



UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II



PhD student: Luigi Maria Giordano Orsini

PhD thesis title:

Linguistical Orientend Automatic Speech
Recognition System
for Museum Applications

Tutor: Francesco Cutugno

Cycle: XXXIX

Year:2024/2025

Candidate's information

- MSc degree: Statistical Science for the Decisions Università degli studi di Napoli Federico II
- DIETI Research group/laboratory: DIETI Research group/laboratory: Interdepartmental laboratory Urban Eco
- PhD: 1/11/2023 31/10/2026
- Scholarship type: PNRR DM 118/2023 Mis. 4.1: Dottorati Pubblica Amministrazione
- Periods in company: Museo e Real bosco di capodimonte, Napoli, Italy

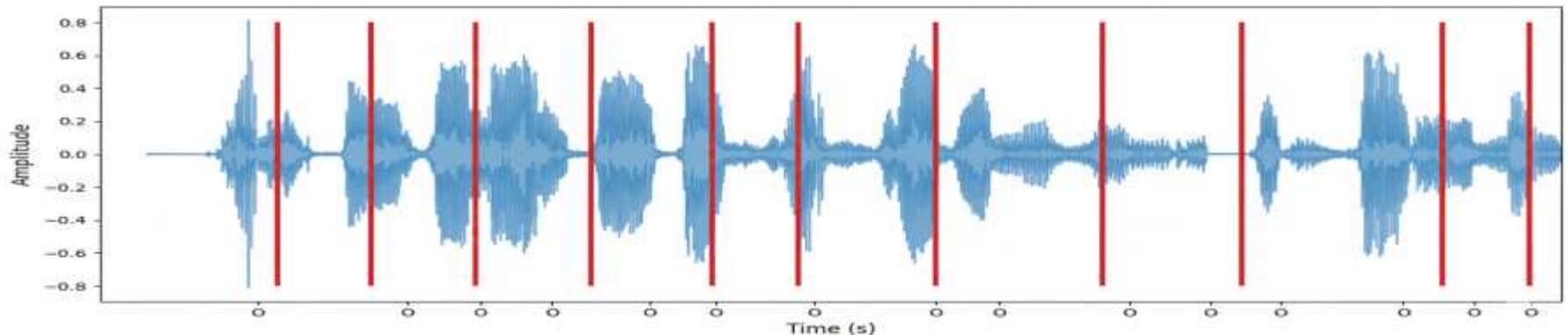
Summary of study activities

	Courses	Seminars	Research	Tutorship
First Year	15	1.2	23	0
Second Year	30.6	5.4	38	1.6
Expected	10 – 20 (30-70 total)	5 – 10 (10-30 total)	10 – 45 (80-140 total)	0 – 1.6 (0-4.8 total)

- Courses attended
 - Conversazioni su Etica e Intelligenza Artificiale
 - Matrix Analysis for Signal Processing with MATLAB Example
 - Speech Processing
 - La Scienza moderna e il problema della disciplina giuridica dell'IA
 - Big Data Architecture and Analytics
 - Text Mining
 - AI Code Generation: Foundations, Evaluation, and Security

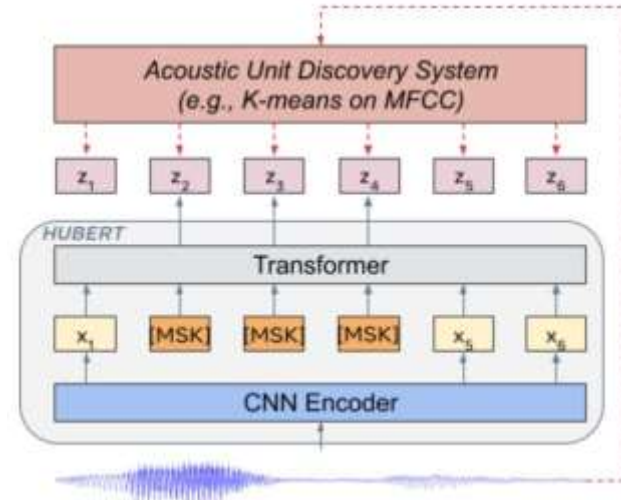
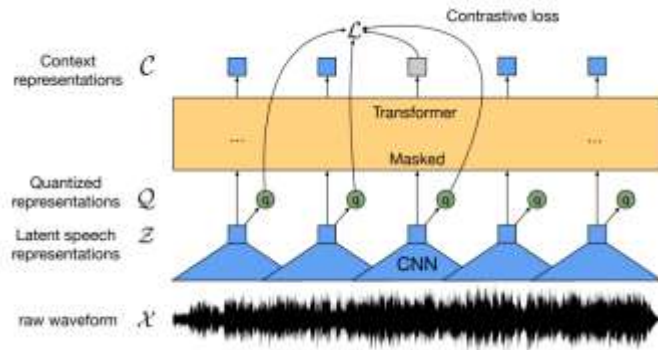
Research area

The long-term goal is to explore novel approaches to Automatic Speech Recognition (ASR) based on syllable-like units, motivated by studies suggesting the salience of syllables in human cognitive development.



