection function prone to fluctuating pitches. This culturally grounded approach not only improves technical accuracy but also supports performance analysis and the development of pedagogical tools that are responsive to innate musical practices.

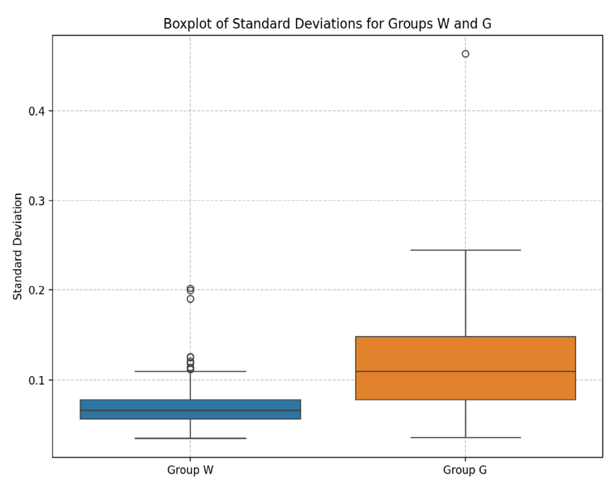


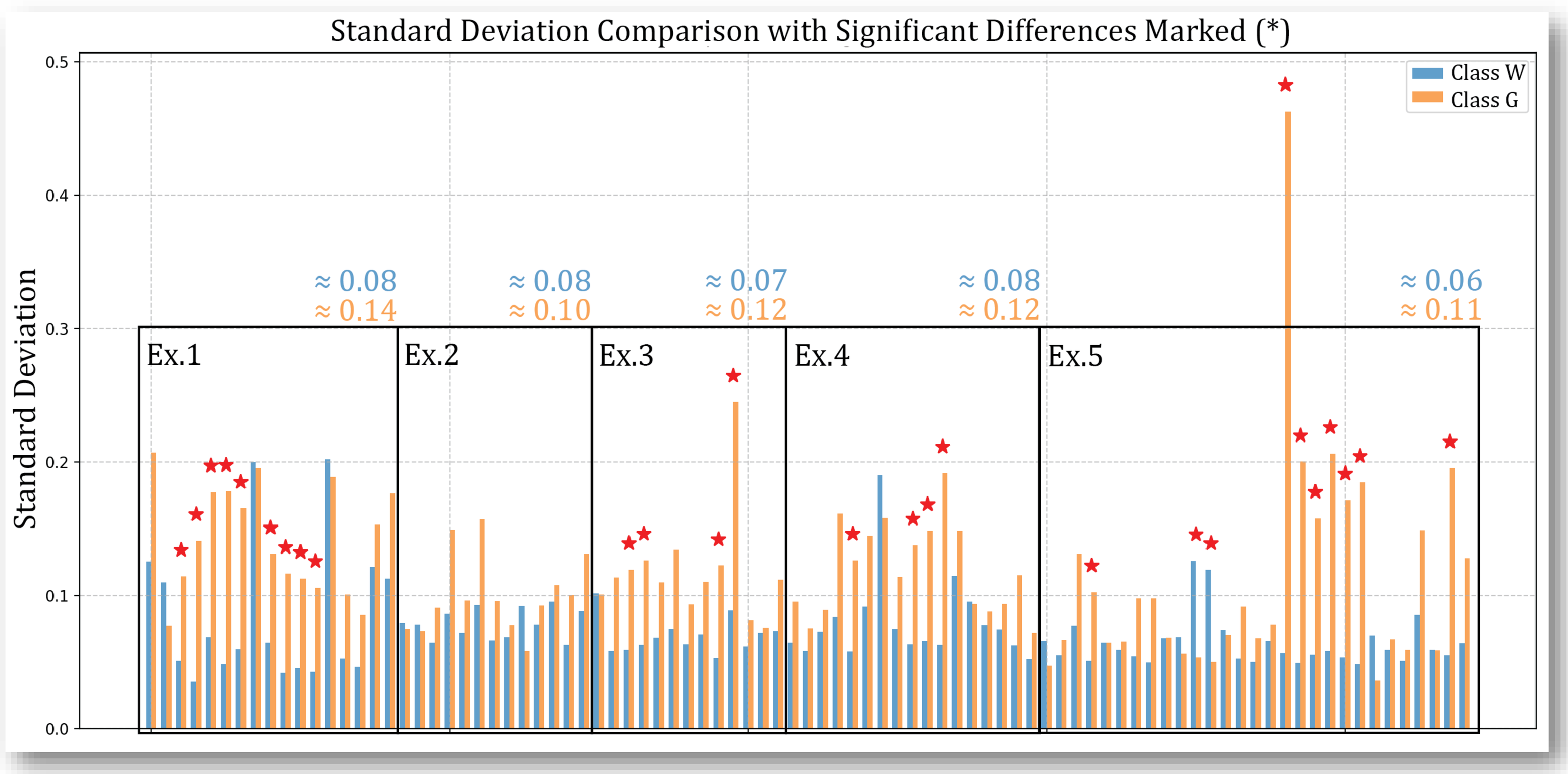
Research Objectives

1. Compare the perceptual onset annotation behaviours of Western-trained musicians (N = 10) with those of Gugak-trained musicians (N = 10).
2. Establish a Ground Truth dataset and tolerance windows specifically designed for the haegeum's gradual tone stabilisation, to evaluate the following system.
3. Develop a culturally informed onset detection framework for the haegeum, aimed at performing specialised pitch detection analysis.

Annotation Analysis

The graphs compare onset annotations from Western and Gugak-trained musicians across five excerpts, including chromatic (1) and pentatonic scales (2, 3), and two beginner pieces with varying rhythmic and pitch complexities (4, 5).

