# Tone Development and Change

**Chenzi Xu**

University of Oxford

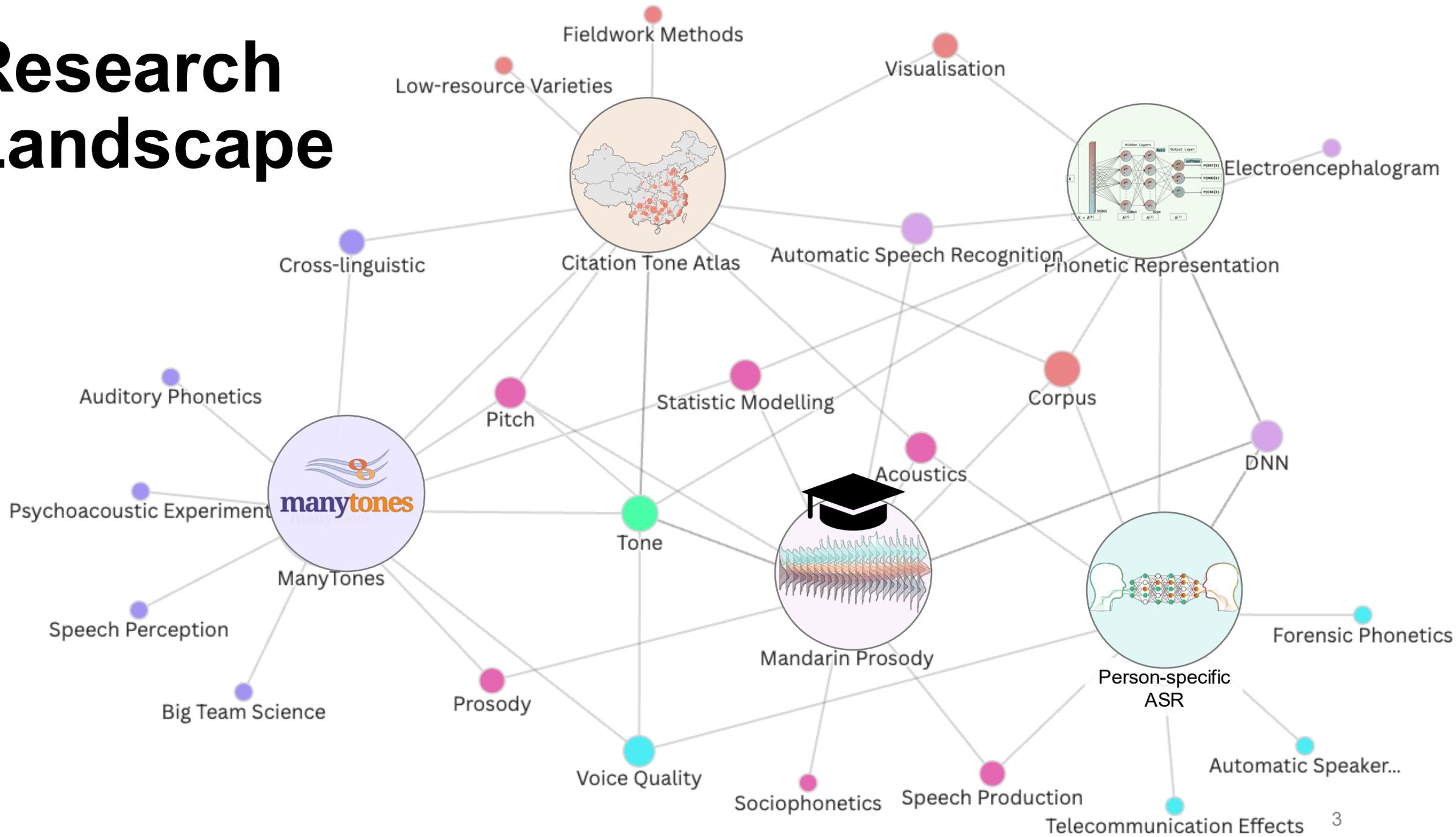🏠 https://chenzixu.rbind.io      ✉ chenzi.xu@ling-phil.ox.ac.uk

# About me

**Dr Chenzi Xu**

Leverhulme Trust Early Career Fellow
Faculty of Linguistics, Philology and Phonetics
University of Oxford

Jan 2026: (Tenure-track) Assistant Professor
Nanya Technological University, Singapore

# Research Landscape

**RoadMap**

**RoadMap**

# What is phonological representation? Why do we need it?

**Abstract** model that

- captures **systematic** patterns in the sound structure of languages
- enables **general** and **predictive** understanding of the patterns
- *aligns well with implicit knowledge or **mental** representation (cognitively plausible)
- *supports computation of some kind (generation, parsing)

# What is meant by tonal representation?

- Model continuous pitch movement in **discrete** categories
- Understand how pitch functions systematically
- Reveals underlying **regularities** across languages and dialects
- Allows us to explain and predict pitch contrasts and variation
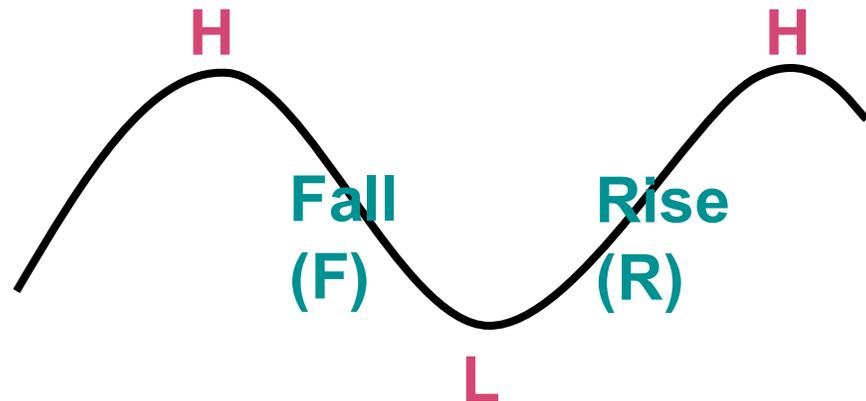
# What is meant by tonal representation?

The Autosegmental-Metrical Prosodic Tree (from Grice, 2022)



continuation rise

final lowering

Tonal features are represented on a separate tier from phonemes

Intonational Phrase

intermediate phrase (Minor intonational phrase)

Phonological phrase

Phonological word (prosodic word)

Foot

Syllable

Segmental structure

Tune
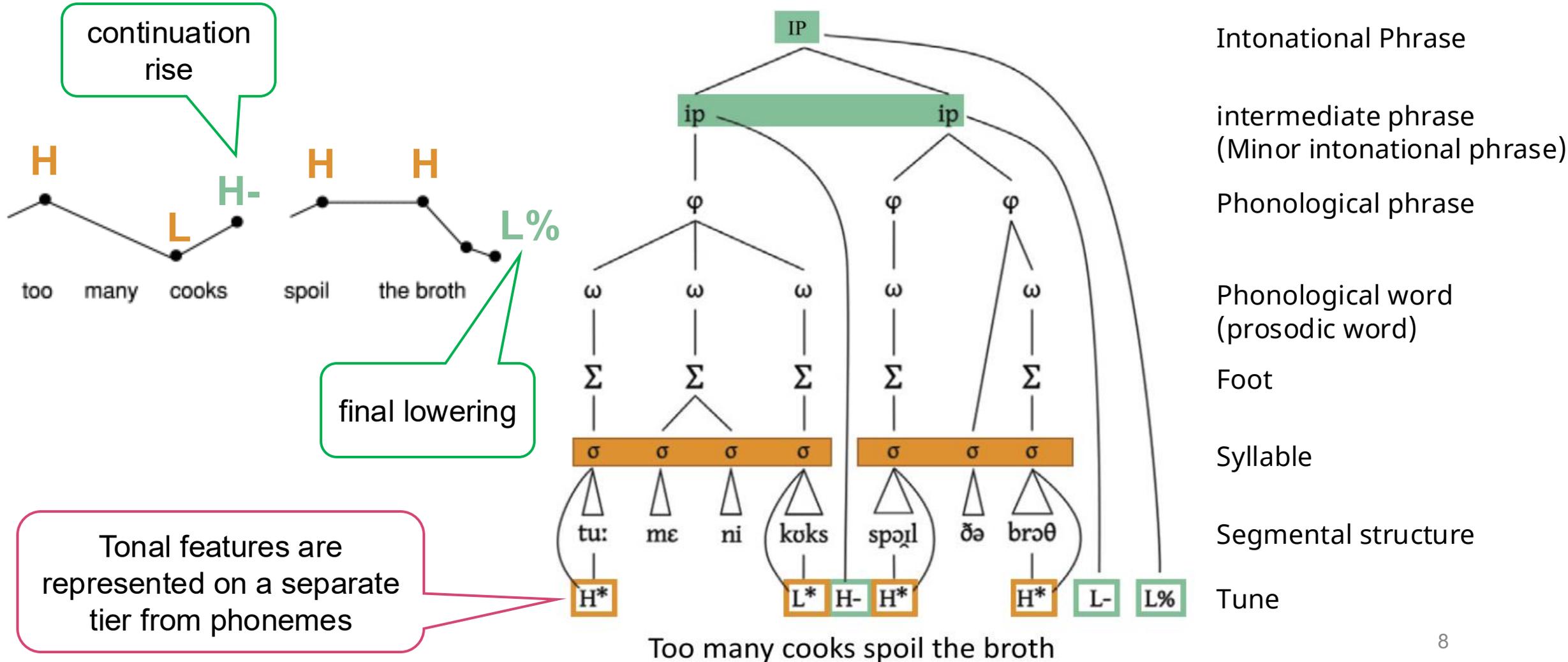
# (Standard) Mandarin Tones

55 ˥ cat **mao** 猫 1
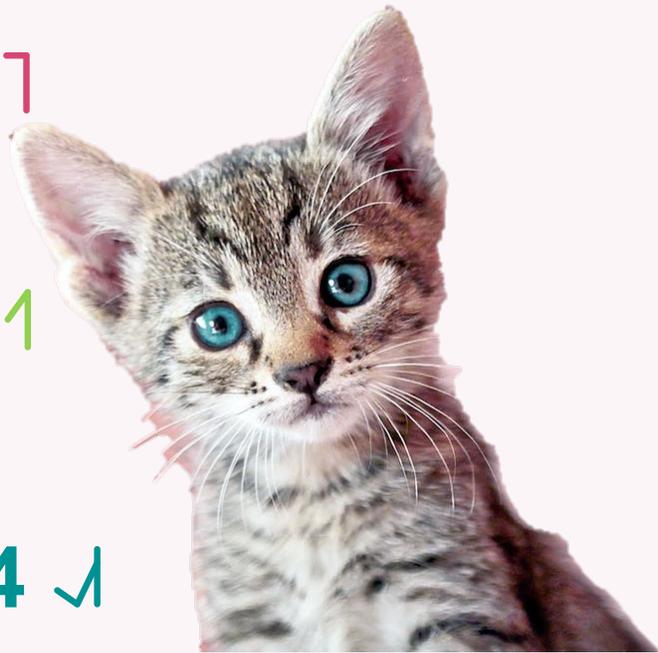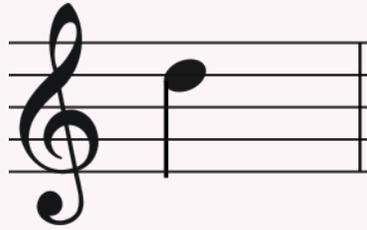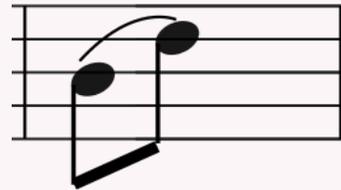
35 ˧˥ cat **mao** 毛 2
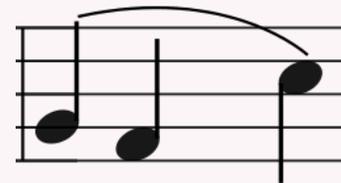
214 ˨˩˦ cat **mao** 卯 3
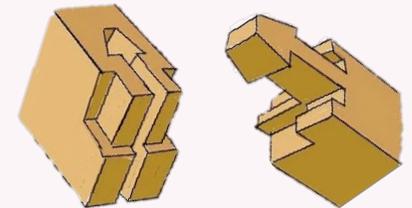
51 ˥˩ cat **mao** 帽 4

# Many Representations of Tone

| 1 | 55 ˥ | H | H |
| 2 | 35 ˧˥ | LH | R |
| 3 | 214 ˨˩˦ | L | L |
| 4 | 51 ˥˩ | HL | F |

(Chao, 1930, 1948, 1968)  (Duanmu, 2000)  (Xu & Wang, 2001)

## Tones change in context

$$T3 \rightarrow T2 / \_\_ T3$$

ni ➕ hao ➡ ni hao

*you*     *good*     *hello*

$$214 \rightarrow 35 / \_\_ 214$$

$$L \rightarrow LH / \_\_ L$$

Underlying Representation

Surface Form

# Competing feature geometry models
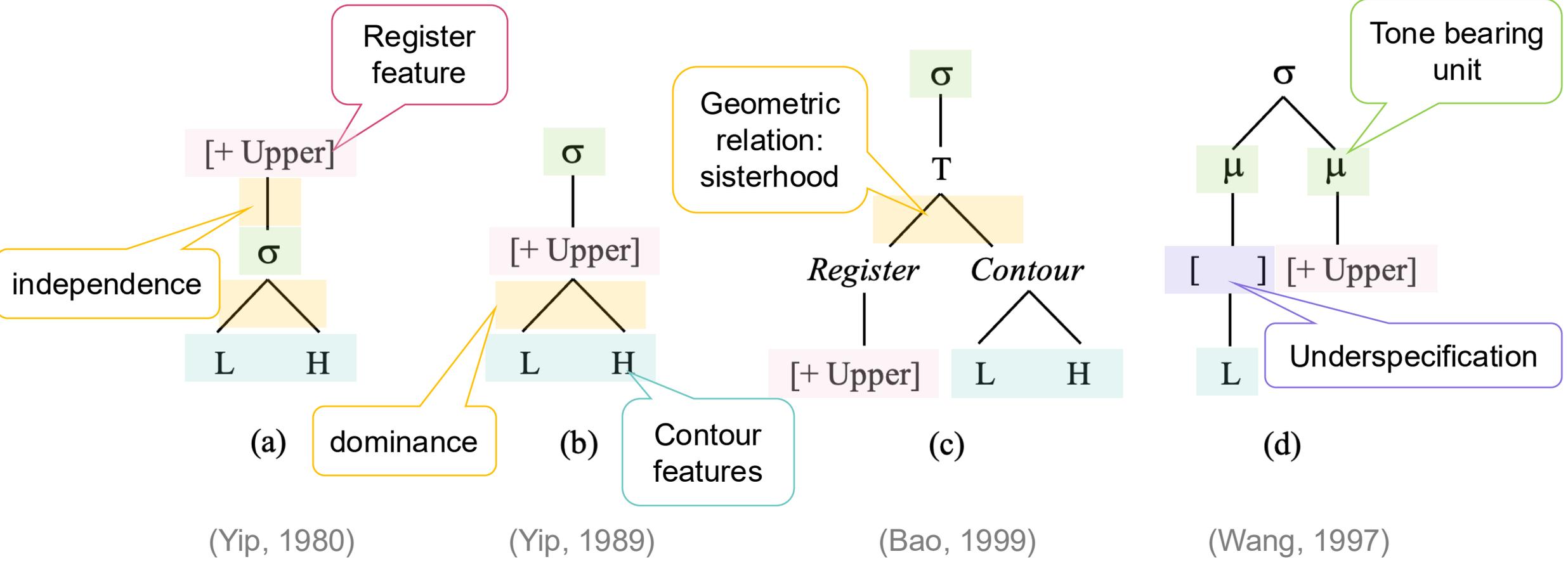


An illustration of **T2** /35 ˥/

# The Puzzle: Neutral tone representation

T**1** | **2** | **3** | **4** → **N**

ma + ma → ma ma

*mother*          *mother*
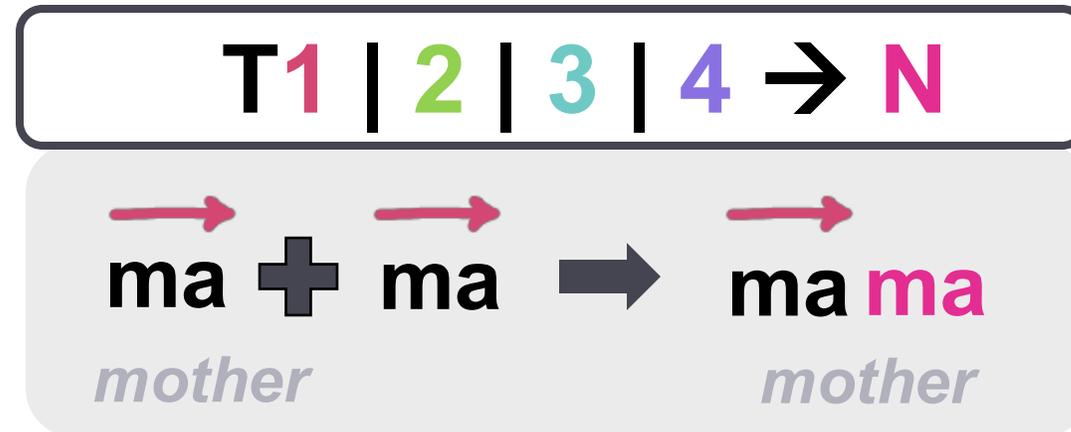
reduplicatives

# Neutral Tone in Standard Mandarin

| | | |
|---|---|---|
| lái **le** 来了 | [laɪ˧˥ lə] | came |
| dōng xī 东西 | [tʊŋ˥ ɕi˥] | east and west |
| dōng **xi** 东西 | [tʊŋ˥ ɕɪ] | thing(s) |
| qí zǐ 棋子 | [tɕʰi˧˥ tsz̩˨˩˦] | chess piece |
| qí **zi** 旗子 | [tɕʰi˧˥ tsz̩] | flag |

## Occurrence and Obligatoriness

- Many are grammatical morphemes
- Never in initial positions
- Some are obligatory and some are optional
- Frequent: One-third of all syllables in colloquial speech (Duanmu, 2007)

# Neutral Tone in Standard Mandarin

| | | | | | |
|---|---|---|---|---|---|
| | *gè* | | *zhèi ge* | | |
| (a) | 个 | /kɤ˥˩/ | 这个 | [ʈʂei˥˩ kə] | |
| | CL | | this one | | |
| | *fāng* | | *dì fang* | | |
| (b) | 方 | /faŋ˥/ | 地方 | [ti˥˩ fə̃] | |
| | direction | | place | | |
| | *fǔ* | | *dòu fu* | | |
| (c) | 腐 | /fu˨˩˦/ | 豆腐 | [toʊf˥˩] | |
| | rotten | | bean curd | | |

## Phonetic Realisation

- Vowel centralisation
- Reduced duration (Lin & Yan, 1980)
- Segment deletion
- Voicing or nasalisation change
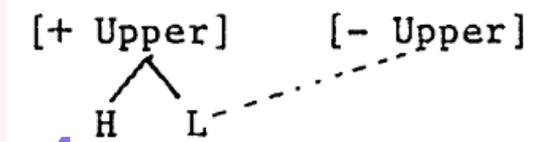- **Varied** but predictable pitch

# Neutral Tone in Standard Mandarin

## Pitch Realisation

- Neutral tone pitch pattern is largely predicted by the **preceding tone**

| Preceding tone | | Neutral tone | | | |
|:---:|:---:|---|---|---|---|
| | | Chao, 1968 | Shih, 1987 | Chen, 2000 | Lee & Zee, 2008 |
| **1** | ˥ | half-low | starts high, then falls | mid | mid falling |
| **2** | ˩ | mid | starts high, then falls, but not as low as after H | mid | high falling |
| **3** | ˨ | half-high | starts fairly low, then rises | half-high | mid-level |
| **4** | ˥ | low | starts fairly low, and falls even lower | low | low falling |

