

Sino-Tibetan Endangered Languages Protection Automated Tone Transcription, Clustering and Representation

Yi Yang, Yiming Wang, and Jiahong Yuan

University of Science and Technology of China

Tips

What Is Tone?

- Syllable-level Pitch Variations

Why Is Tone Important?

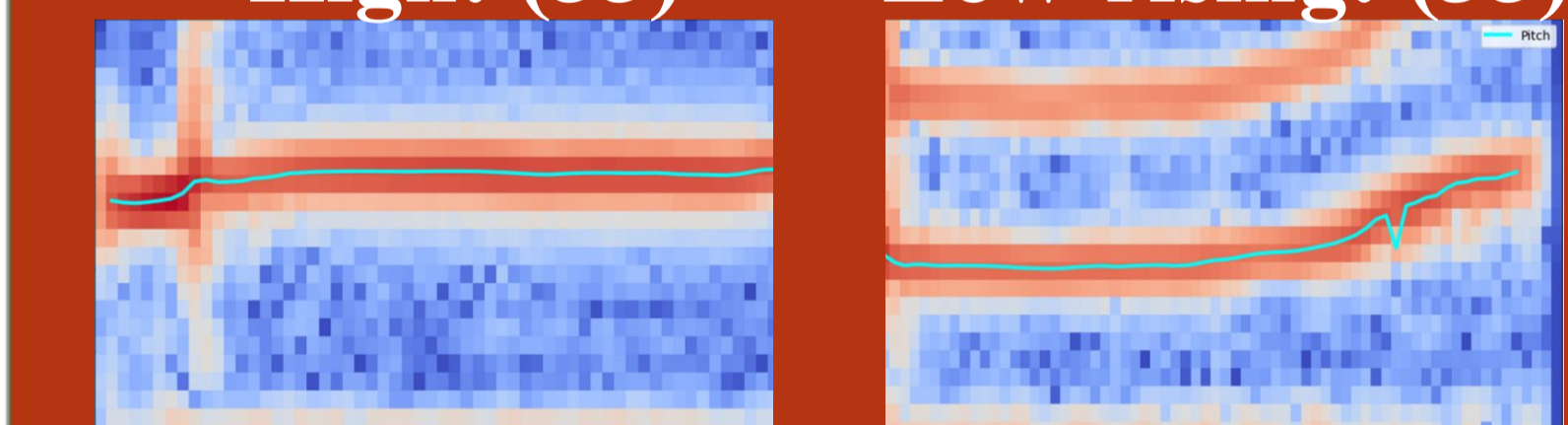
- Most Sino-Tibetan languages are tonal languages.
- Different tones represent different meanings.
- Documenting tones is crucial for phonetic fieldwork.

Five-scale Marking System

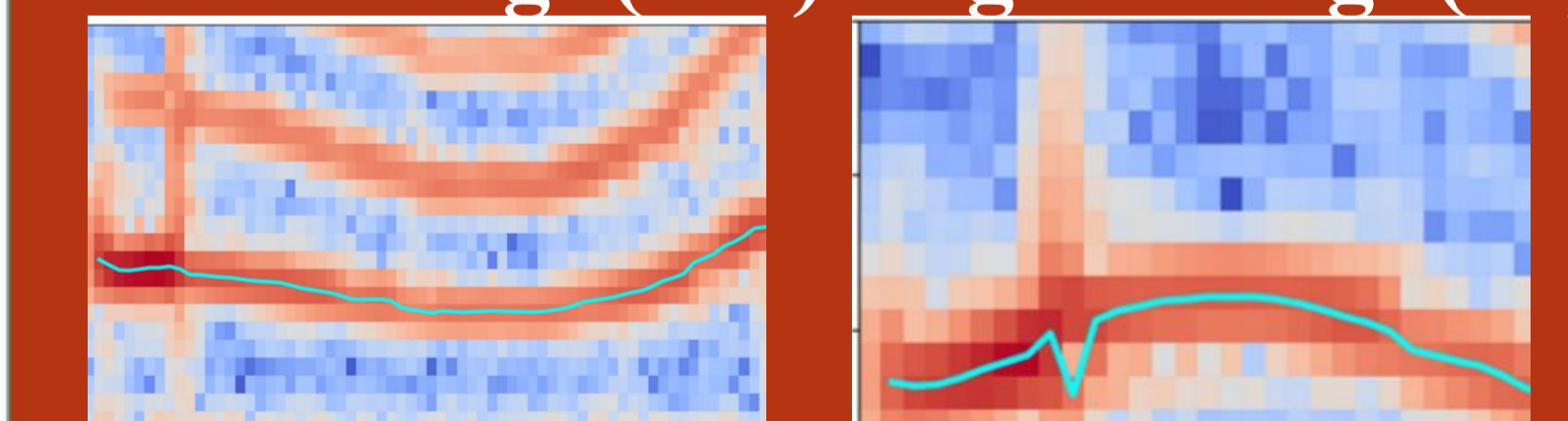
- Transcribe Tones into strings.
- Split pitch into 5 relative levels.