- Variable length
- Time dependent
- Complex informative content

Research area



Current State-of-the-Art (SOTA) models often circumvent the problem of variable-length inputs.

- Enforcing a fixed-length analysis window
- Transforming the audio into alternative representations
- Applying a neural approach without a deep, foundational analysis of the data's intrinsic structure

Research activity

- 3 Phases Framework



Calibration:

- Defining the Space: Benchmarking the tools that measure similarity.



Grouping:

- From Map to Meaning: Which algorithm best translates distance into discovery?



Representation:

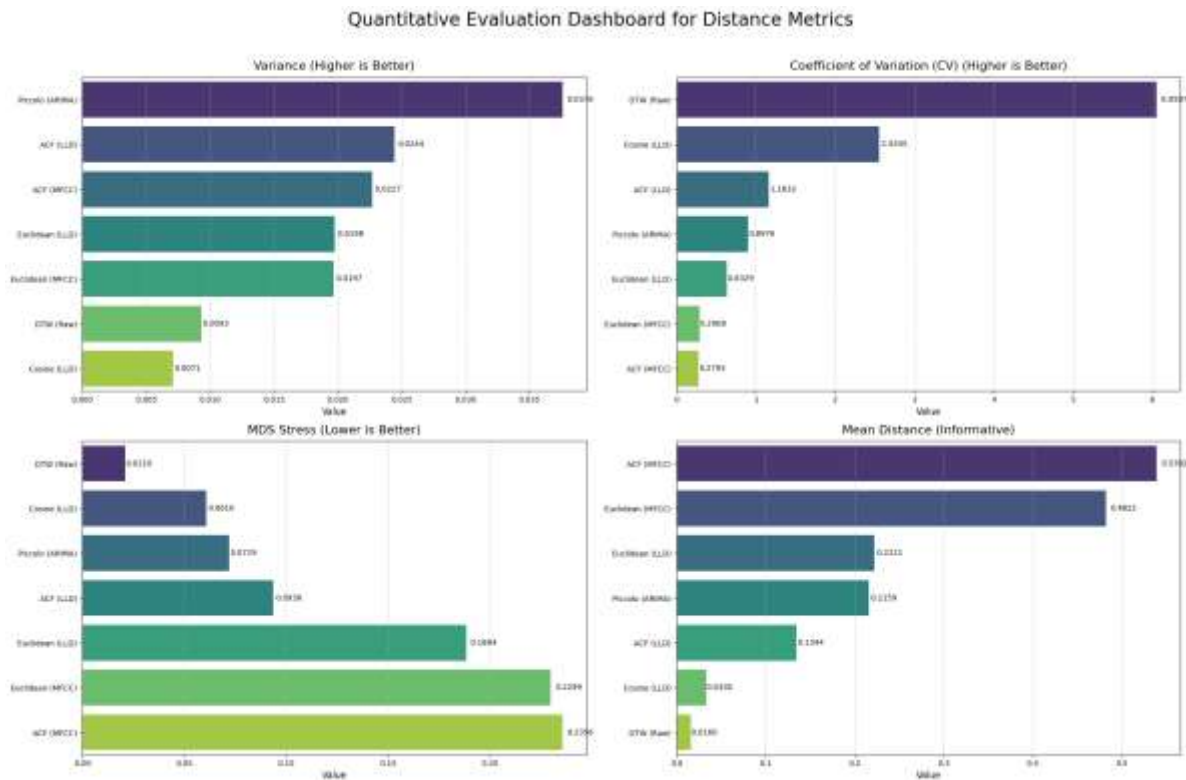
- Designing a Better Universe: The quest for the most mappable representation.

Research activity

- **Conducted a Large-Scale Comparative Study:**
 - Processed and prepared a large population of over 5,000 syllabic segments from an annotated speech dataset.
 - Computed seven distinct distance matrices on a representative sample, evaluating metrics ranging from standard (Euclidean, Cosine) to time-aware (DTW) and model-based (ARIMA/Piccolo).
- **Established a Quantitative Evaluation Methodology:** Implemented a robust framework to assess each metric's quality via its statistical properties (Coefficient of Variation) and its performance in **Multidimensional Scaling (MDS)**, evaluated using Kruskal's Stress-1.
- **Developed a New Analysis Framework:** Designed and implemented a complete software pipeline in Python to systematically analyze and compare time series distance metrics from data extraction to final evaluation.
- **Designed a Multi-Experiment Validation Plan:** Architected a quantitative validation phase based on clustering and three distinct sampling strategies (weighted, balanced phonotactic, balanced phonetic-articulatory) to objectively test the hypotheses of metric specialization.

Research results

Identified Evidence of "Metric Specialization": The choice of a distance metric is not a trivial decision but is **highly task-dependent**. This finding challenges the "one-size-fits-all" approach often seen in standard vector database implementations.



Future Work



- *Quantitatively Validate Metric Specialization via Clustering*
 - The key outcome will be a definitive, multi-faceted performance analysis, confirming which metrics are superior for tasks that are structural, content-based, or reflective of natural distributions.
- *Extend the Framework to Advanced Representation Learning*
 - This will quantitatively answer the research question: "Can a learned representation offer superior performance for the defined clustering tasks?"

Thank you for your time
and attention