# Approaches to Neutral Tone

## Is it the fifth lexical tone?

| | | Underlying Representation | |
|---|---|---|---|
| | | Specified | Underspecified |
| Explaining surface Variability | Feature spreading | Yip (1980): **[-upper]**<br><br>Lin (2006): **L** | Shen (1992)<br><br>Li (2003): boundary **L%** |
| | Target Interpolation or approximation | Van Santen et al.(1998): A single **mid** target<br><br>Xu & Wang (2001)'s PENTA model: A static **mid** target | Kochanski & Shih (2003)'s STEM-ML model: No soft template |

**1** [+ Upper]   [− Upper]
     H   H

**2** [+ Upper]   [− Upper]
     L   H

**3** [− Upper]   [− Upper]
     L   L        H
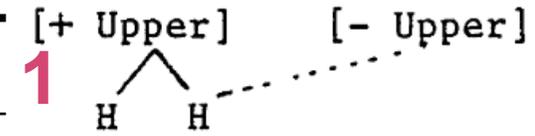
**4** [+ Upper]   [− Upper]
     H   L
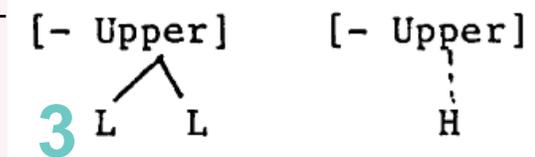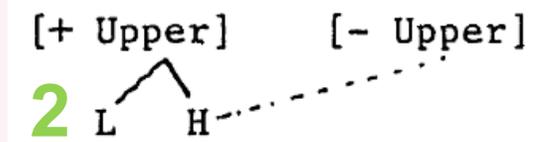
Yip (1980)

# Neutral Tone in Changsha

## Pitch Realisation

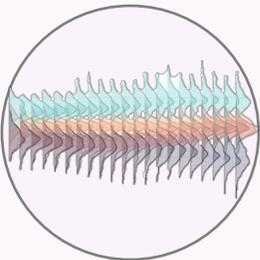| Citation tone | | Neutral tone | |
|:---:|:---:|:---:|:---:|
| | | Pattern A | Pattern B (*s w*) |
| **1** | ˥ | 3 | 5 |
| **2** | ˩ | 3 | 3 or 5 |
| **3** | ˩ | 4 | 3 |
| **4** | ˥ | 5 | 3 or 5 |
| **5** | ˩ | 2 | 3 or 5 |
| **6** | ˩ | 4 | 3 or 5 |

Zhong (2003)

- Both left-dominant (*s w*) and right-dominant (*w s*) words (Lin, 2011)
- Neutral tone pitch pattern A depends on the **underlying tone** (Zhong, 2003)
- Word-final neutral tone: short and level (Guo & Chen, 2022)

**RoadMap**

# Tone change:
## Case Study of Plastic Mandarin

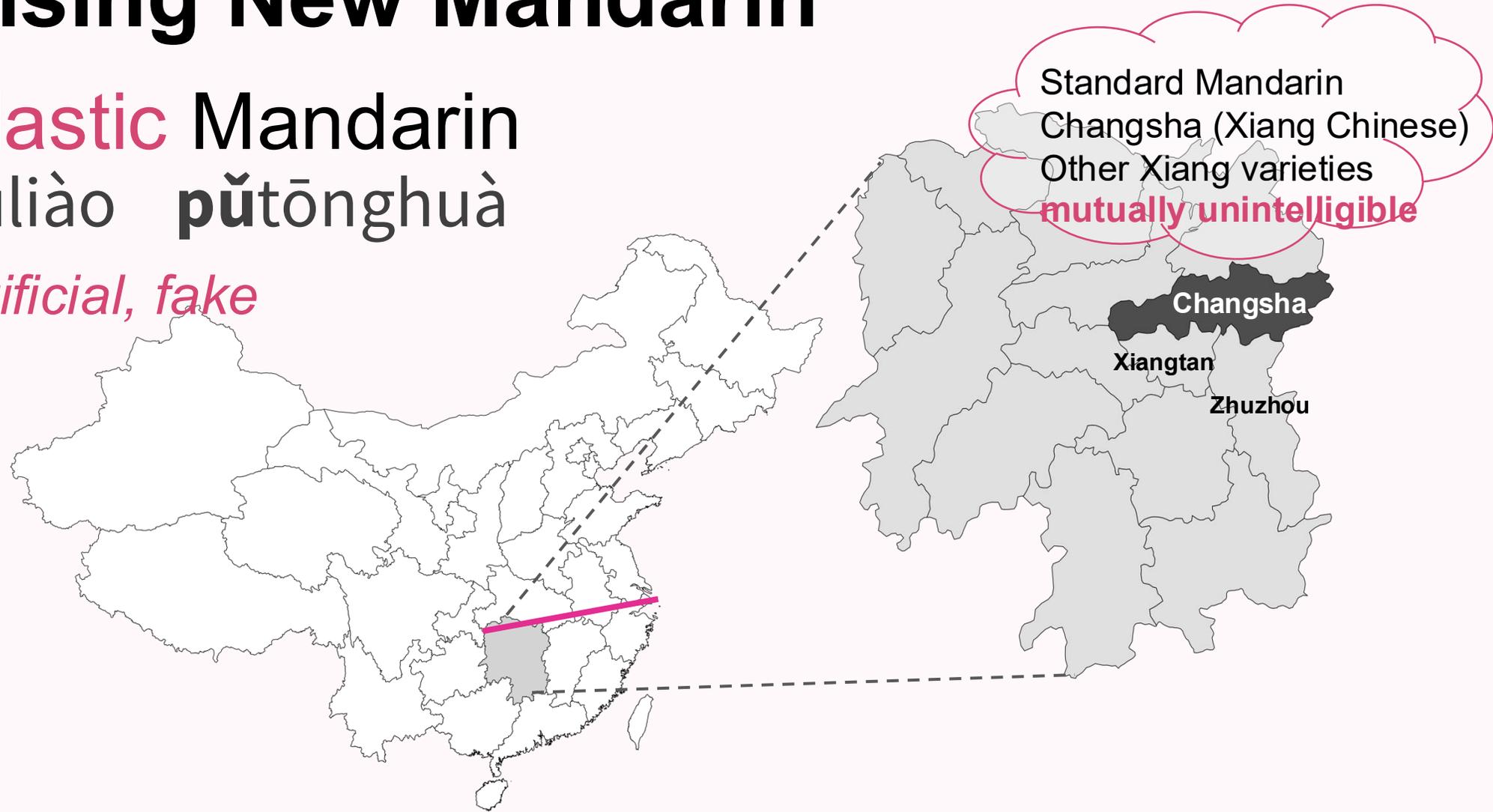In this talk: **Citation tones (A) and neutral tone (B)**

**Central propositions**

1. Mandarin Neutral Tone is **underspecified** and interacts with tonal features at different prosodic hierarchy

2. Mandarin Neutral Tone associated with post-lexical boundary tone is more **pertinacious** than lexical tones in language variation and change

# Crystalising New Mandarin

**Plastic** Mandarin

**sù**liào  **pǔ**tōnghuà

*artificial, fake*

Standard Mandarin
Changsha (Xiang Chinese)
Other Xiang varieties
**mutually unintelligible**

Changsha

Xiangtan

Zhuzhou

# Crystalising New Mandarin

## Language contact

- Large-scale migration of newcomers
- Existing variety "swamped" by the incomer varieties

e.g. **South African English** (Lass, 1990)

- Mutually unintelligible varieties of migrants
- No language shared by a large enough minority
- An economically or politically powerful language as a remote target

e.g. **Portuguese-Lexified Creoles**

(Cardoso, 2020)

## Less obvious contact

- A population adopts a different language
- A standard or official language
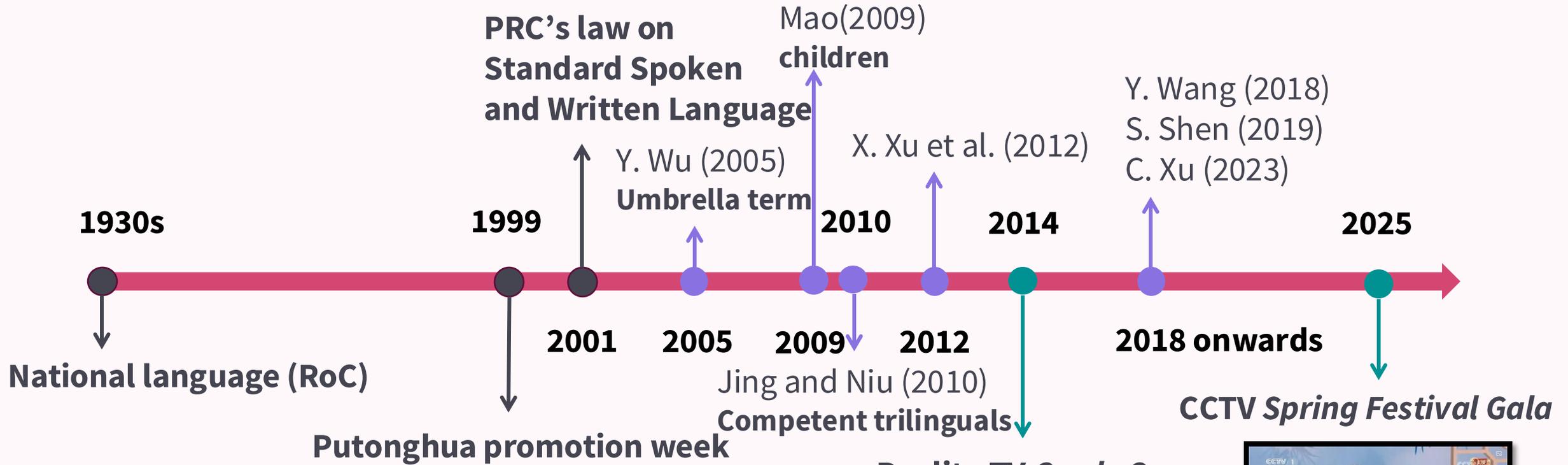- Digital communication

**High Standard Mandarin**
Education and other formal domains

↕

**Plastic Mandarin**

**Low Changsha**
Conversations at home and wet market
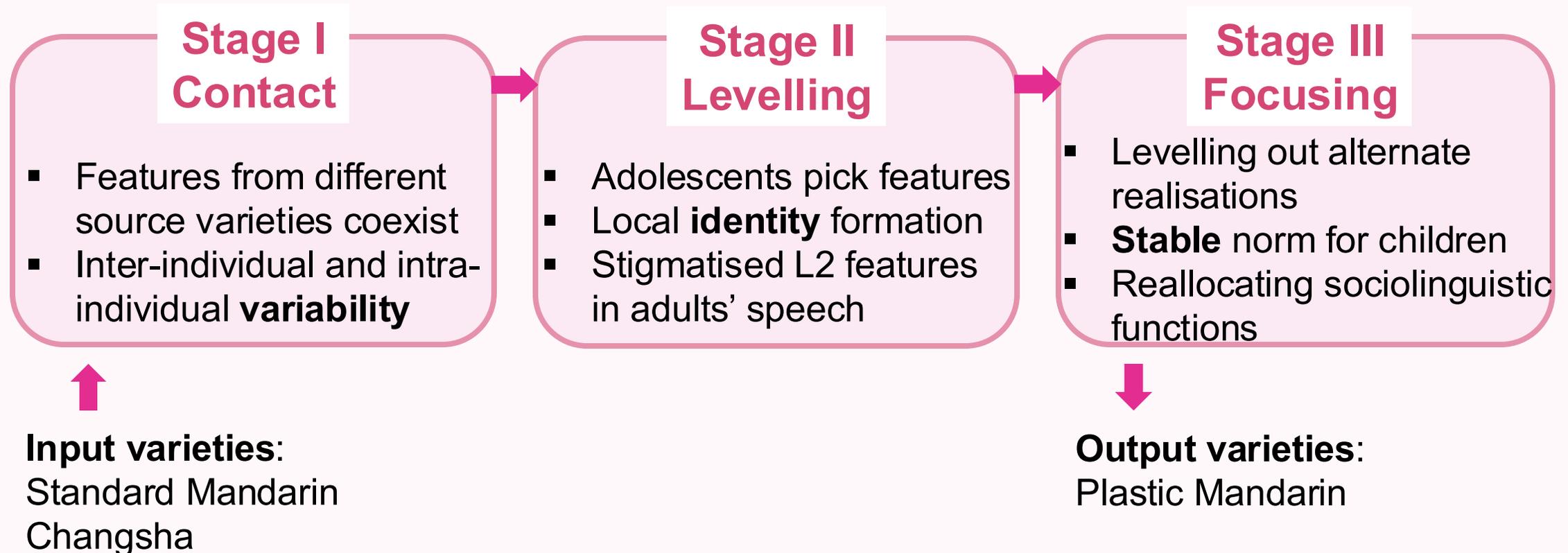
# Plastic Mandarin Timeline

PRC's law on Standard Spoken and Written Language

Mao(2009) children

Y. Wang (2018)
S. Shen (2019)
C. Xu (2023)

Y. Wu (2005) Umbrella term

X. Xu et al. (2012)

**1930s**

**1999**

**2010**

**2014**

**2025**

**2001**

**2005**

**2009**

**2012**

**2018 onwards**

National language (RoC)

Jing and Niu (2010)
Competent trilinguals

CCTV *Spring Festival Gala*

Putonghua promotion week

Reality TV *Grade One*



- ● **Policy**
- ● **Academic**
- ● **Media**

# Crystalising New Mandarin

Trudgill's three-stage model of New Dialect Formation (Trudgill et al., 2000)
**Three-stage model for Plastic Mandarin**

| **Stage I Contact** | **Stage II Levelling** | **Stage III Focusing** |
|---|---|---|
| ▪ Features from different source varieties coexist<br>▪ Inter-individual and intra-individual **variability** | ▪ Adolescents pick features<br>▪ Local **identity** formation<br>▪ Stigmatised L2 features in adults' speech | ▪ Levelling out alternate realisations<br>▪ **Stable** norm for children<br>▪ Reallocating sociolinguistic functions |

**Input varieties**:
Standard Mandarin
Changsha

**Output varieties**:
Plastic Mandarin

# Crystalising Tonal Change

师生

*'teacher(s) and student(s)'*

**shī shēng**

读书

*'read book(s)'*

**dú shū**

| | 师生 | 读书 |
|---|---|---|
| Standard Mandarin | /ʂʐ˥ ʂəŋ˥/ | /tu˧˥ ʂu˥/ |
| Changsha | [sʐ˨ sən˨] | [təʊ˨ ɕy˨] |
| Hunan-accented Mandarin | [sʐ˥ sən˥] | [təʊ˧˥ ɕy˥] |
| **Plastic Mandarin** | [sʐ˨ sə˨] | [tu˨ ʂu˨] |

Velar nasal **/ŋ/**
- Alveolar [n]

Retroflex fricative **/ʂ/**
- Alveolar [s]
- Alveolo-palatal [ɕ]

Hunan-accented examples taken from a transcript in Y. Wu (2005, p.29)