Four Basic Mandarin Tones

High: (55) Low-rising: (35)



Low-falling: (214) High-falling: (53)



Hear the Difference!



Background

- The Sino-Tibetan language family includes over **400** languages, connecting **1.4 billion** people culturally.
- **Most** Sino-Tibetan languages are **endangered**.

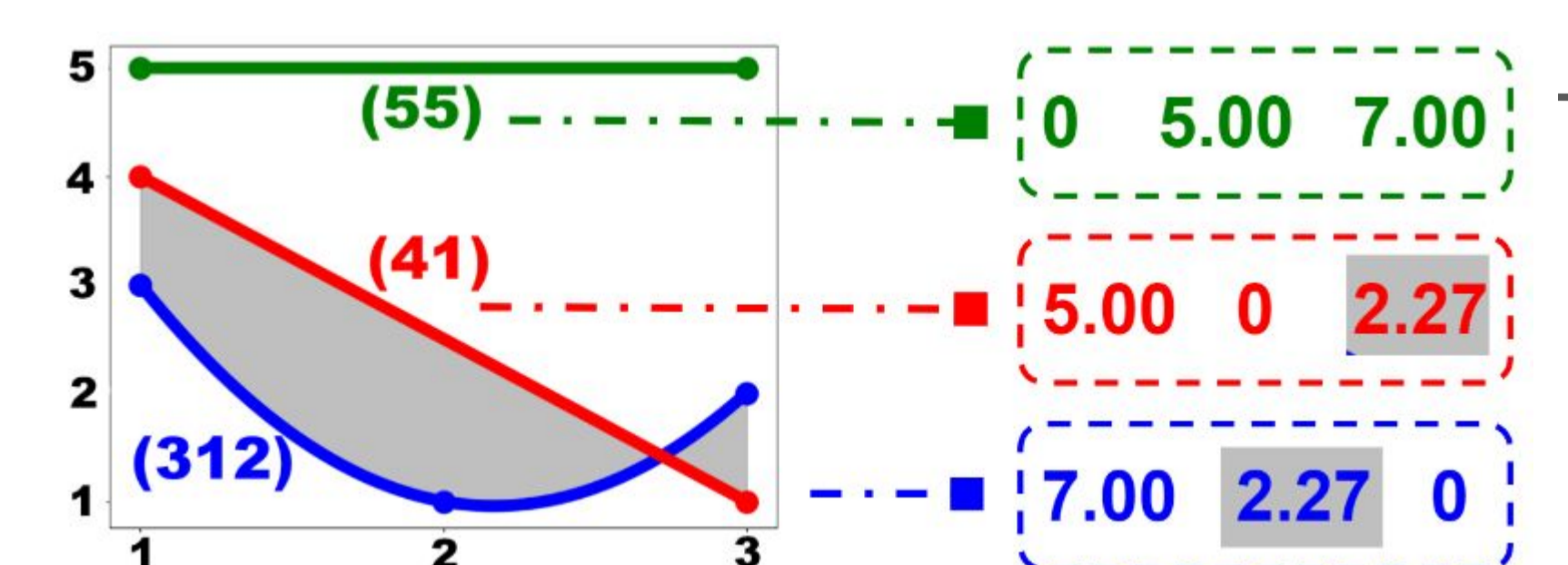
Two Obstacles:

- **Documenting:** Phonetic fieldwork relies on **manual** transcription of tones, which is costly, slow, and subjective.
- **Analysis:** Although fieldwork gathers extensive data, building **comparable tonal features** and conducting quantitative analysis have long remained unexplored areas.