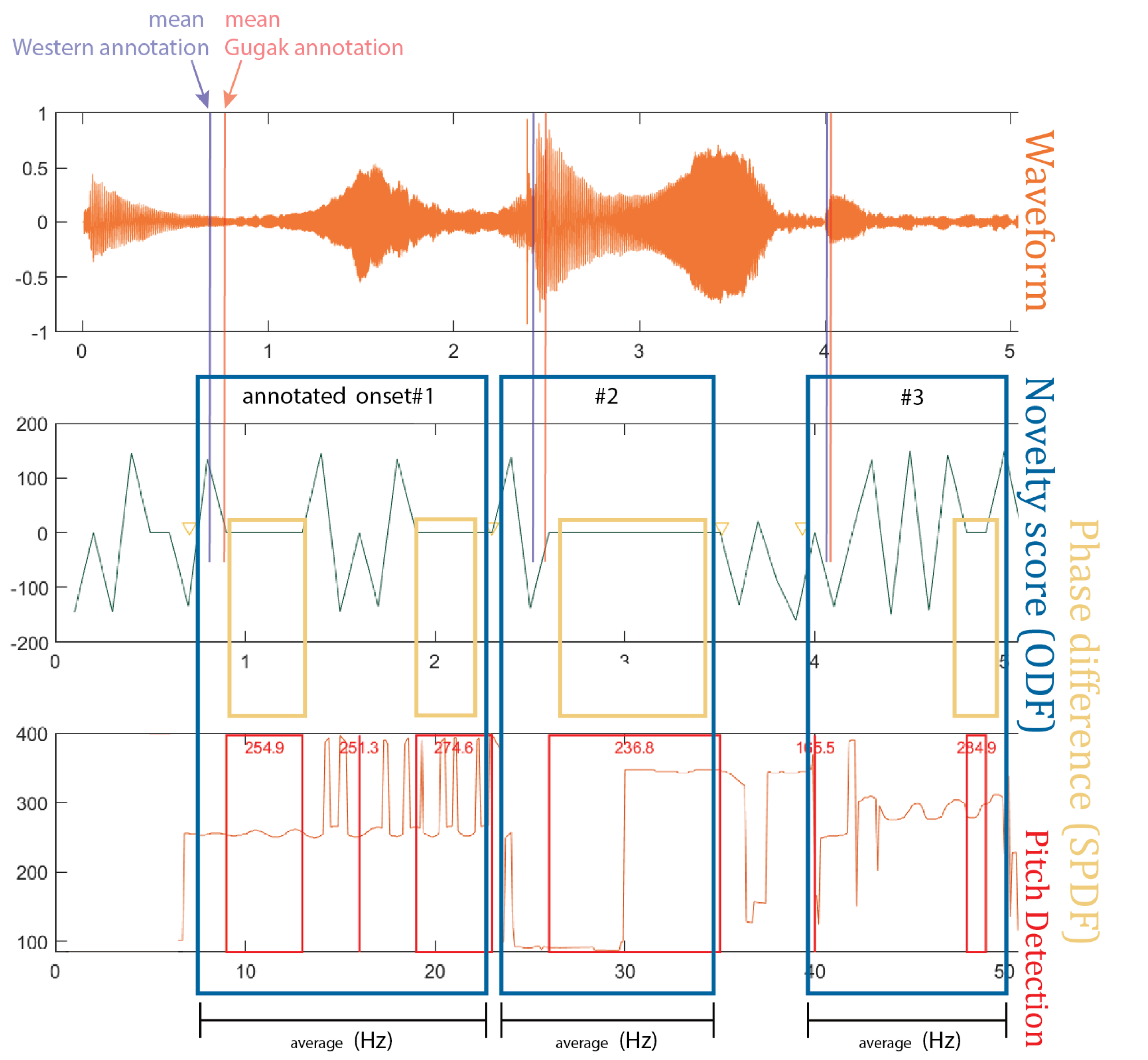




Pooled Standard Deviation (variance)	**Significant disparity in standard deviations between the groups:** F-value = 3.63
Paired T-test (mean differences between the two dependent classes)	Western mean ≈ 0.07 Gugak mean ≈ 0.11
Negative T-statistic (Group 1= W / Group 2 = G)	**70/89 onsets, W annotated earlier than G:** Negative T-statistic = -6.97
The observed later onset labelling by Gugak annotators—paired with higher variability — suggests that performance traditions, stylistic flexibility, and listening strategies likely shape timing perception.	

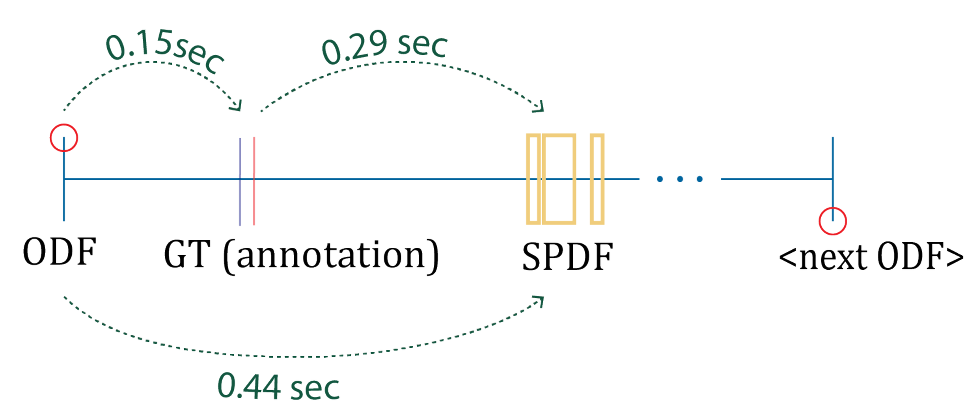
Multi-windowing Results

ODF = Global window of one intended note	**Bow friction's white noise spikes @ bow direction change:** DFT → power spectrum → cross-compare → extract local maxima
SPDF = Localised window for pitch detection function	• First-order phase difference = rate of change in each freq bin • Second-order phase difference = across consecutive 0.1s frames **ID prominent spectral peaks**



Comparison of Ground truth (GT), ODF, and the multiple SPDF windows within each ODF frame shows the timing from motion onset to PAT and highlights that several short SPDF windows can occur even without intentional ornamentation or pitch bending.

Method*/Excerpt >	1	2	3	4	5
ODF to GT	-0.13s	-0.21s	-0.21s	-0.06s	0.13
Human perceptual onset (GT) detected later than bow onset (ODF)					
SPDF to GT	0.17s	0.31s	0.31s	0.42s	0.25s
ODF to SPDF	0.29s	0.51s	0.53s	0.48s	0.34s



Conclusion

ODF identifies the global windowing parameters of an intended note, while SPDF refines this by locating the most stable pitch regions and filtering out unintentional fluctuations arising from instrument design challenges. This study employed a corpus with minimal ornamentation to establish best-case haegeum onset instances. Although the first SPDF window for haegeum aligns closely with the previously established Western PAT, the occurrence of multiple SPDF windows within a single ODF frame further demonstrates the applicability of this methodology for beginner Gugak musical analysis. Detailed pitch detection algorithms (normalised correlation function) and their evaluation lie beyond the scope of this paper and will be presented in a forthcoming study.