# A. Research Questions

**RQ1** What is the prototypical pitch contour of each tone in Plastic Mandarin?

**RQ2** How similar the pitch contours are between the tones of Plastic Mandarin and Changsha?

# RQ1. Plastic Mandarin Tones
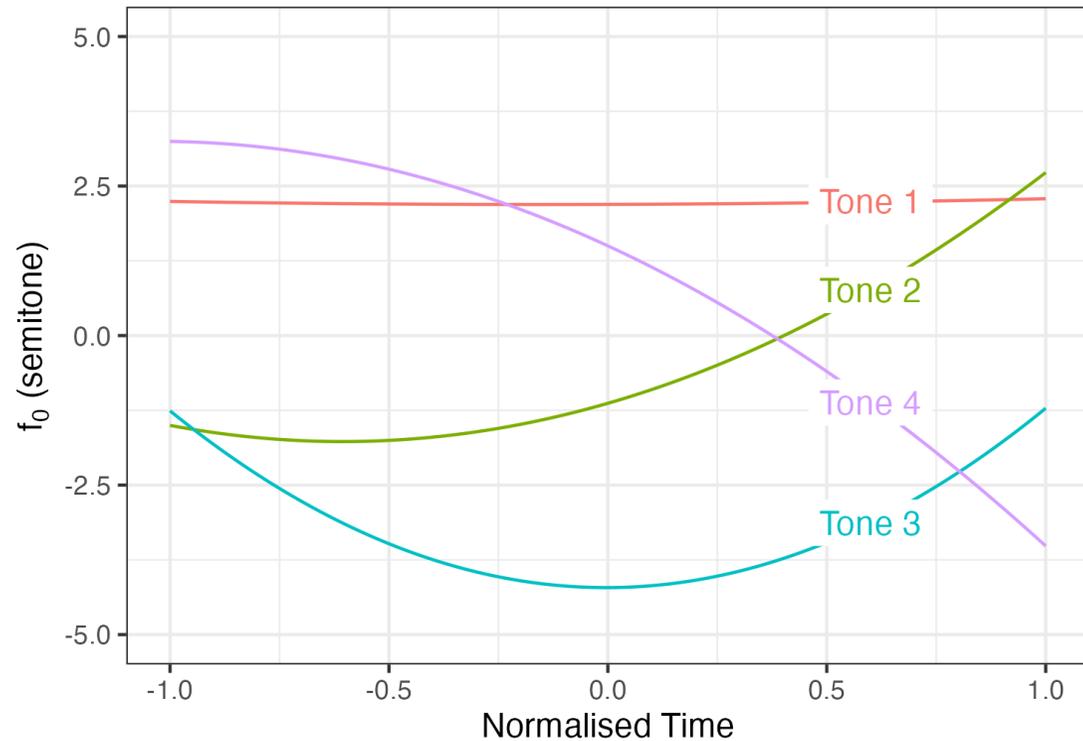

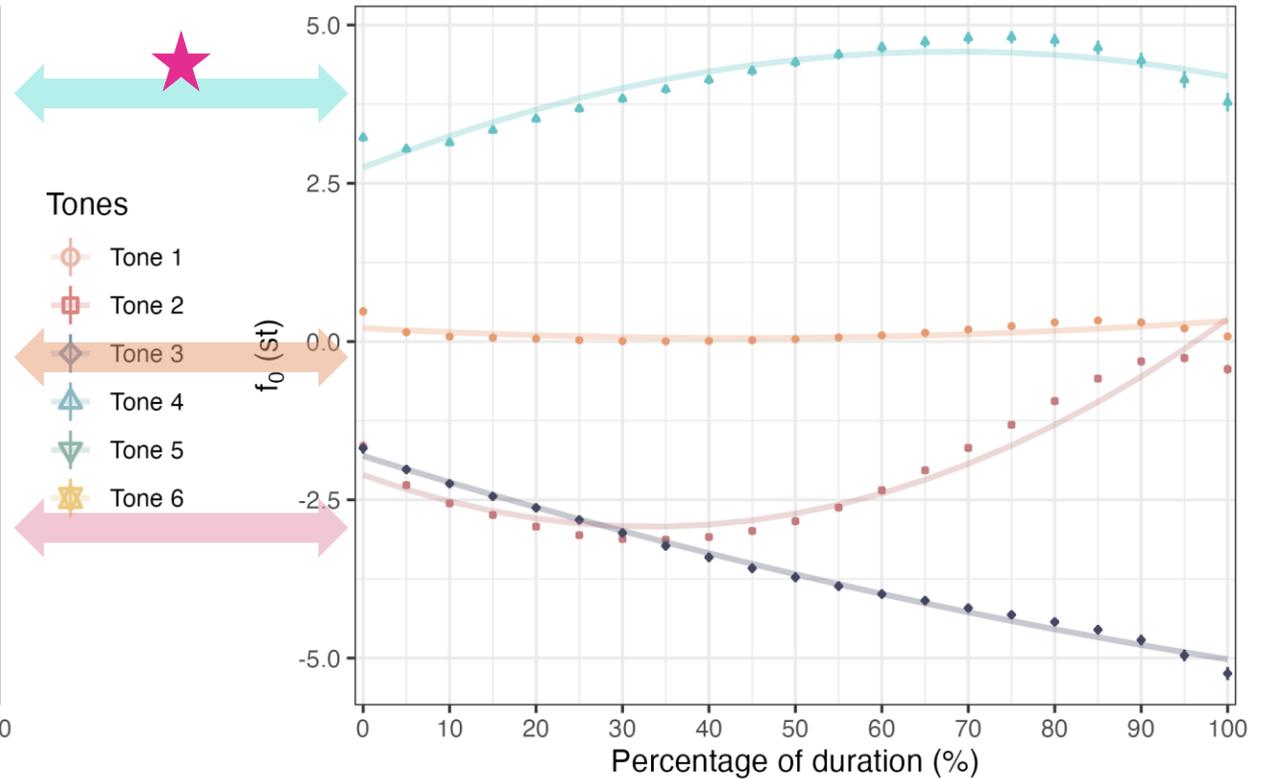
Standard Mandarin    (Xu, 2025)    Plastic Mandarin
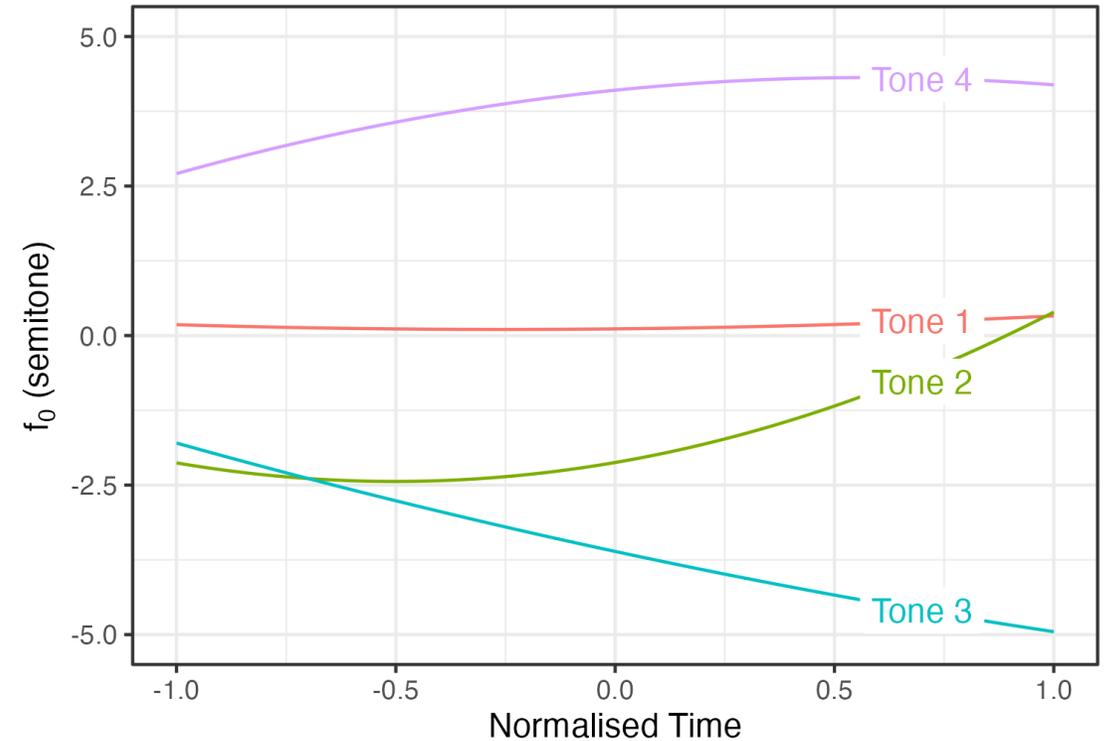
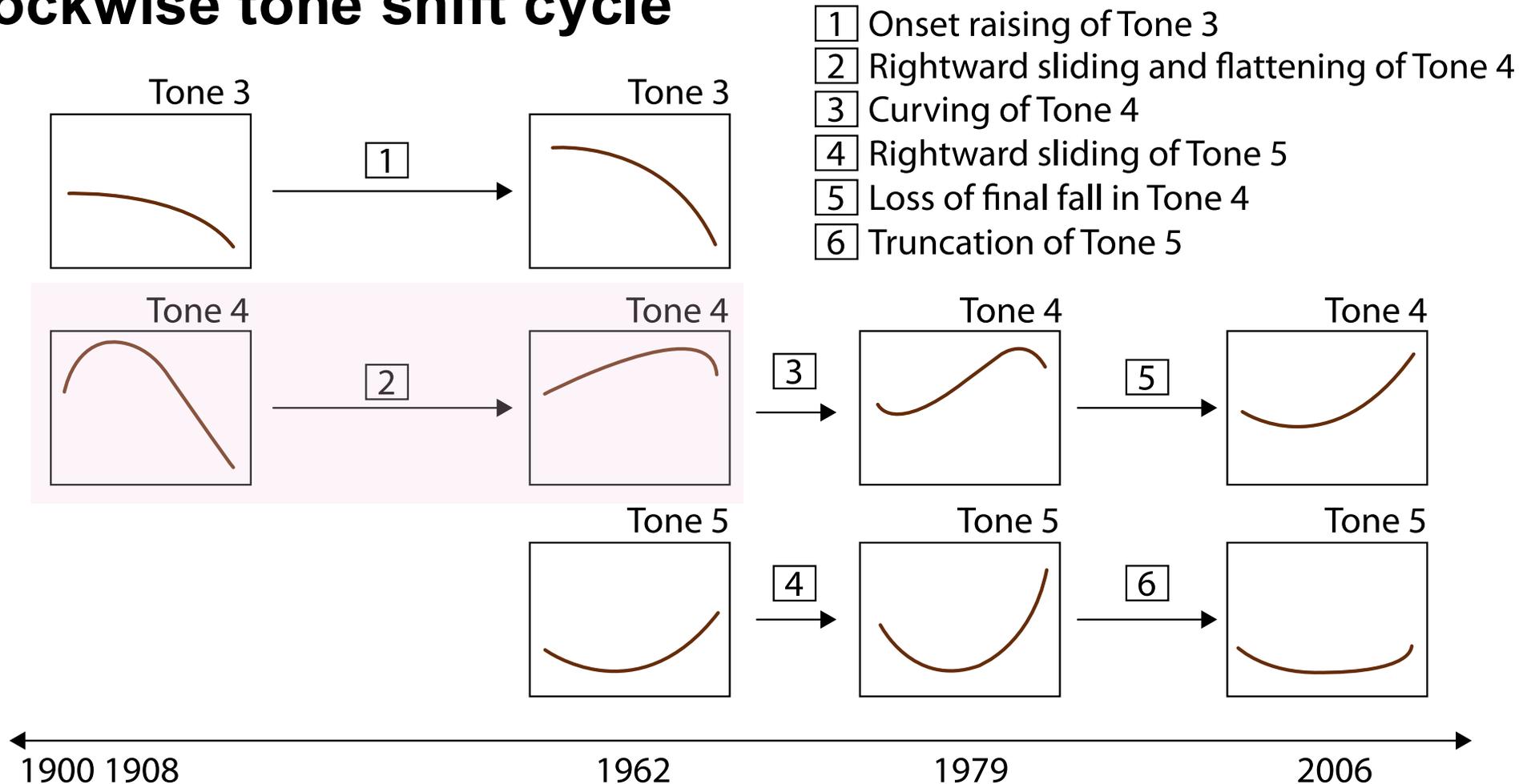# RQ2. Changsha VS Plastic Mandarin Tones
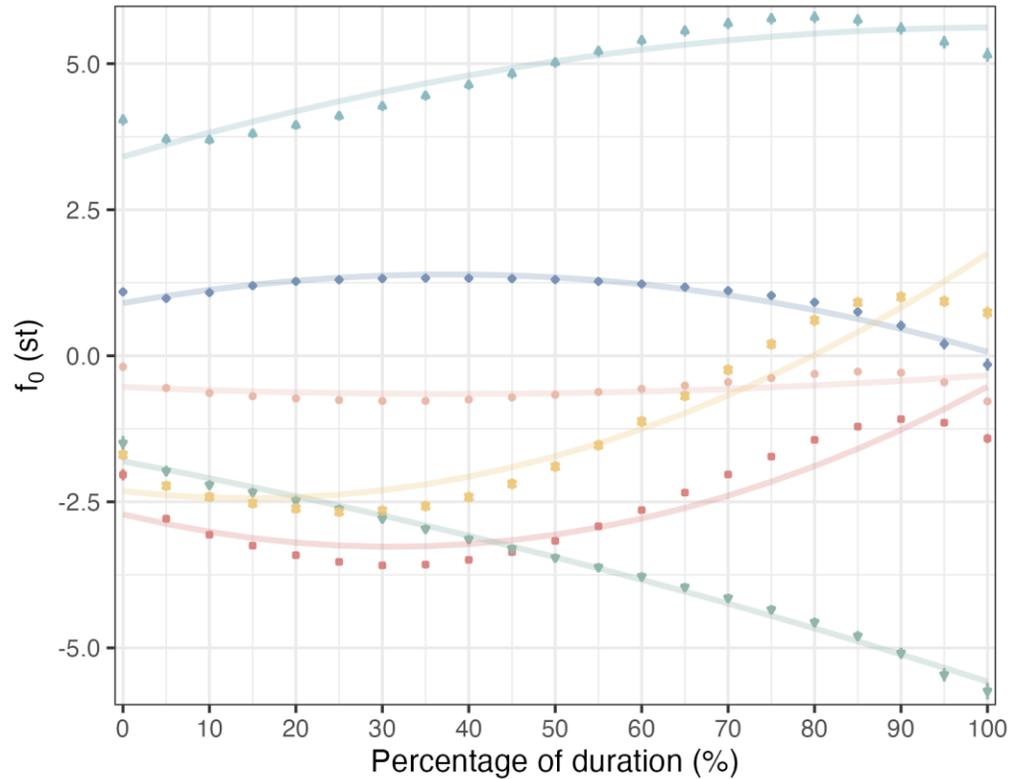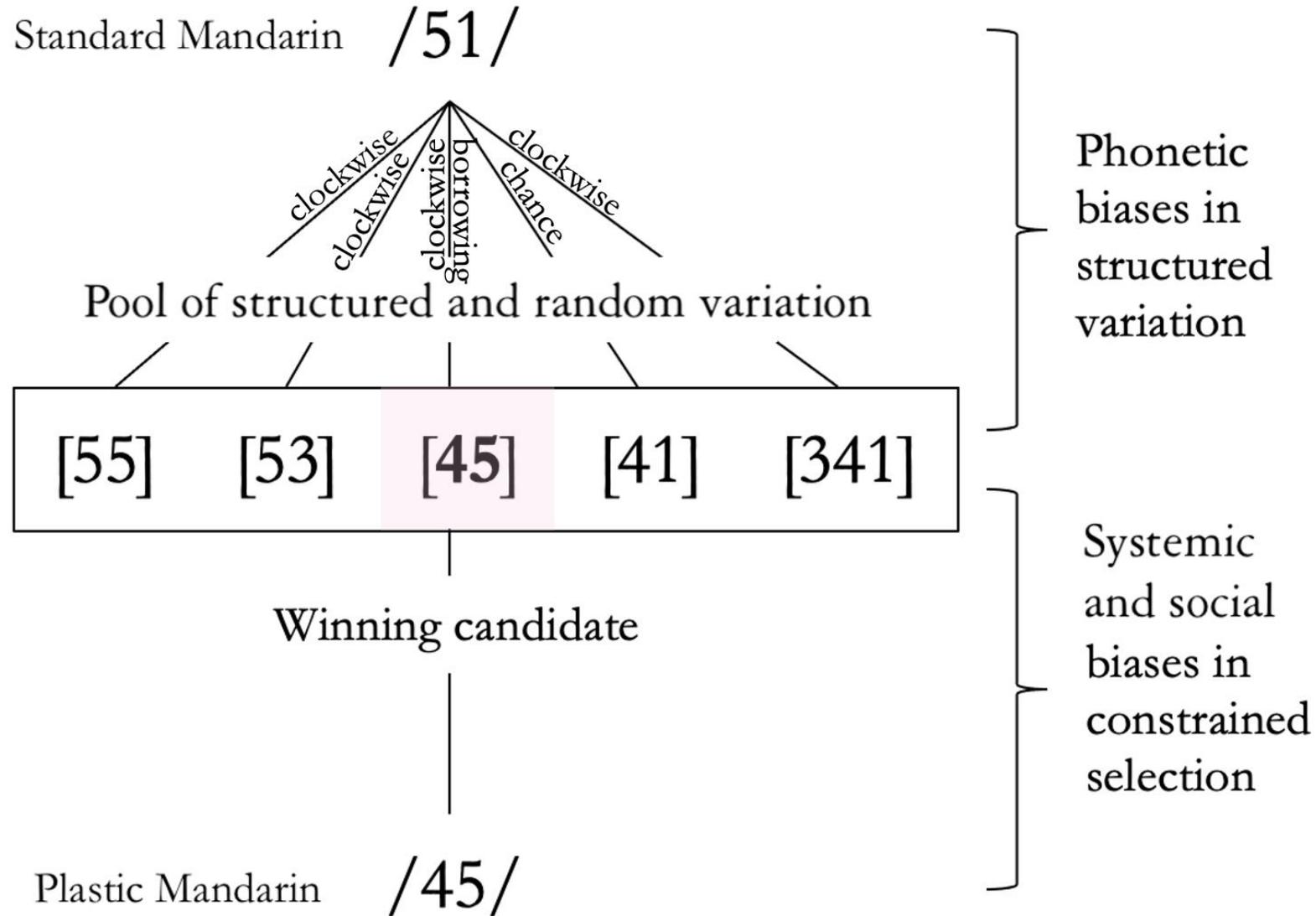


Changsha        (Xu, 2025)        Plastic Mandarin

# Recurrent tone contour changes
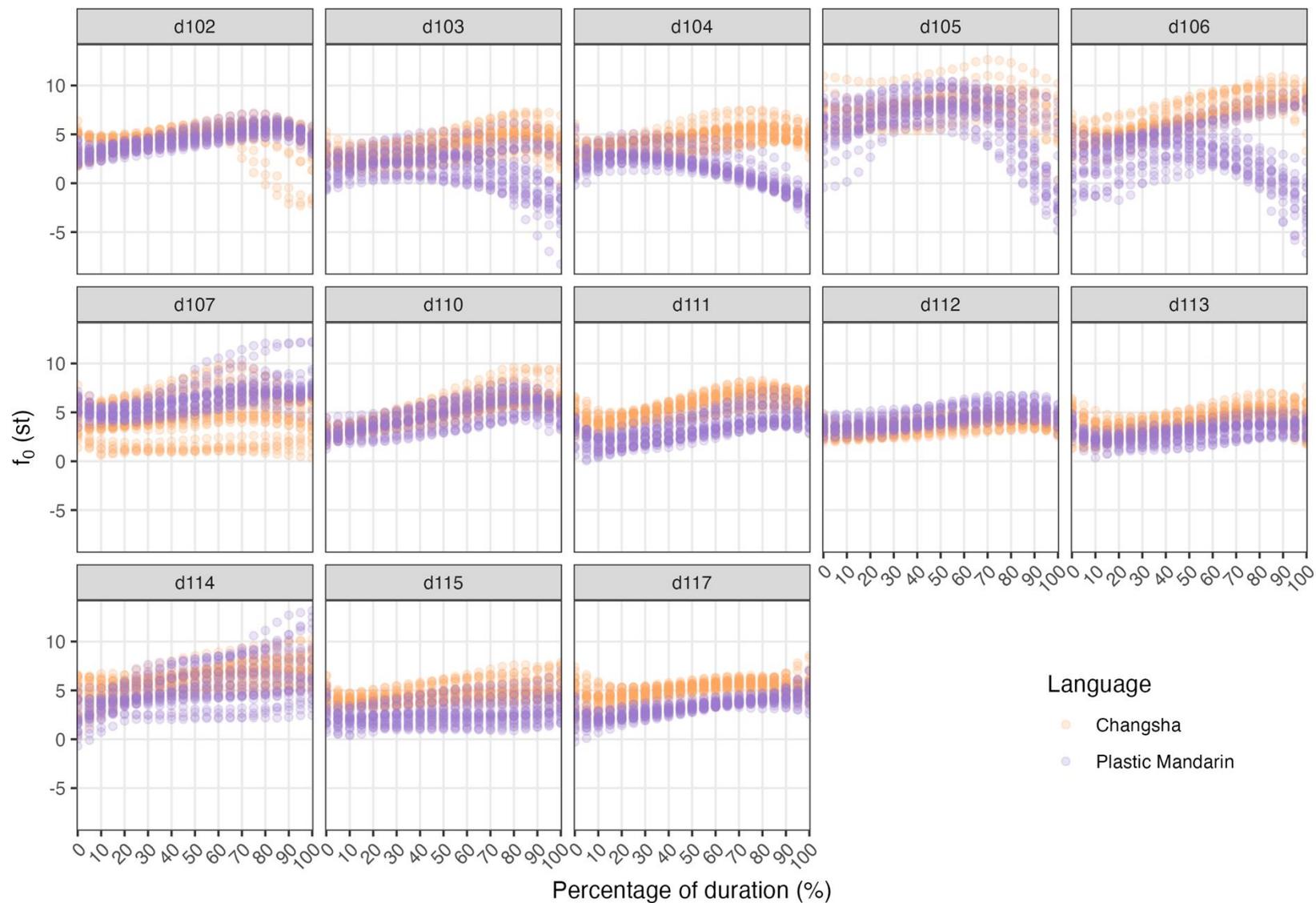
## Clockwise tone shift cycle

1 Onset raising of Tone 3
2 Rightward sliding and flattening of Tone 4
3 Curving of Tone 4
4 Rightward sliding of Tone 5
5 Loss of final fall in Tone 4
6 Truncation of Tone 5



Tone changes in Bangkok Thai (Pittayaporn, 2018)

# Mandarin lexical tone variation

# Mandarin lexical tone variation

# B. Research Questions

**RQ1** How is a neutral tone realised in various tonal contexts in Plastic Mandarin?

**RQ2** How do neutral tone patterns in Plastic Mandarin (PM) compare to those in Standard Mandarin (SM)?

**RQ3** Is there a pitch target for neutral tone?

# B. Research Questions

**RQ1** How is a neutral tone realised in various tonal contexts in Plastic Mandarin?

**RQ2** How do neutral tone patterns in Plastic Mandarin (PM) compare to those in Standard Mandarin (SM)?

**RQ3** Is there a pitch target for neutral tone?

**Preceding tone**

**Duration**

**Following tone**

# B. Research Questions

RQ1 How is a neutral tone realised in various tonal contexts in Plastic Mandarin?

**RQ2** How do neutral tone patterns in Plastic Mandarin (PM) compare to those in Standard Mandarin (SM)?

RQ3 Is there a pitch target for neutral tone?