Automatic fieldwork tools and tonal representations are urgently needed to protect and uncover our linguistic diversity!

Algorithms and Designs

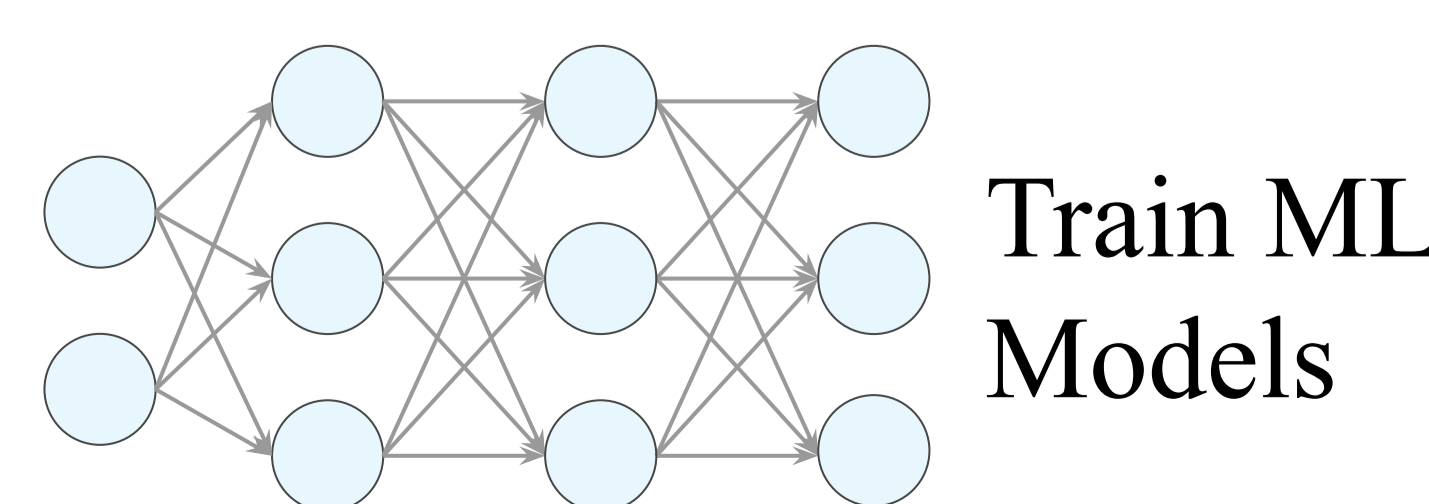
Map Pitches into Simulated Curves



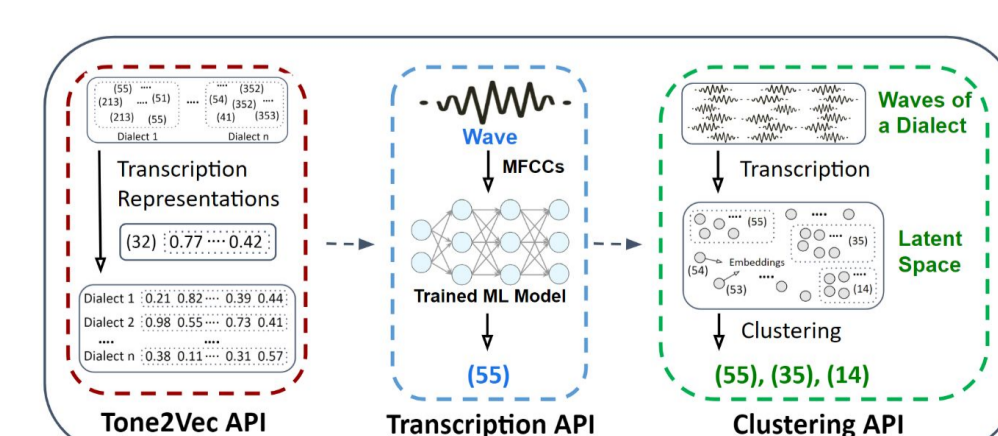
Tone to Vectors

Subsequent Analysis

$$\mathcal{L}_{pitch}(\mathcal{X}, \mathcal{Y}) = - \sum_{j=1}^N \hat{D}(\mathcal{M}(x_j), y_j)$$