### H1
**PM ≈ SM neutral tones**

### H2
**PM ≈ Changsha neutral tones**

### H3
**PM neutral tones are unique**

# B. Research Questions

**RQ1** How is a neutral tone realised in various tonal contexts in Plastic Mandarin?

**RQ2** How do neutral tone patterns in Plastic Mandarin (PM) compare to those in Standard Mandarin (SM)?

**RQ3** Is there a **pitch target** for neutral tone?

**H1**
**Yes, underlyingly a mid or low tone**

**H2**
**No**

**H3**
**Yes, attracts a low boundary tone**

# Data and Method

## Plastic Mandarin

- 16 females, 5 males
- Age: 17 ± 0.7 years
- On average 15.7 years in Changsha

## Standard Mandarin

- 9 females, 5 males
- Age: 24 ± 2 years
- Mandarin region in northern China
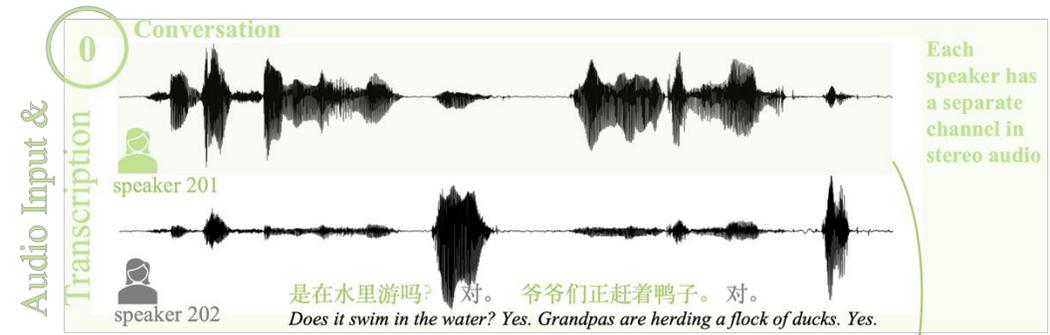- 8/14 spent more than 10 years in Beijing

## Disguised friendship game

- Peer group pair design
- Carefully design materials
- Increasing number of neutral tones
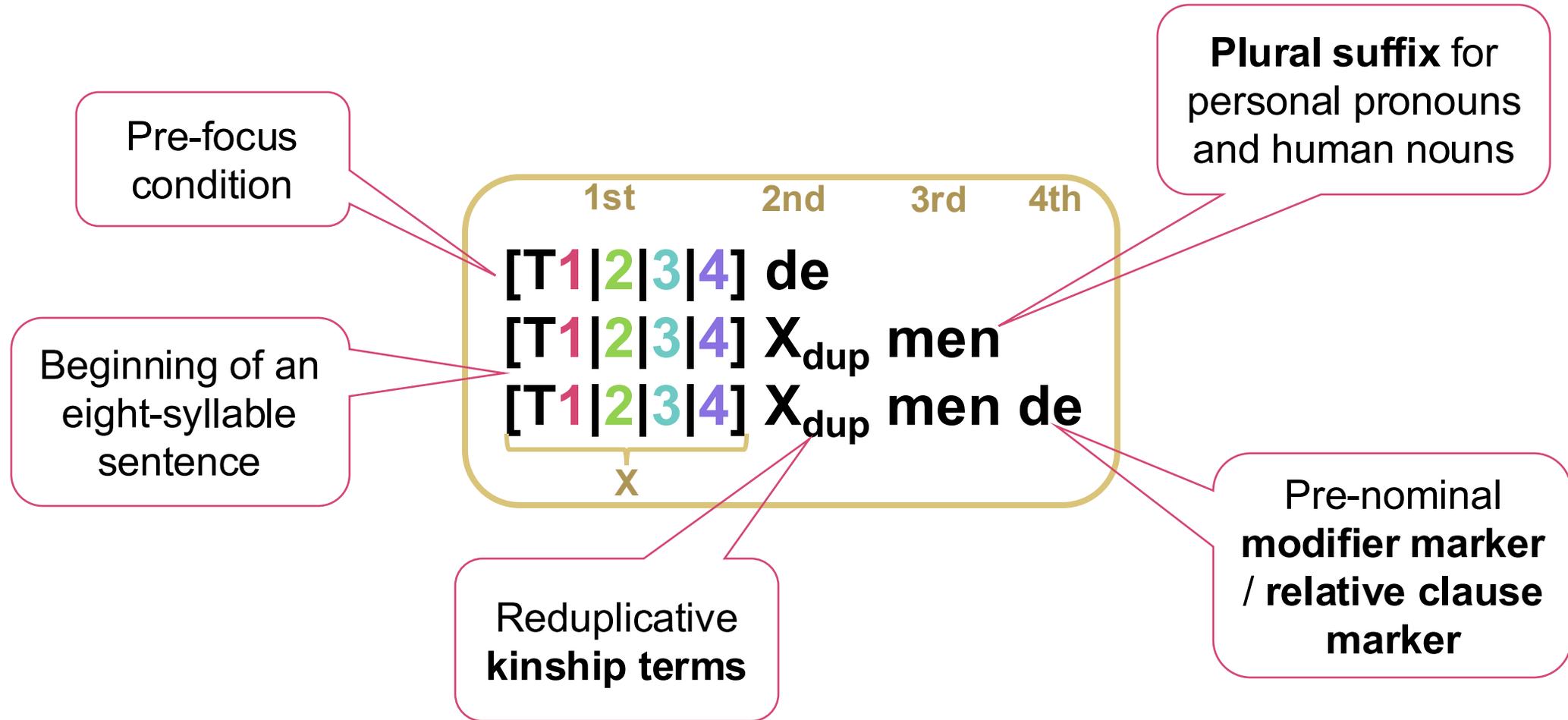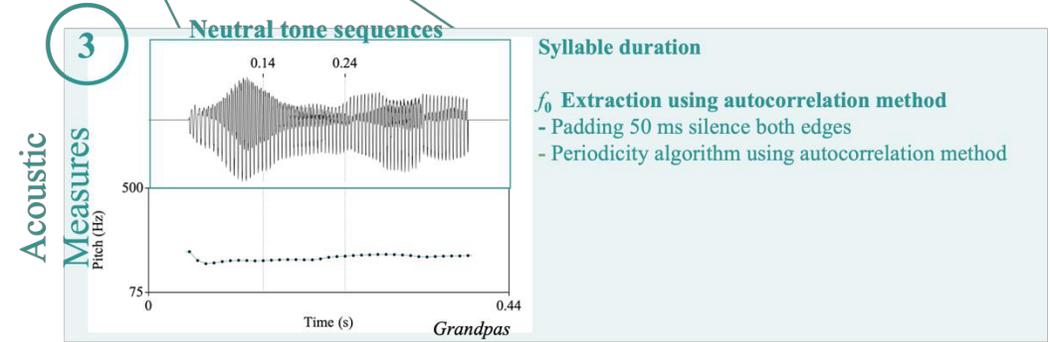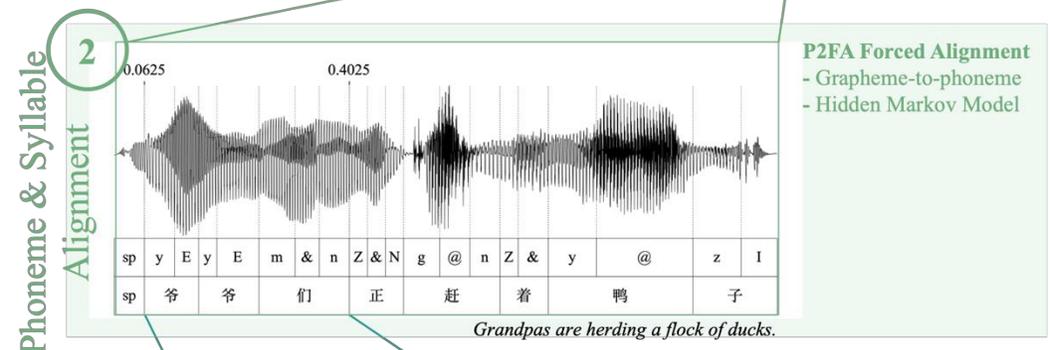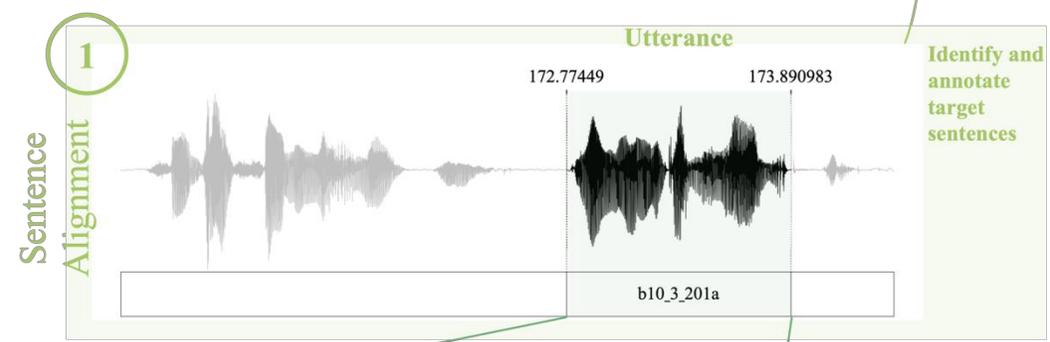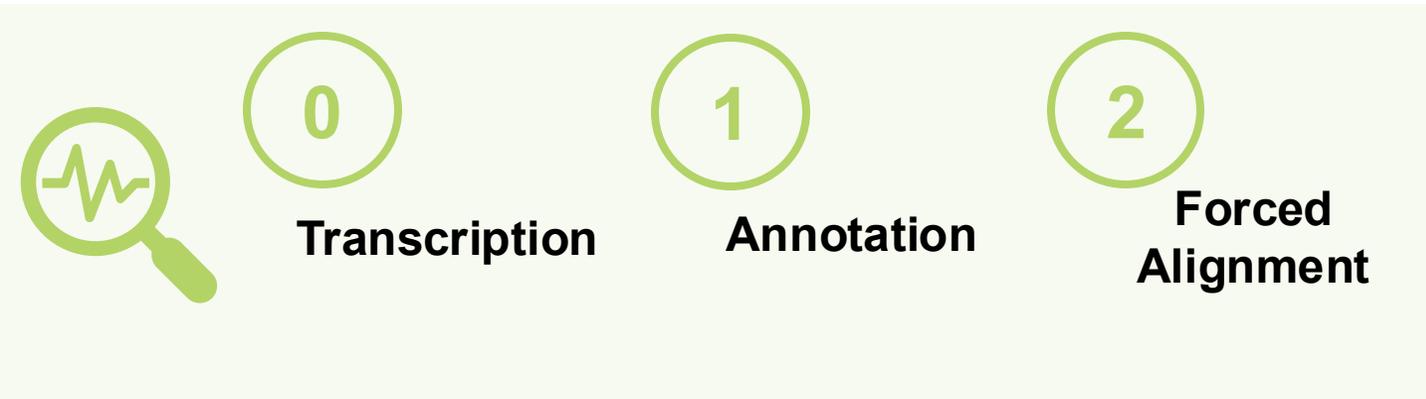
|  |  |  |  |
|---|---|---|---|
| 1st | 2nd | 3rd | 4th |

$[T1|2|3|4]$ **de**

$[T1|2|3|4]\ X_{dup}$ **men**

$[T1|2|3|4]\ X_{dup}$ **men de**

X

# Data and Method

Pre-focus condition

Plural suffix for personal pronouns and human nouns

Beginning of an eight-syllable sentence

|  | 1st | 2nd | 3rd | 4th |
|---|---|---|---|---|

$[T1|2|3|4]$ de
$[T1|2|3|4]$ $X_{dup}$ men
$[T1|2|3|4]$ $X_{dup}$ men de

X

Reduplicative kinship terms

Pre-nominal modifier marker / relative clause marker

# Data and Method



**0** Transcription

**1** Annotation

**2** Forced Alignment

**3** Acoustic Measurement

**4** Data Normalisation

**5** Generalised Additive Mixed Models

(Xu, 2024)

# Results (RQ1): Neutral tones in Plastic Mandarin

*How is a neutral tone realised in various tonal contexts in Plastic Mandarin?*
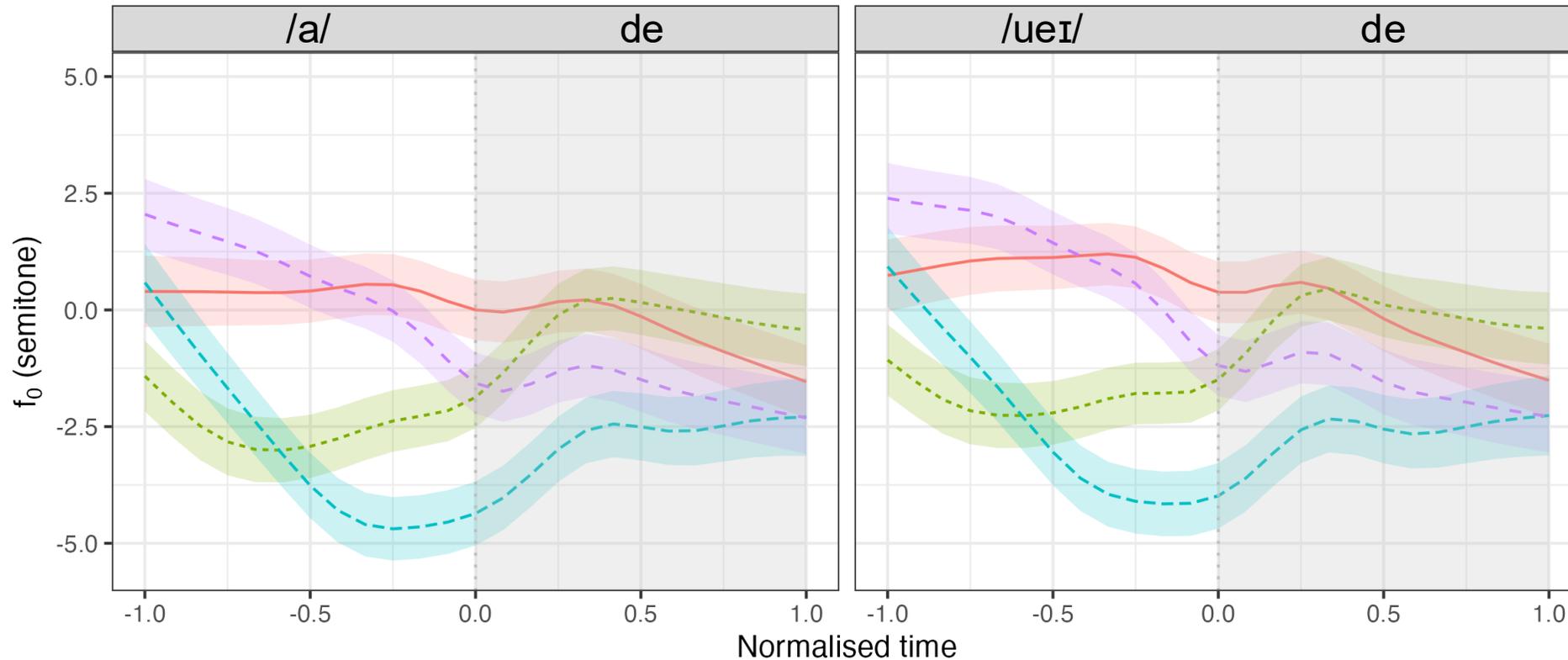
**Preceding tone matters!**
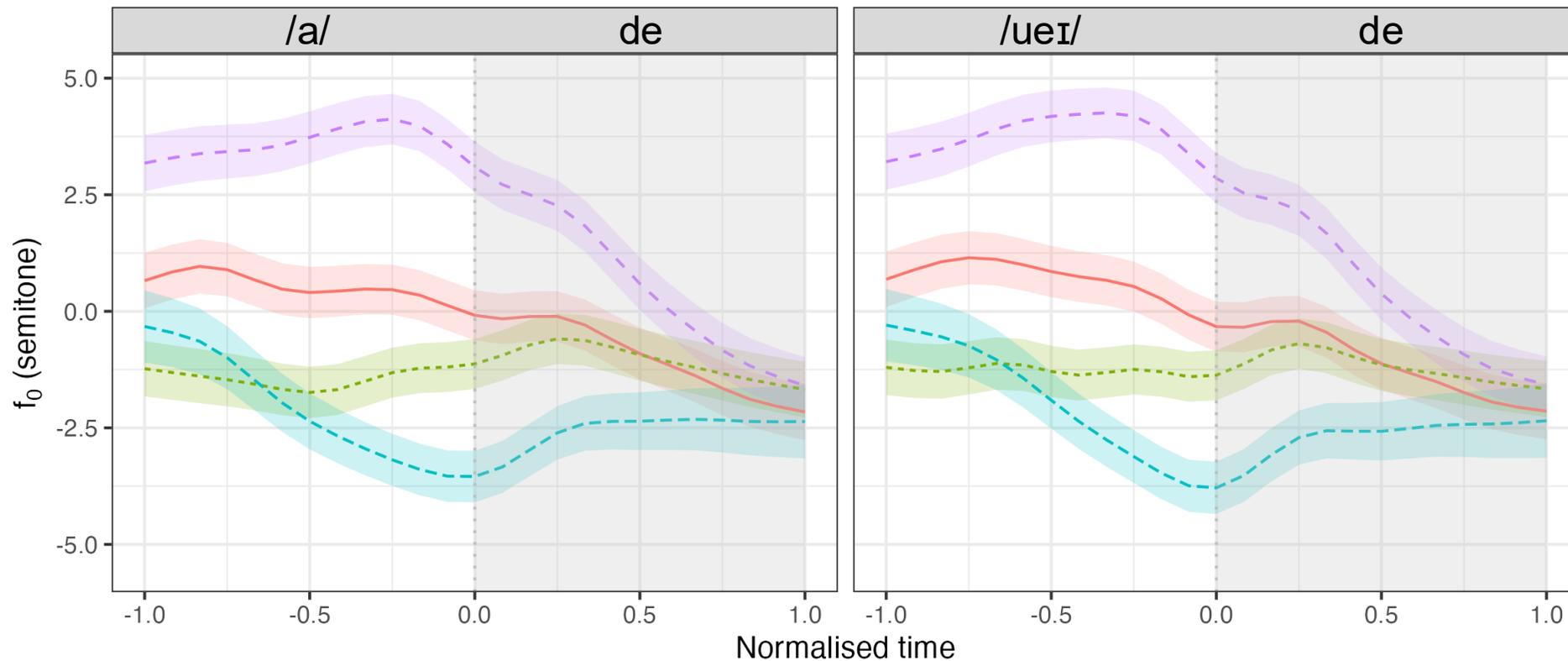
[T1|2|3|4] de

# Results (RQ1): Neutral tones in Plastic Mandarin

*How is a neutral tone realised in various tonal contexts in Plastic Mandarin?*

**Duration matters!**

da4 de
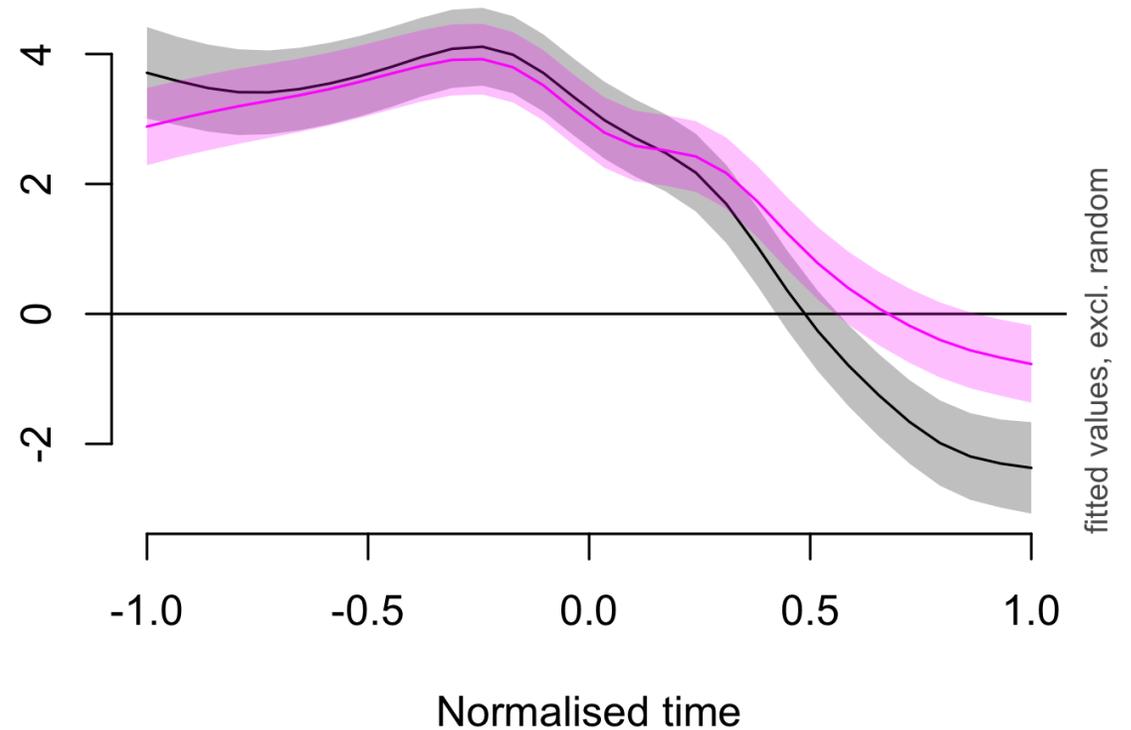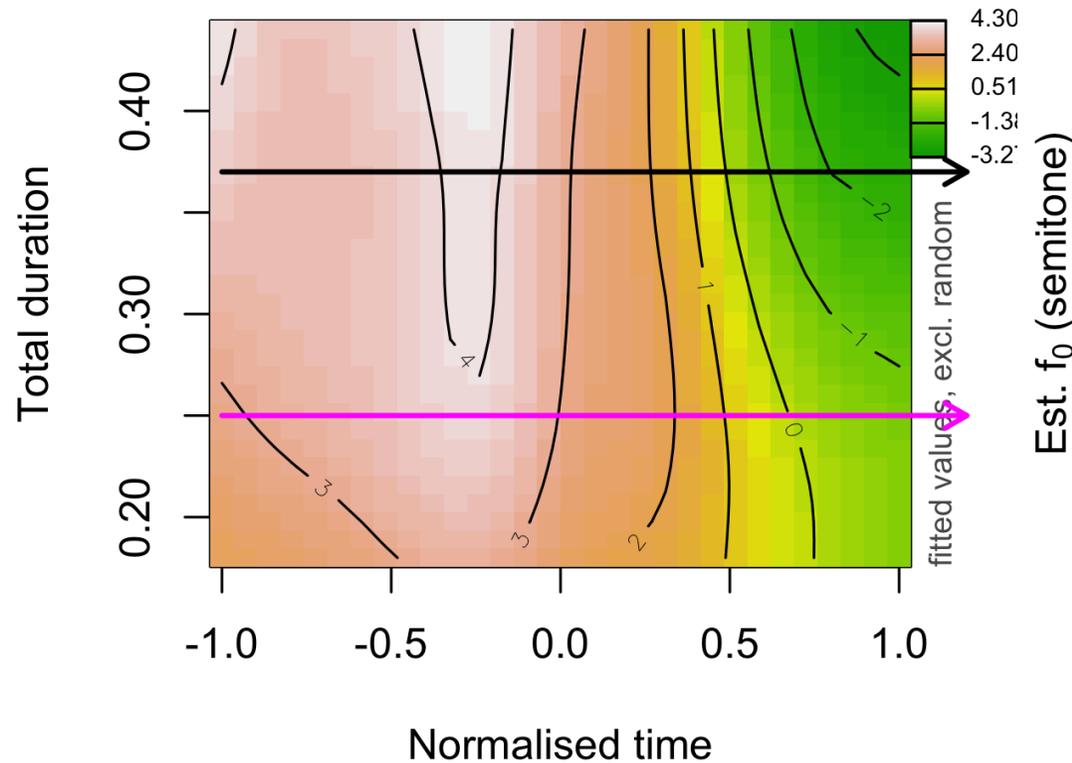
# Results (RQ1): Neutral tones in Plastic Mandarin

*How is a neutral tone realised in various tonal contexts in Plastic Mandarin?*
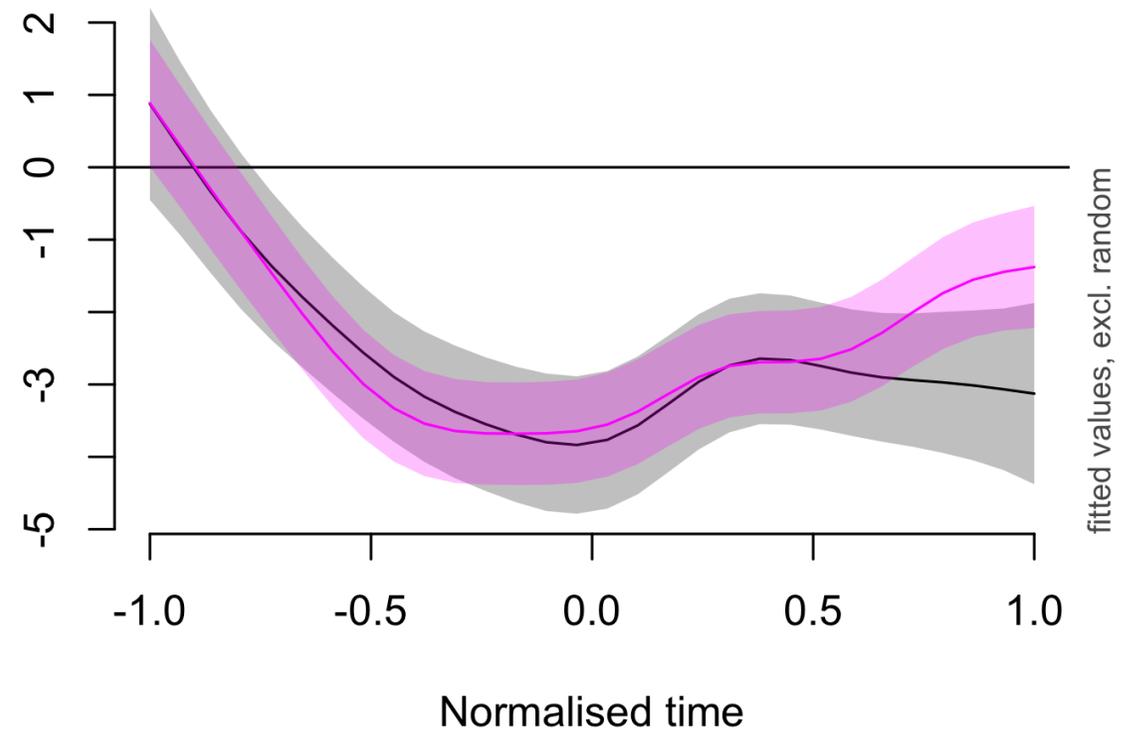
**Following tone: not so much**

[T1|2|3|4] X$_{dup}$ men
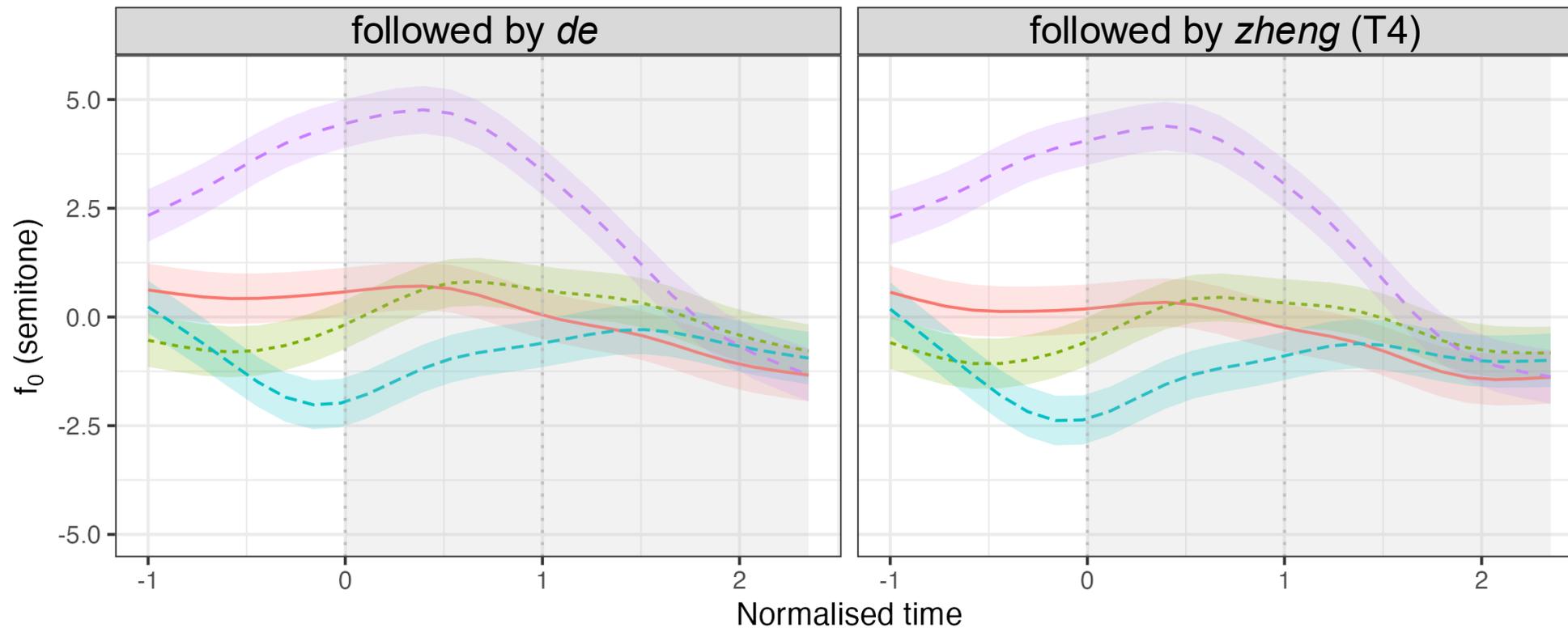
# Results (RQ2): Neutral tones in Standard Mandarin

*How do neutral tone patterns in Plastic Mandarin compare to those in Standard Mandarin?*

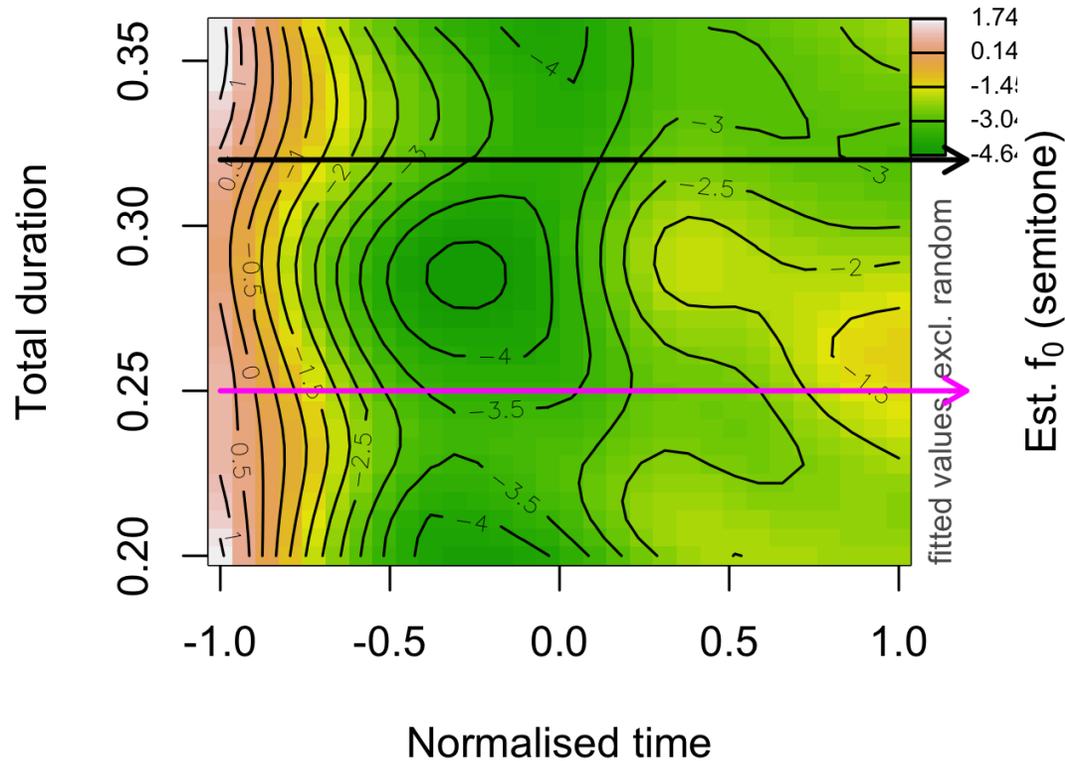**Similarly preceding tone matters**

[T1|2|3|4] de

# Results (RQ2): Neutral tones in Standard Mandarin

*How do neutral tone patterns in Plastic Mandarin compare to those in Standard Mandarin?*
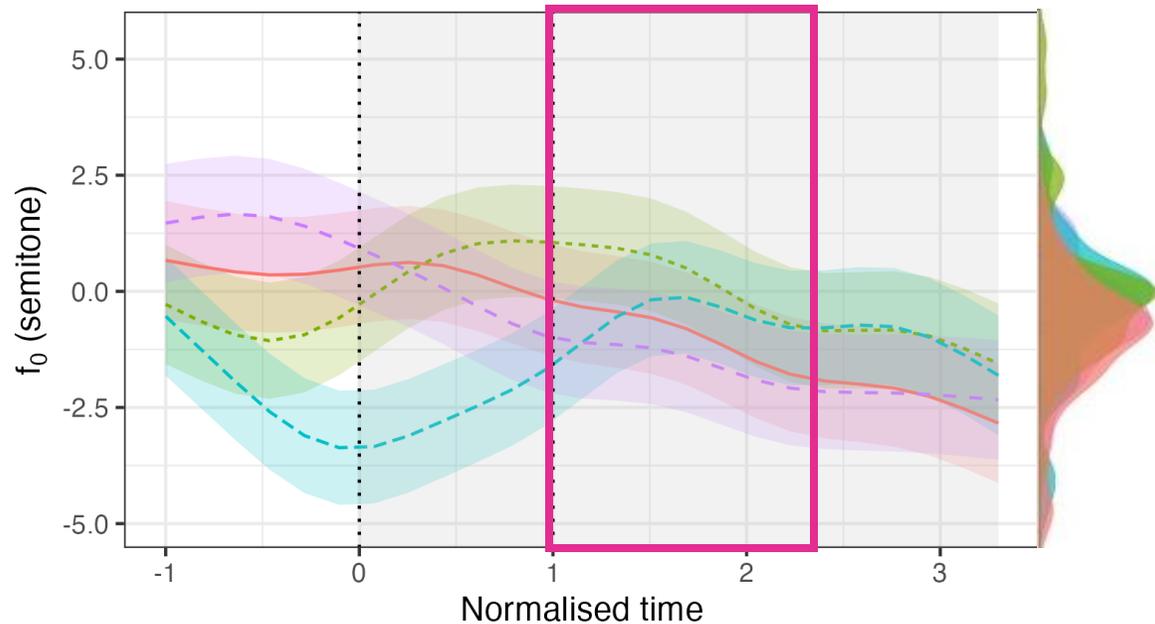
**Similarly duration matters**

da3 de

# Results (RQ2): Neutral Tone Comparison

*How do neutral tone patterns in Plastic Mandarin compare to those in Standard Mandarin?*
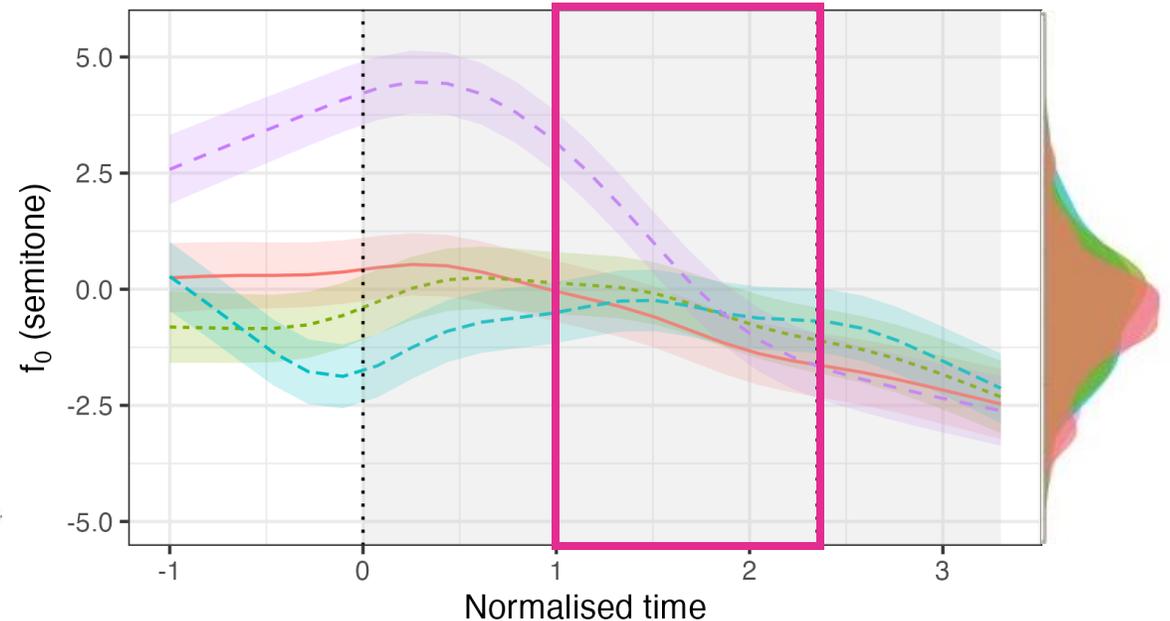
**Similarly converging contours**