Train ML Models



Proposed Package: ToneLab

Transcription API

Clustering API

Tone2Vec API

Func1: Automatic Tone Transcription

- Accept speech from any dialect as input and outputs a five-scale transcription

Model	Method	Accuracy (%)	Variance
	F0	10.07	0.2165
ResNet	Tone2Vec	55.99	0.1222
VGG	Tone2Vec	<u>56.08</u>	0.1052
DenseNet	Tone2Vec	61.01	<u>0.1083</u>

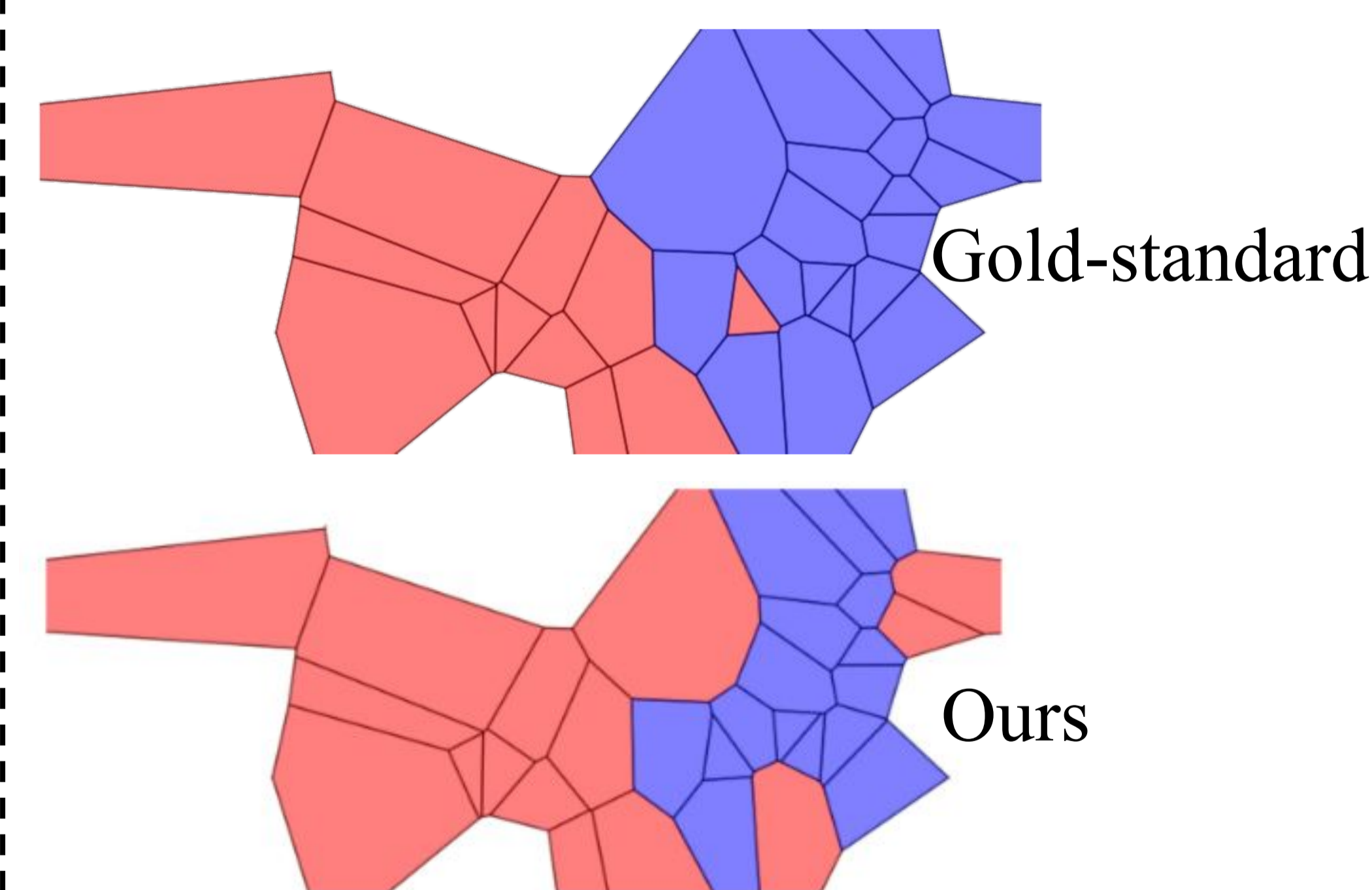
Func2: Automatic Tone Clustering

- Group a series of signals into N categories

SPK	Type	Tone 1	Tone 2	Tone 3	Tone 4
OF	Lab.	(213)	(24)	(41)	(53)
	Pred.	(313)	(45)	(51)	(42)
YF	Lab.	(212)	(24)	(51)	(55)
	Pred.	(213)	(34)	(52)	(44)
OM	Lab.	(213)	(24)	(41)	(51)
	Pred.	(212)	(34)	(31)	(32)