[T1|2|3|4] $X_{dup}$ men de



Standard Mandarin
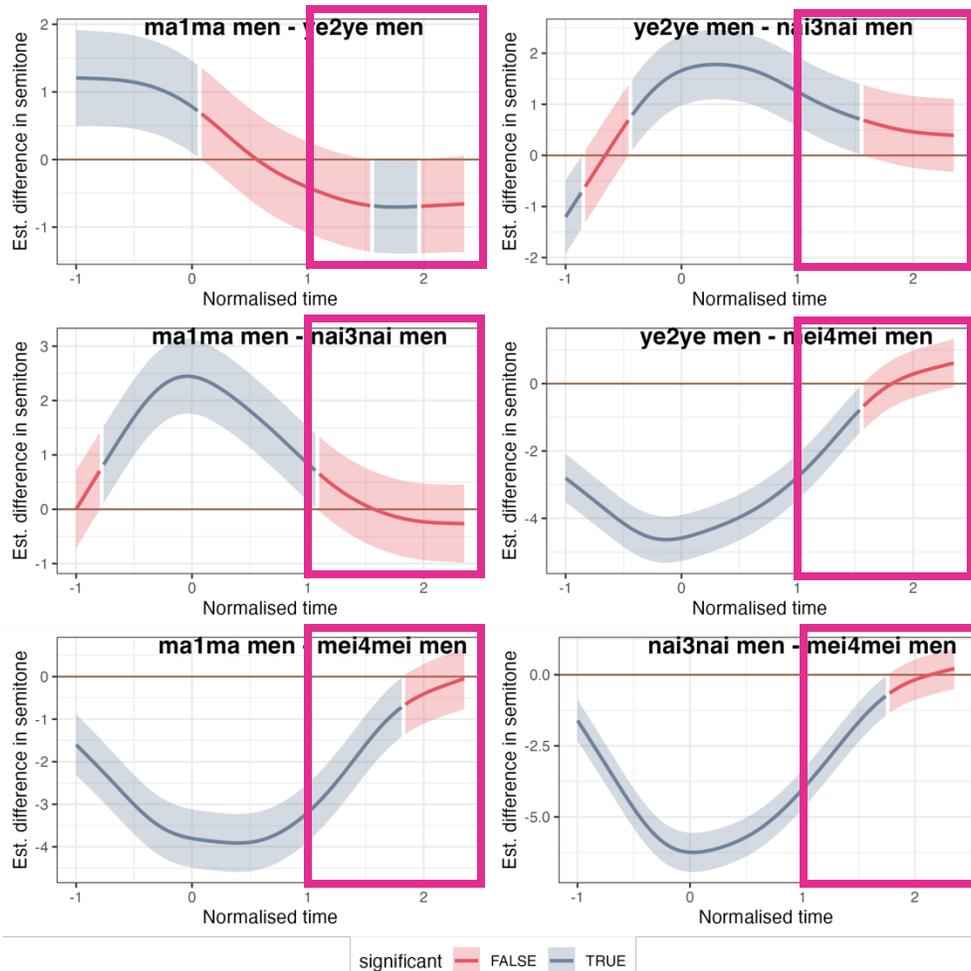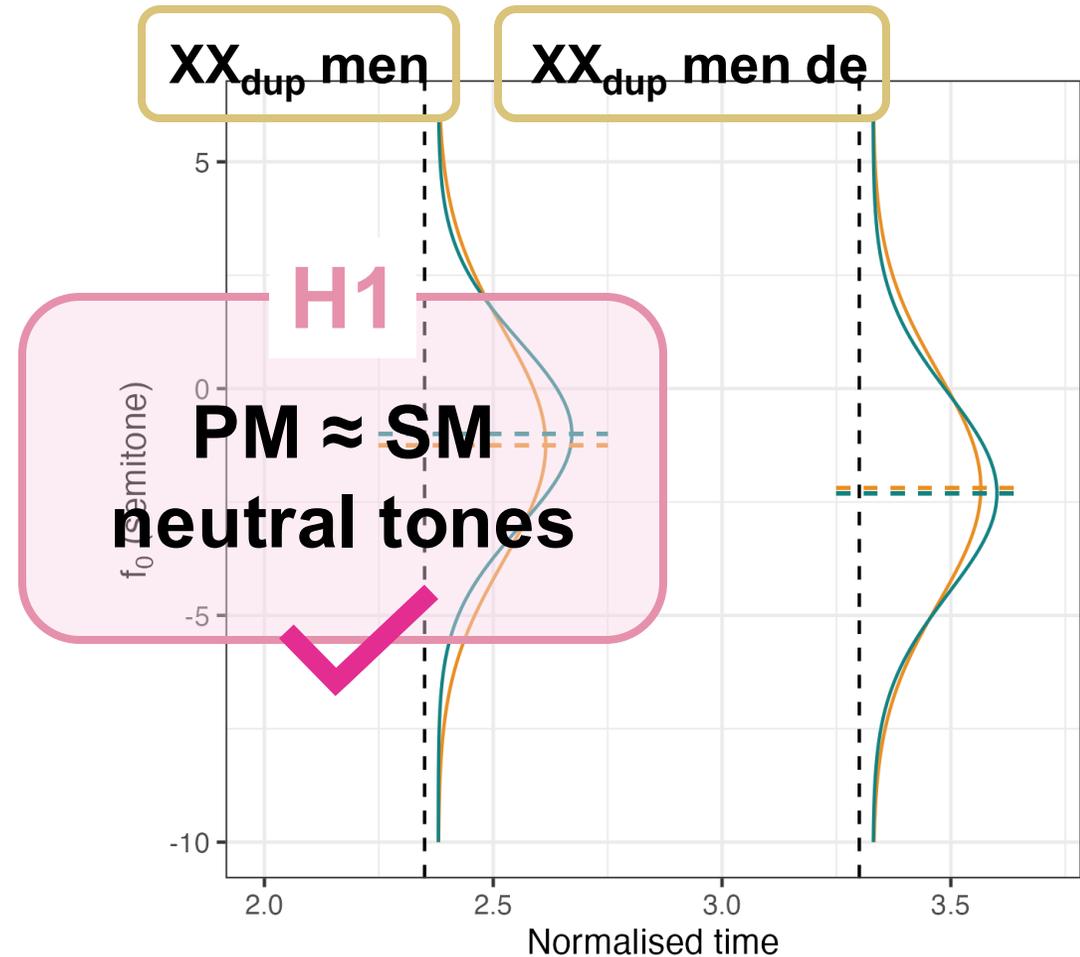
Plastic Mandarin

# Results (RQ2): Neutral Tone Comparison

*How do neutral tone patterns in Plastic Mandarin compare to those in Standard Mandarin?*
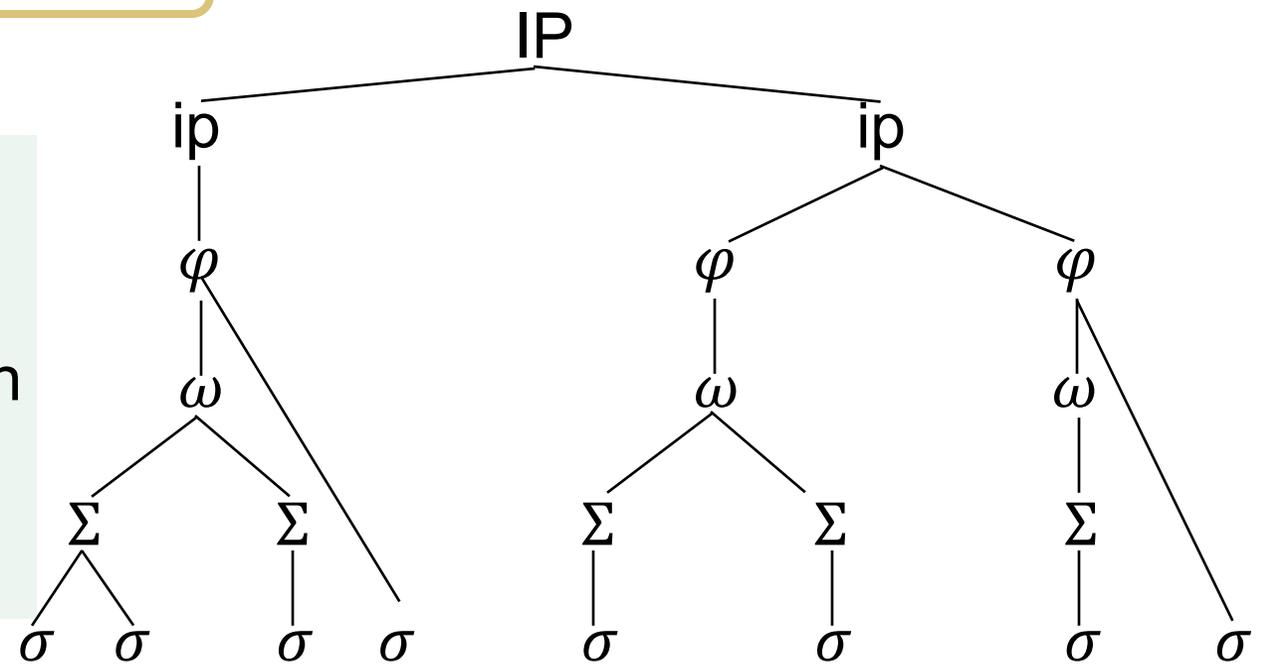


Plastic Mandarin



XX$_{dup}$ men

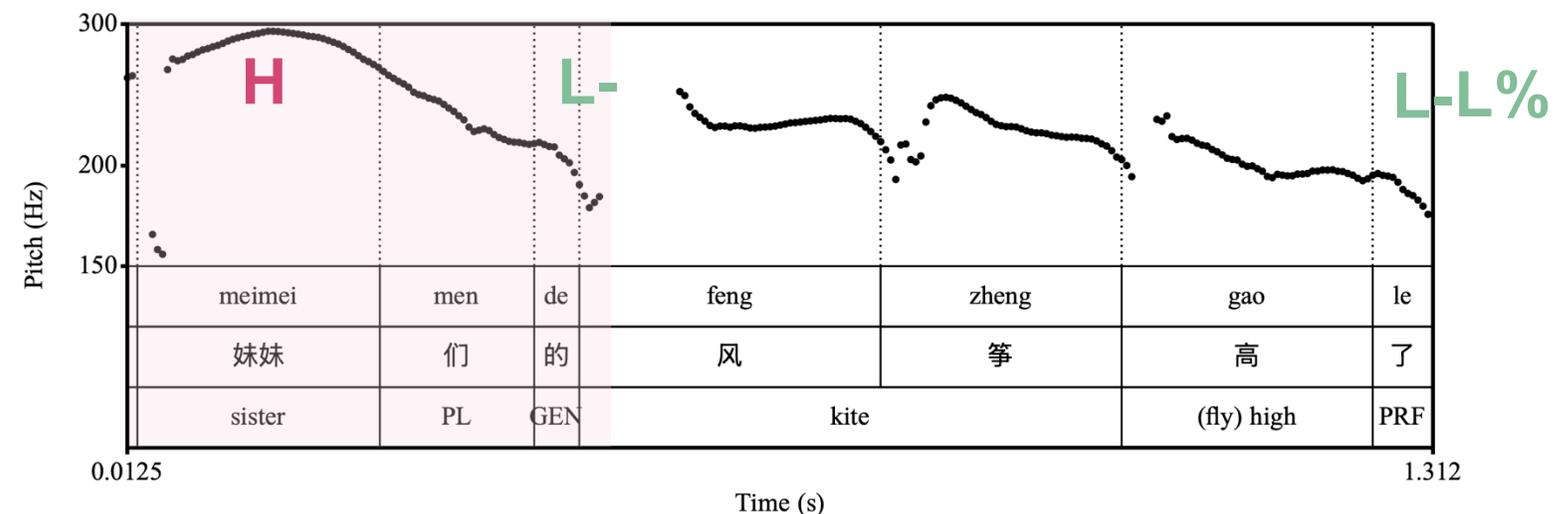XX$_{dup}$ men de

H1

PM ≈ SM neutral tones

Standard Mandarin    Plastic Mandarin

**mei4mei men de**

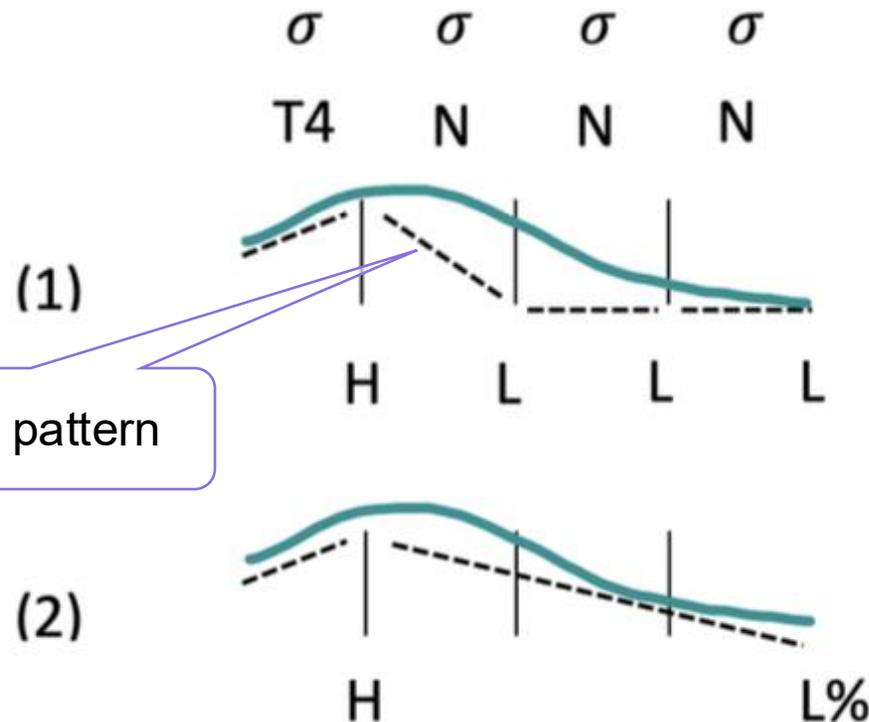The location of a neutral tone syllable tends to coincide with the **right edge** of a prosodic constituent

IP

ip                                    ip

φ                         φ              φ

ω                      ω              ω

Σ          Σ        Σ        Σ        Σ

σ   σ    σ    σ    σ       σ      σ    σ

Intonational Phrase

intermediate phrase

Phonological phrase

Phonological word

Foot

Syllable

H          L-                              L-L%

| meimei | men | de | feng | zheng | gao | le |
|--------|-----|-----|------|-------|-----|-----|
| 妹妹 | 们 | 的 | 风 | 筝 | 高 | 了 |
| sister | PL | GEN | kite | | (fly) high | PRF |

Pitch (Hz)
300
200
150

0.0125                                          1.312
Time (s)

*Is there a pitch target for neutral tone?*

# 2. Discussion

## Standard-Plastic Mandarin variation

**What's changed?**

**Lexical tones (A)**:
**Systematic** tone variation

**What remains persistent?**

**Neutral tone (B)**:

Prosodically conditioned in a similar fashion
- **Low** pitch target at the end of a **sequence**
- Strength of effects of neighboring tones
- Asymmetric influence of preceding and following tones

Attracting a **boundary low tone** (L%)

**Cross-dialectal insight: Constancy at higher prosodic levels?**

**RoadMap**

# Tone development:
## *ManyTones* project

In this talk: **Perception of f0 perturbations across many languages**

## Aims

1. To enhance our knowledge on microprosodic pitch perception in speech

2. To create an online framework for large-scale auditory perceptual research

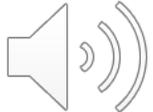# Consonant-related $f_0$ perturbations (CF0)

**English**

- Aspirating
- Non-tonal

**Ukrainian**

- True voicing
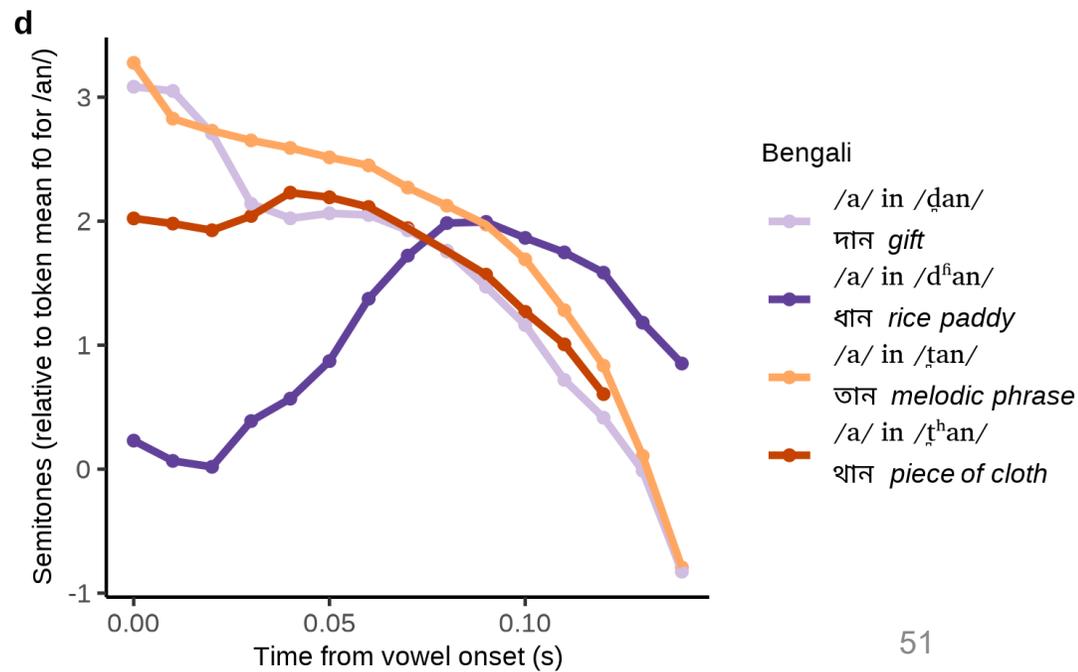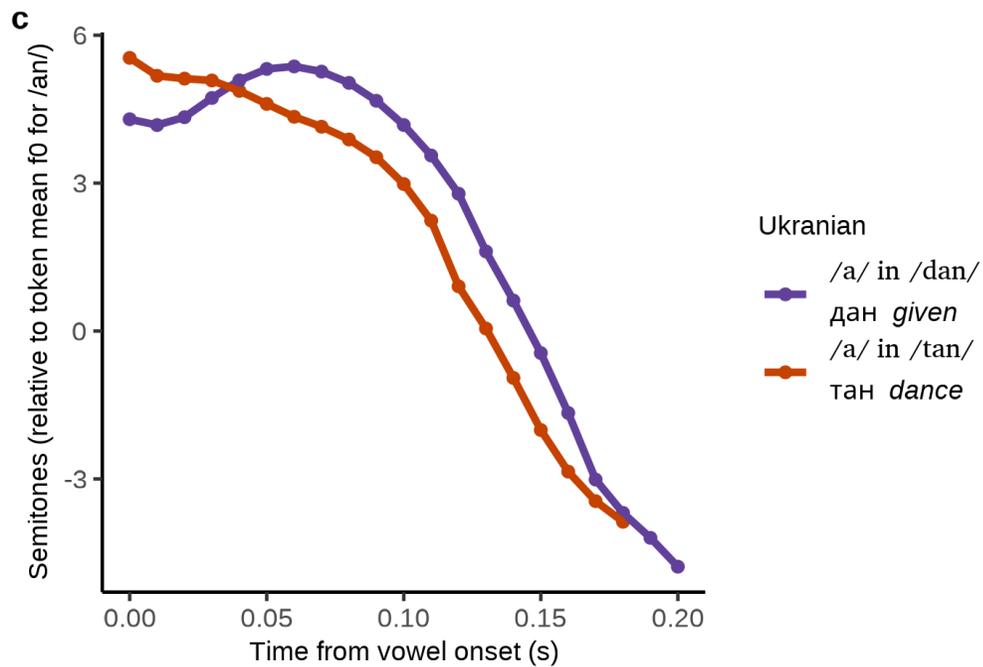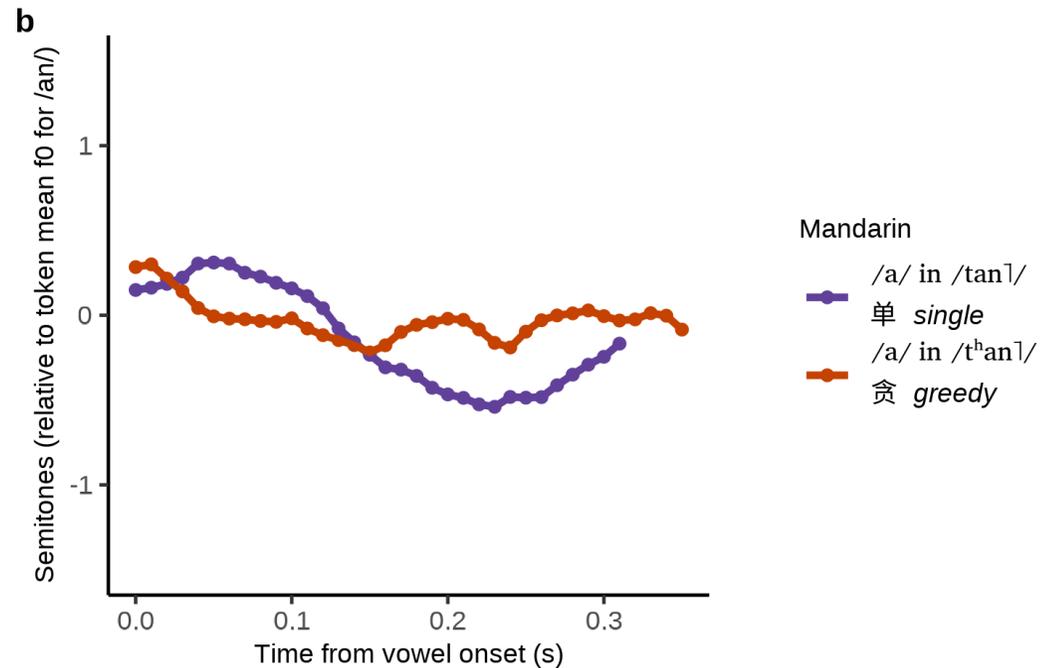- Non-tonal

/tan/   /dan/

**Mandarin**

- Aspirating
- Tonal

**Bengali**

- Four-way contrast
- Non-tonal

**CF0**

**a** English
- /a/ in /tan/ *Dan*
- /a/ in /tʰan/ *tan*

**b** Mandarin
- /a/ in /tan˥/ 单 *single*
- /a/ in /tʰan˥/ 贪 *greedy*

**c** Ukranian
- /a/ in /dan/ дан *given*
- /a/ in /tan/ тан *dance*

**d** Bengali
- /a/ in /d̪an/ দান *gift*
- /a/ in /dʱan/ ধান *rice paddy*
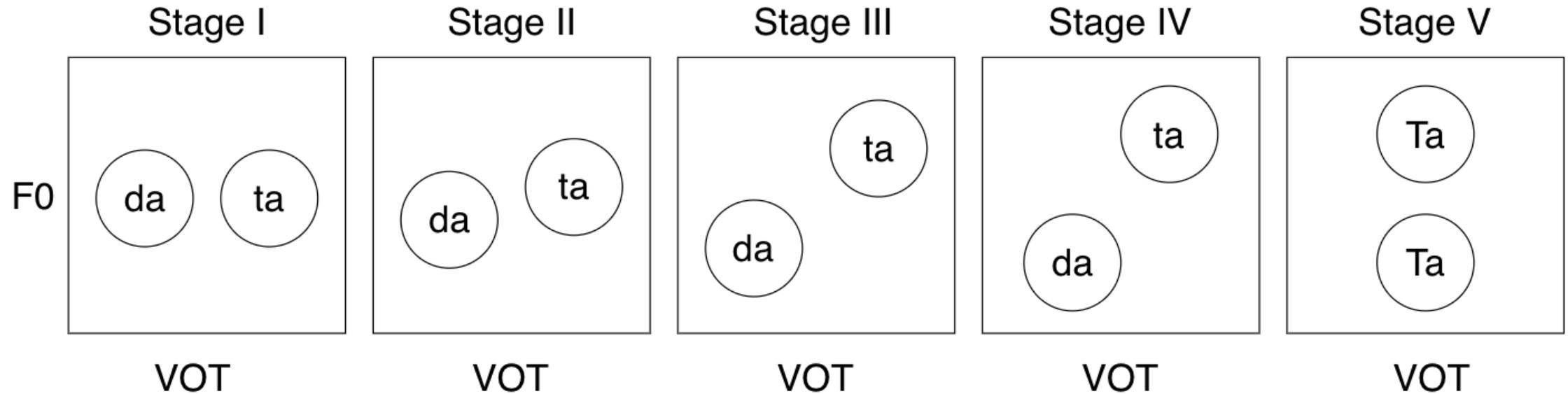- /a/ in /ʈan/ তান *melodic phrase*
- /a/ in /ʈʰan/ থান *piece of cloth*
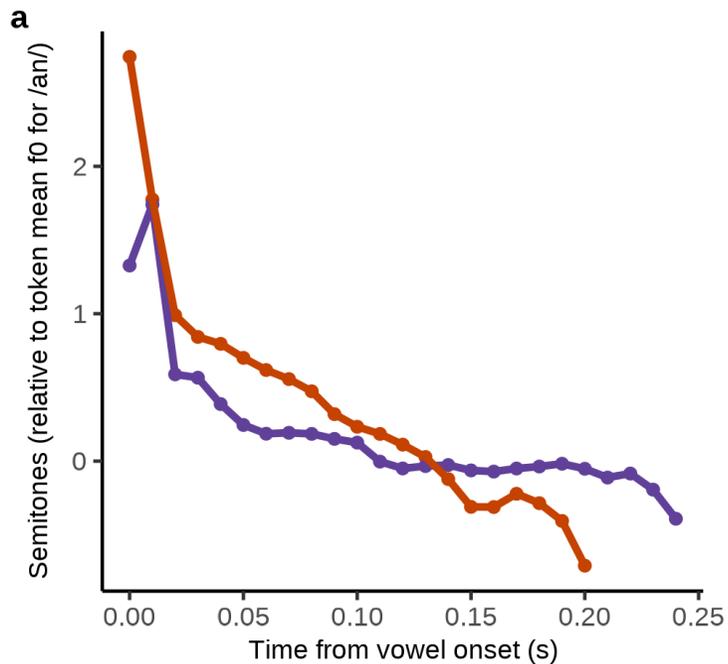
# CF0

- $f_0$, especially at vowel onset, is **higher following a voiceless** obstruent than following a voiced one.
- The **temporal extent** and **magnitude** of CF0 **vary** considerably.
  - Effect size: 0.4 to 3.9 semitones (Ting et al., 2025)
  - Duration*: 20 to 140 ms (various studies)
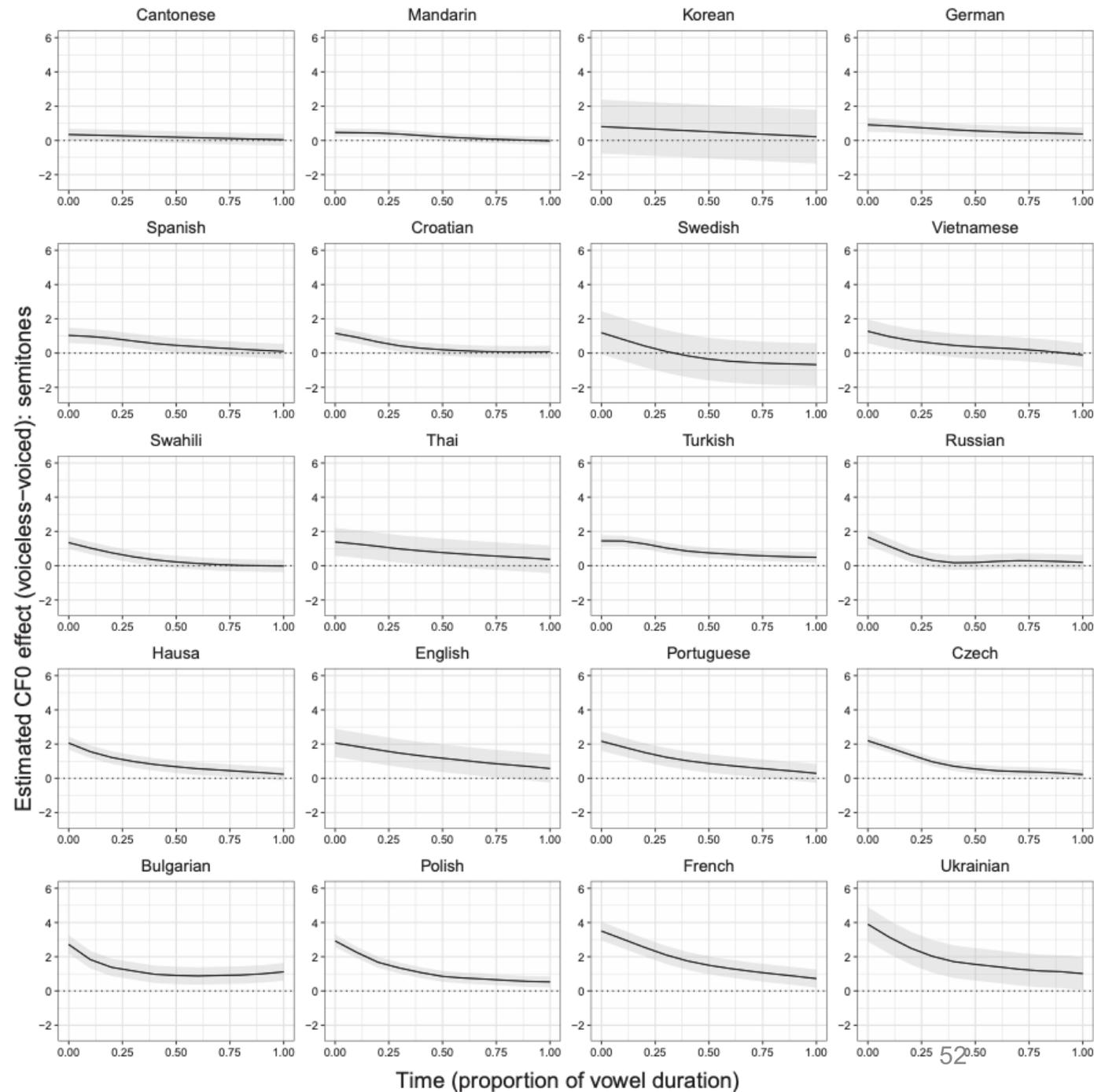
# The Development of contrastive tones



Five stages of tonogenesis based on Maran (1973). VOT = voice onset time.
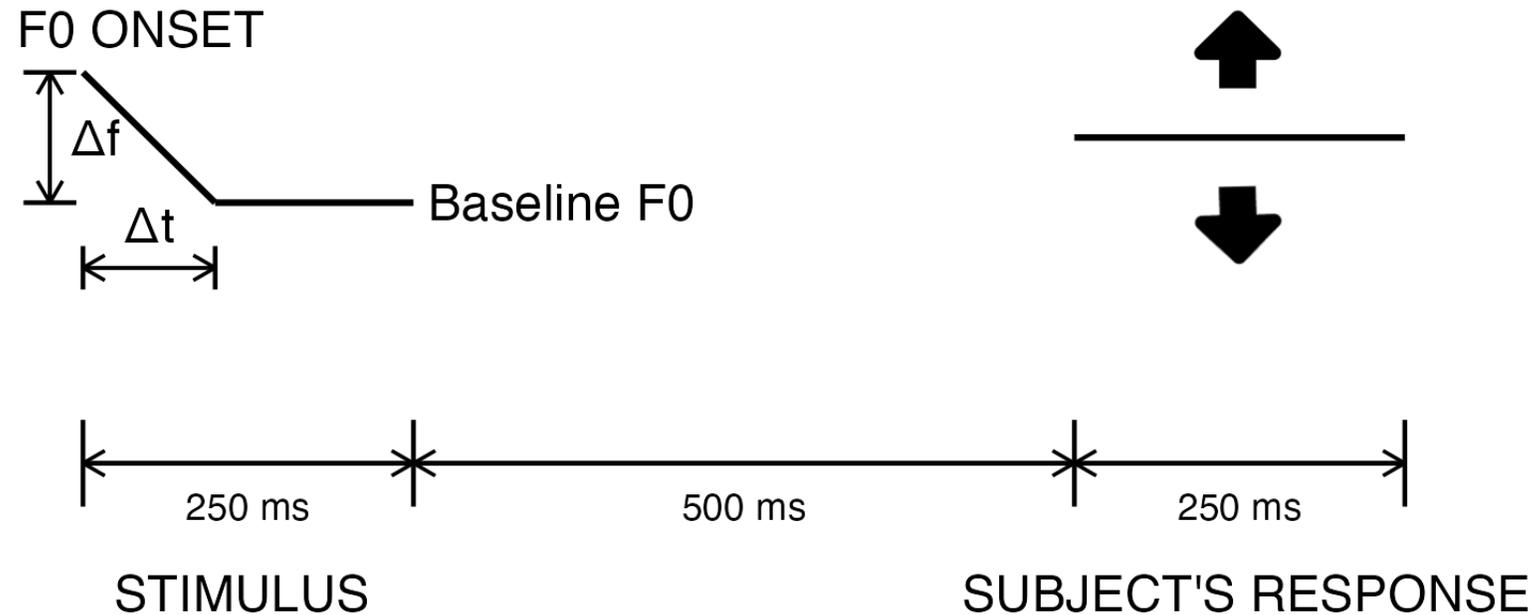Source: Kang, 2014

# Research Questions

**RQ1** To what extent can $f_0$ perturbations can be perceived depending on the **duration** and **extent** of the perturbation?

**RQ2** Does the perception of $f_0$ perturbations vary across listeners with different language experience? If so, how?
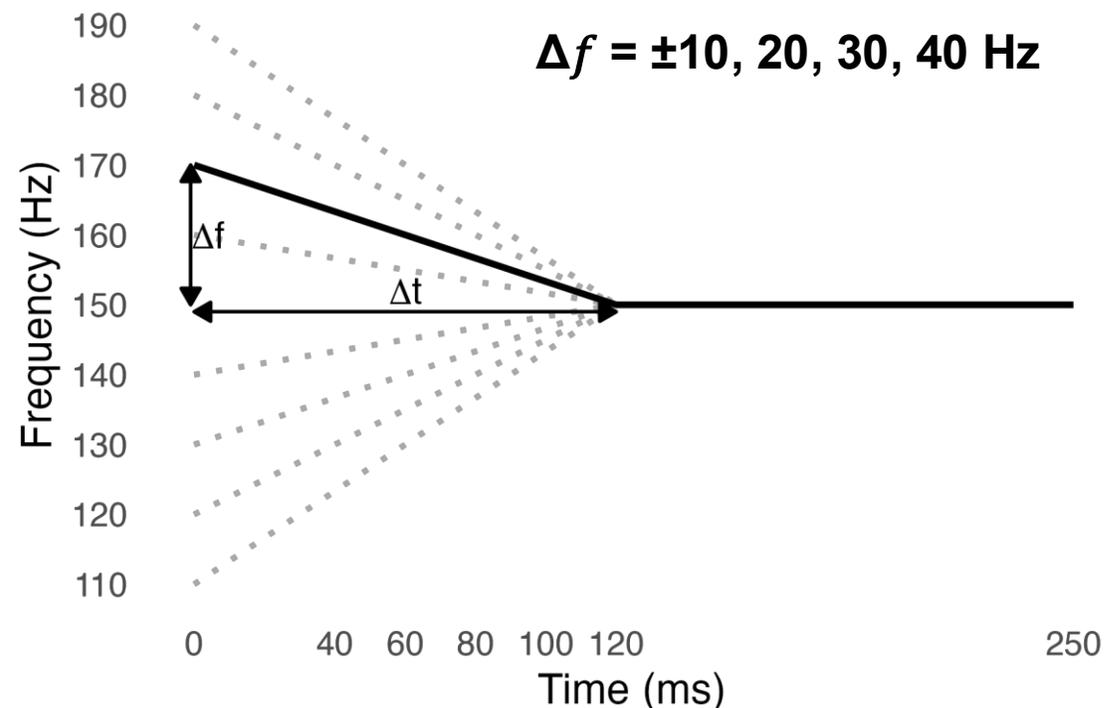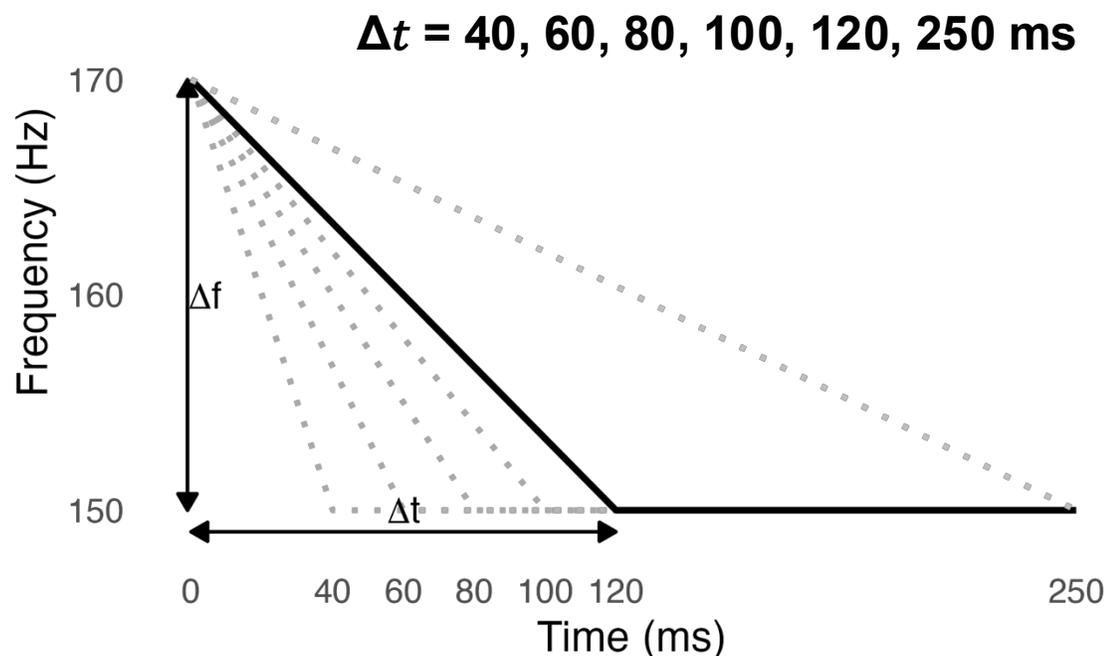
# Method: Experiment Paradigm

The Pitch-matching Paradigm



Hombert (1975, p.223 Part I)