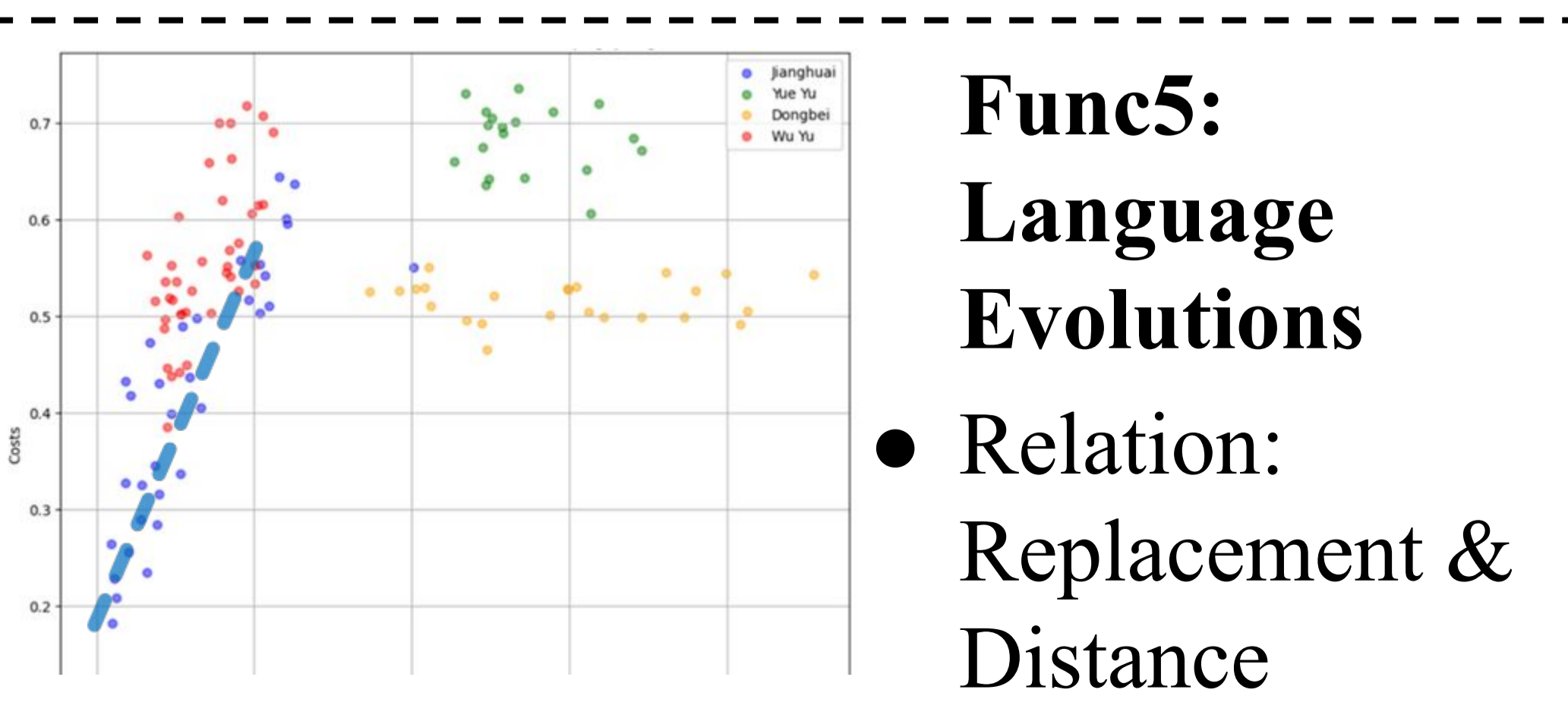
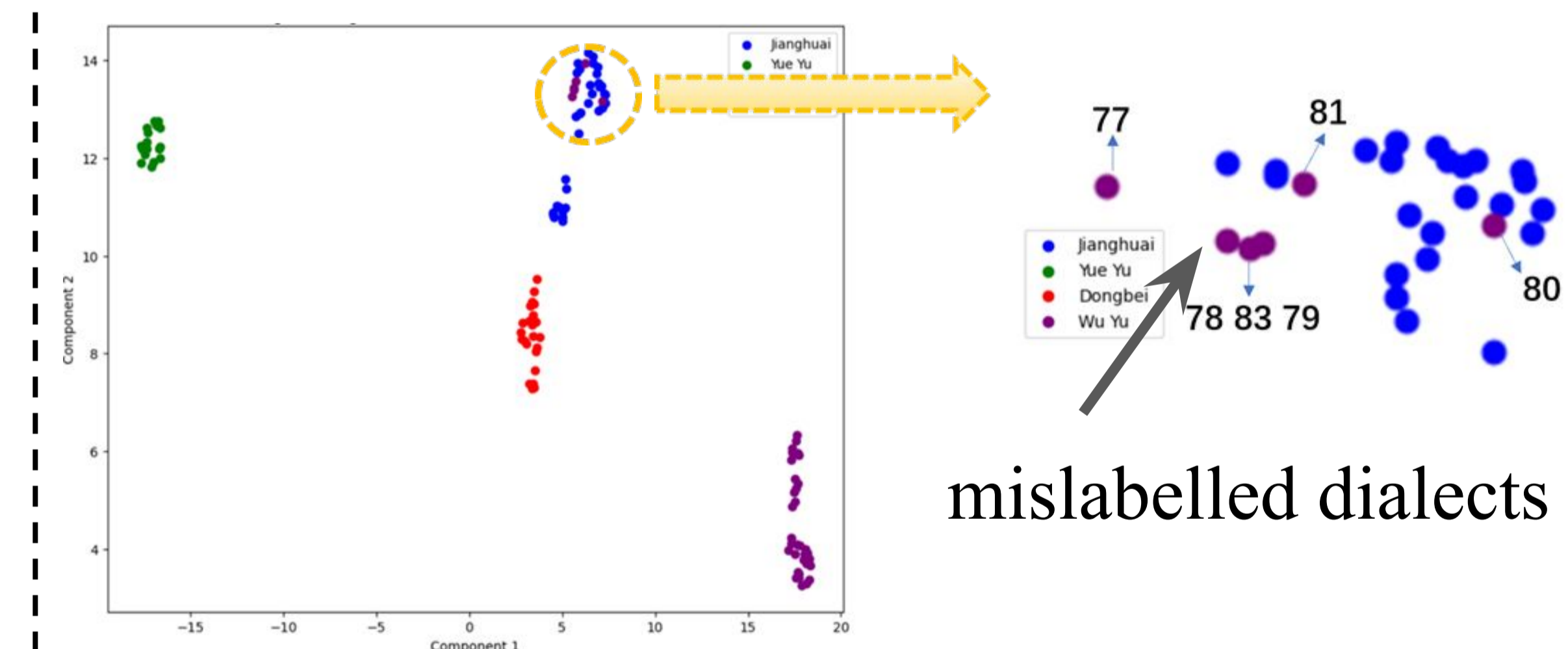
Func3: Dialect Clustering and Variance

- Use **comparative** tonal representations to categorize and analyze tonal variance across dialects.



Func4: Data Correction

- Correct data labels through anomaly detection in features.



Func5: Language Evolutions

- Relation: Replacement & Distance

Know more about our work!

Homepage

Full Paper