# Method: Stimuli (Pilot)

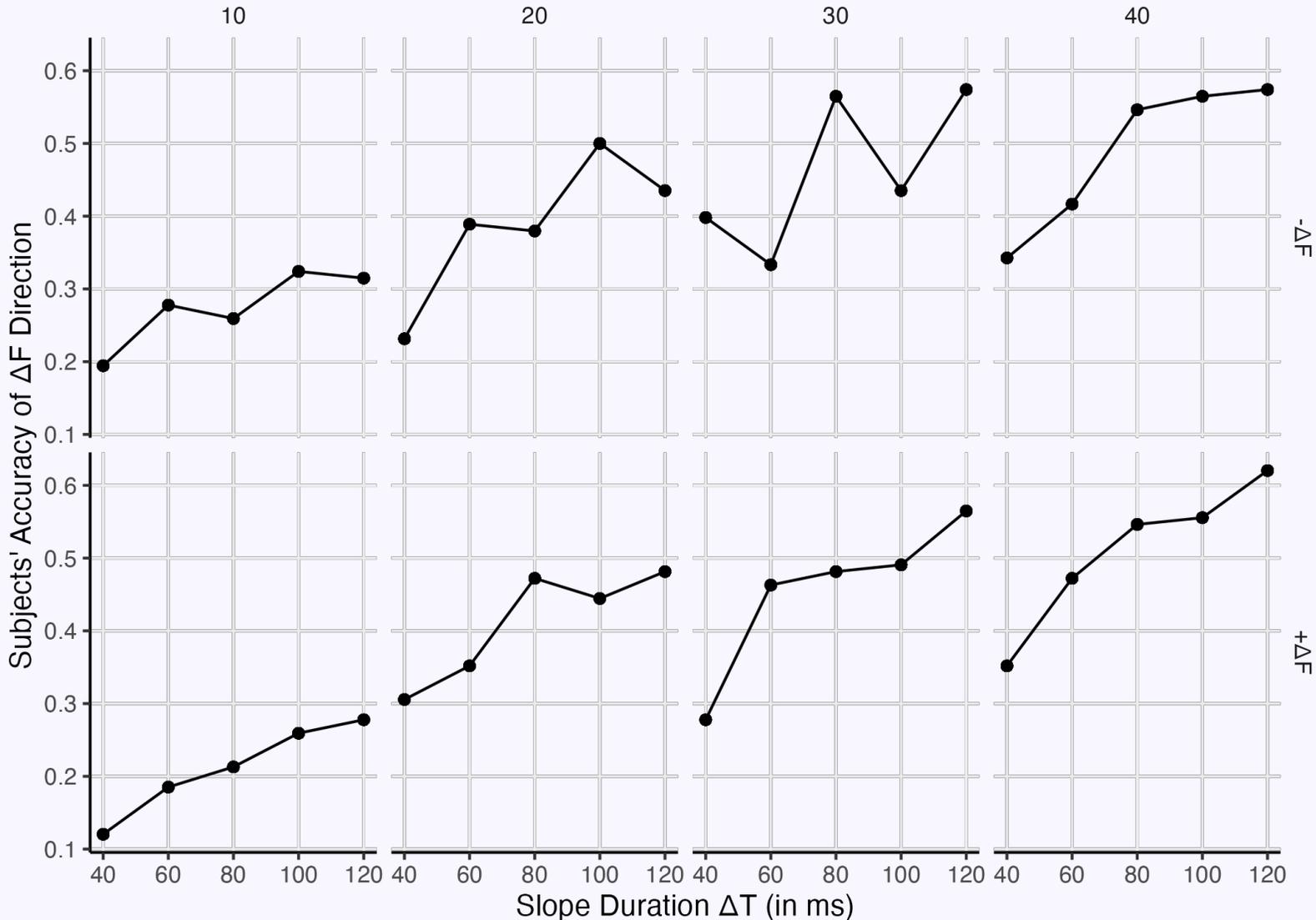**Δ*t* = 40, 60, 80, 100, 120, 250 ms**

**Δ*f* = ±10, 20, 30, 40 Hz**



- 3 sound token type: baseline $f_0$ = 150 Hz
  - Complex tone with 12 harmonics
  - Vowel [i:]
  - Syllable [ti:] (with short-lag VOT ≈12 ms)
- Resynthesised from a male recording (44.1 kHz, 16 bit, mono)
- Intensity normalised to 75 dB
- Fixed token length of 250 ms
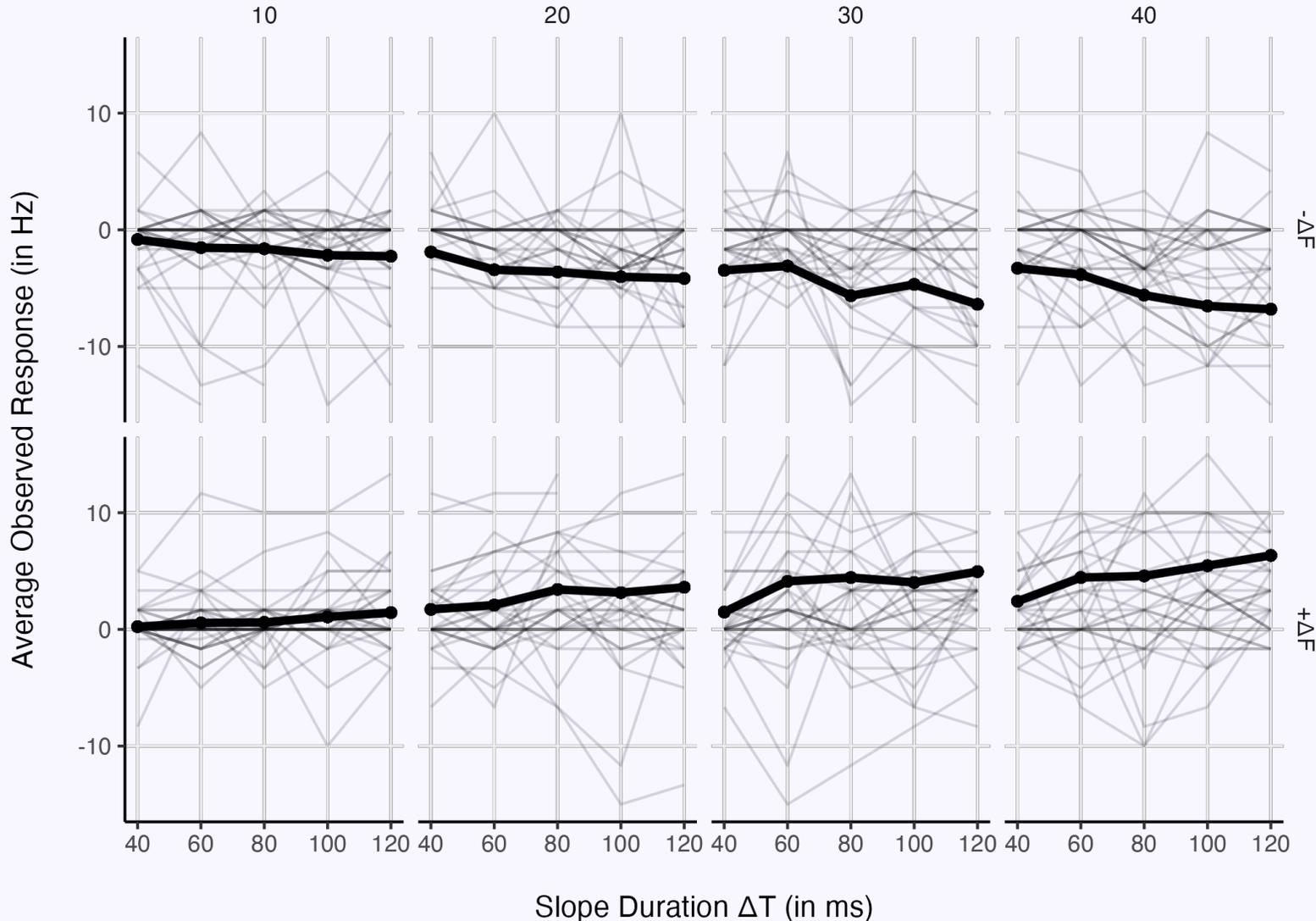
# Method: Procedure
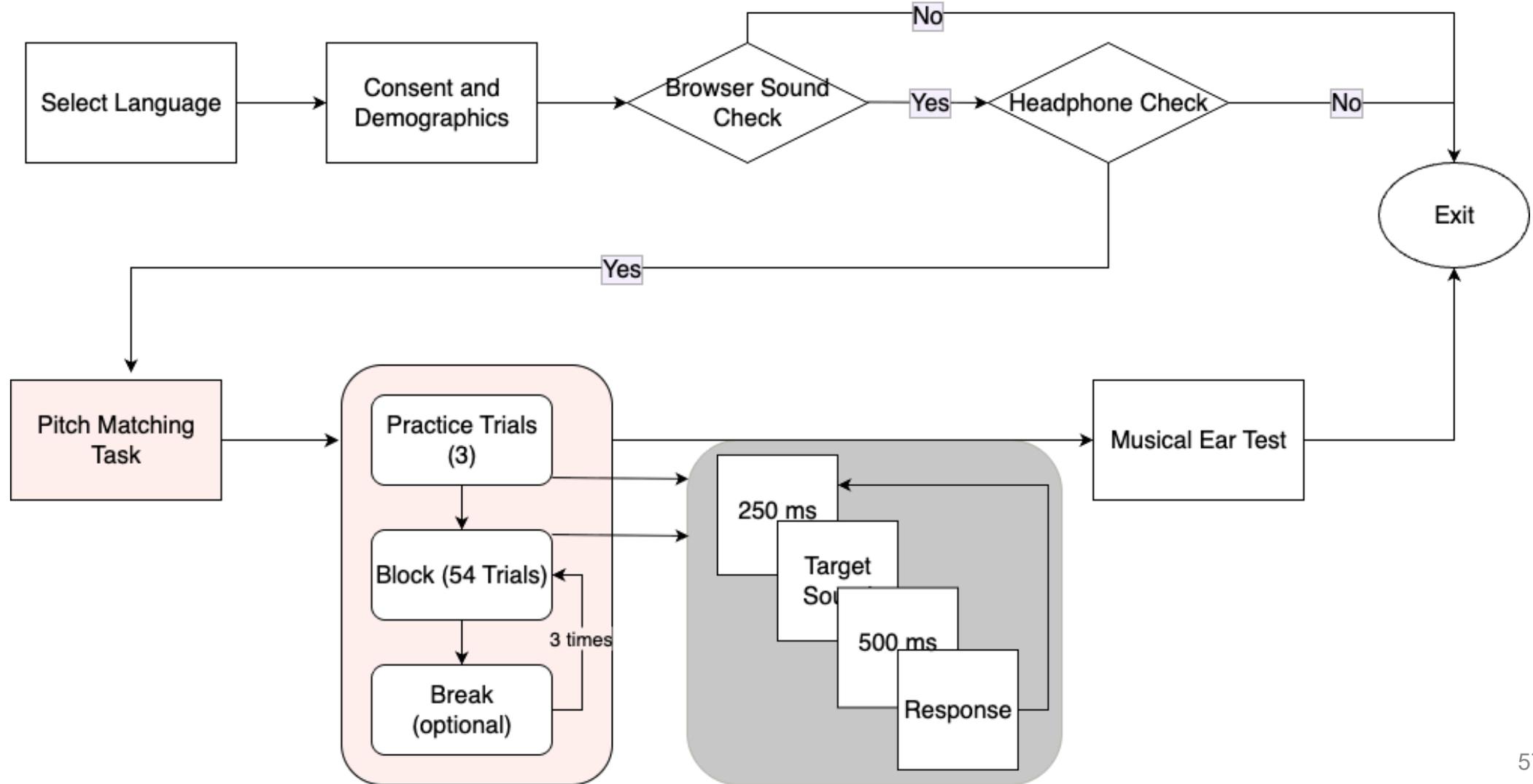
# Pilot results I



**Accuracy (A) of judging ΔF direction**

- Generally:
  $$A \propto (|\Delta F|, \Delta T)$$
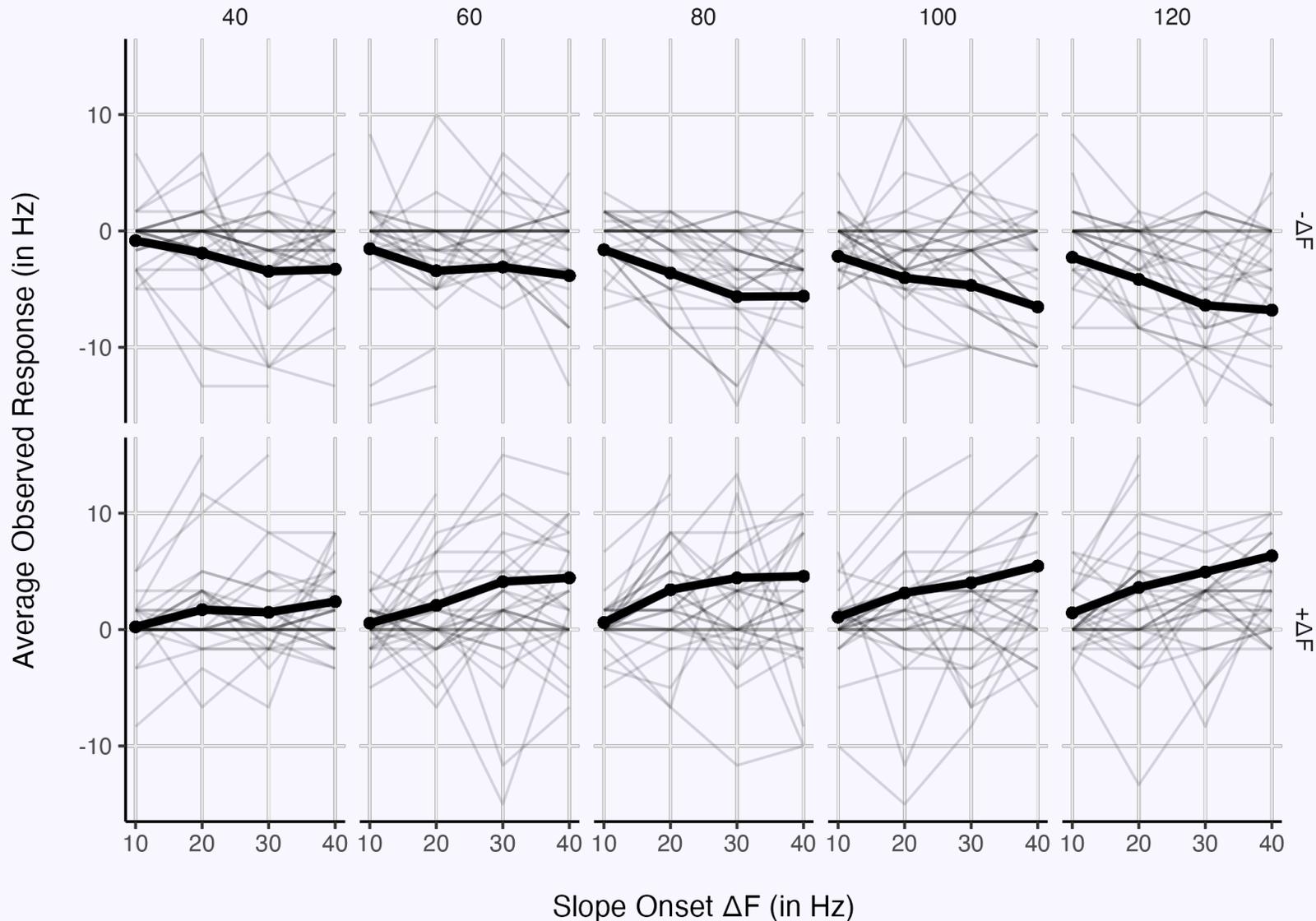- When |ΔF| =10Hz, accuracy is very low even with ΔF =120 ms.
- The accuracy increase is not linear

# Pilot results II



## Average perceived ΔF (R)

- The average perceived ΔF is relatively small |R|<10 Hz.
- In some cases (e.g. 20 ms), R reaches a plateau even with increased ΔT.
- There is considerable individual variation.

# Pilot results III



**Average perceived ΔF (R)**

- R is roughly linear with respect to |ΔF|, (only) when ΔT is large.

$$\frac{\partial^2 R}{\partial(\Delta F)^2} \to 0$$

# Big Team Science



Erin M. Buchanan

**Data Lead**

Timo B. Roettger

**Analysis Lead**

Chenzi Xu

**Method Lead**

**Scientific Lead**

Xinbing Luo

**Project Manager**
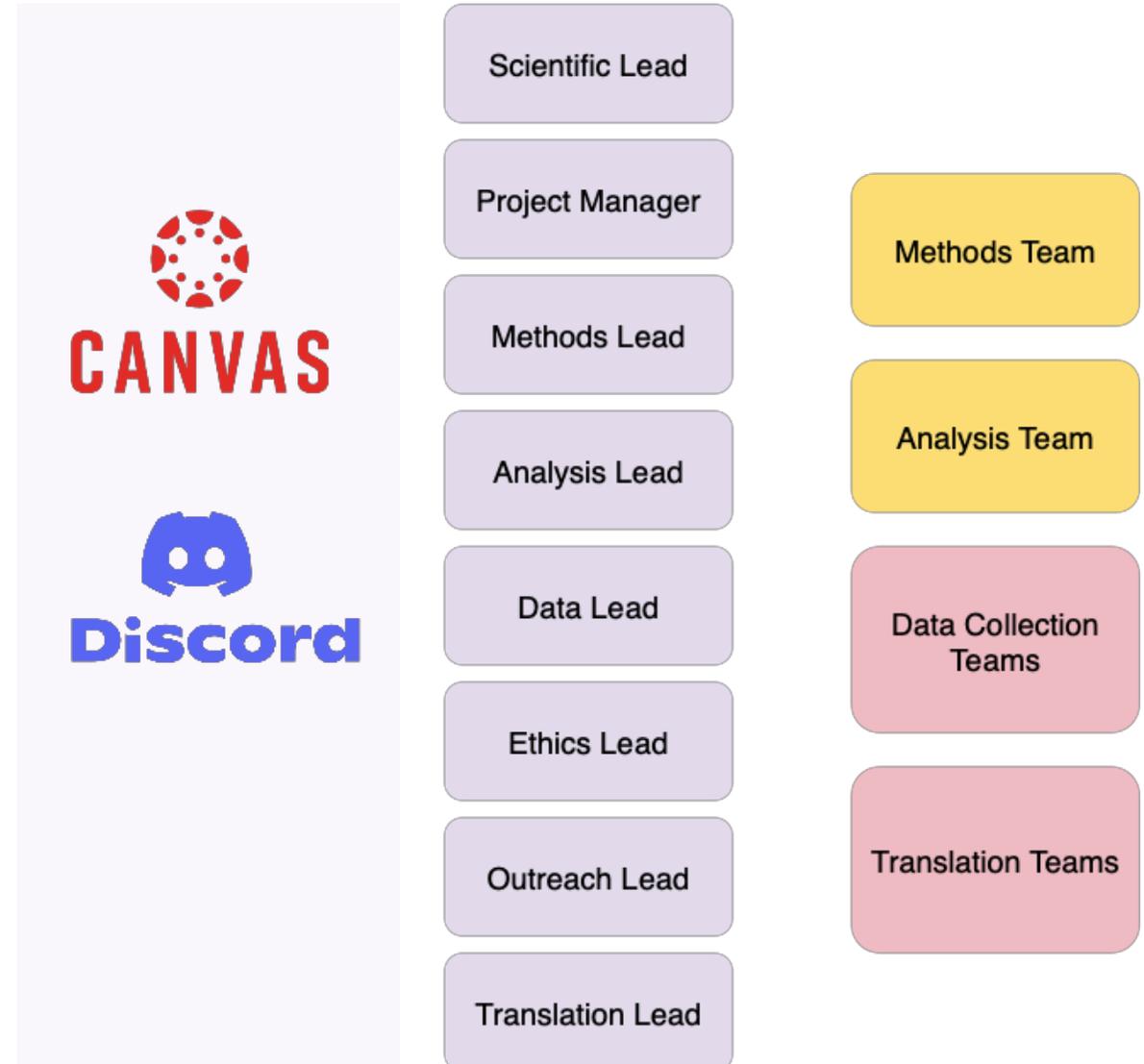
Cong Zhang

**Ethics Lead**

Indranil Dutta

**Outreach Lead**

# Big Team Science

## Join as a collaborator

- Open to research labs, fieldworkers, linguists, cognitive scientists, and musicians
- Collaborate in data collection and analysis
- Co-author in high-impact, multi-author publications



Scientific Lead

Project Manager

Methods Lead

Analysis Lead

Data Lead

Ethics Lead

Outreach Lead

Translation Lead

Methods Team

Analysis Team

Data Collection Teams

Translation Teams

# Take-home Messages

- Tones vary and change!

- Neutral tone in Mandarin may be underlyingly **underspecified** and acquire a target via **prosodic structure**

- Tone features at higher prosodic hierarchy are more **pertinacious** (resistant to change)

- Perspectives of language **variation and change** help us understand tonal phenomena

- The number of tone targets in tone languages is **fewer** than the number of syllables in connected speech

# Thank you!

🏠 https://chenzixu.rbind.io

✉️ chenzi.xu@ling-phil.ox.ac.uk