ALS 2020 Building bridges

14-15 December 2020
(Sydney time)

One million km^2: Broad coverage of Mandarin Chinese dialects using smartphone technology

Liang Zhao, Philip Harrison, Paul Foulkes and Eleanor Chodroff

University of York

December 2, 2020

Goals of this study:

We propose to test a remote audio collection method using free smartphone recording applications controlled by participants on six mainstream Mandarin dialects. Our primary goals are:

- To provide a qualitative analysis of at-home smartphone recording quality
- To examine the suitability of recordings for analysis of *relative* acoustic-phonetic patterns in
  - F1–F2 vowel space
  - Tone system
- To provide a preliminary set of recommendations for remote audio data collection based on our experiences

Overview

1. Introduction
2. Methods
   - Participants
   - Tasks and stimuli
   - Procedures
3. Results and preliminary analysis
   - Recording quality
   - Vowel space
   - F0 measurements—tone categories of Mandarin dialects
4. Conclusion

1. Introduction

	lab & field data	remote audio data
EXPERIMENTAL CONTROL	+++	+
COST	+++++	+
TIME	+++++	+
REACH	+	++++
ENVIRONMENT	+++++	+
COVID	+	++++
…		

1. Introduction

Studies on potential sources of variability in remote audio collection:

- Conflicting findings in the effects of **recording devices** on acoustic measurements including f0 values and vowel formants (De Decker & Nycz 2011, Grillo et al. 2016)

- Such a difference in findings between 2011 and 2016 might be attributed to technical advances and improved quality of the built-in recorder over time

- Also, device-derived variability may be dependent on the **acoustic parameter** to be investigated (Jannetts et al., 2019)

1. Introduction

Uncertainty in the resulting absolute values of specific speech sounds or phonetic categories

Can we assume an approximately equal impact of device and environment noise on recording quality, such that the relative patterns of acoustic-phonetic realizations are still reasonably trustworthy?

- Are the absolute values within an expected range and do the patterns conform with previous findings?
  - F1–F2 vowel space
  - Tone systems

1. Introduction

Goals of this study:

We propose to test a remote audio collection method using free smartphone recording applications controlled by participants on six mainstream Mandarin dialects.

Our primary goals are:

- To provide a qualitative analysis of at-home smartphone recording quality
- To examine the suitability of recordings for analysis of *relative* acoustic-phonetic patterns in
  - F1–F2 vowel space
  - Tone system
- To provide a preliminary set of recommendations for remote audio data collection based on our experiences

Overview

1. Introduction
2. Methods
   - Participants
   - Tasks and stimuli
   - Procedures
3. Results and preliminary analysis
   - Recording quality
   - Vowel space
   - F0 measurements—tone categories of Mandarin dialects
4. Conclusion

2. Methods

Six Mandarin dialects (Mandarin-group dialects)

Dialect code	City	Province
BEI	Beijing	
TYN	Taiyuan	Shanxi
JNN	Jinan	Shandong
XIA	Xi’an	Shaanxi
CHD	Chengdu	Sichuan
WHN	Wuhan	Hubei

2. Methods

Participants

N = 36, aged 18-40 (mean=28.29)

	Beijing	Taiyuan	Jinan	Xi'an	Chengdu	Wuhan	total
M	3	3	2	2	1	2	13
F	6	4	3	2	4	4	23
total	9	7	5	4	5	6	**36**

2. Methods

Stimuli (reading material) & Tasks

- Presented through an online experiment builder—***Gorilla***
- Participants received a web link to the whole experiment after agreeing to participate
- Participants carried out recording process without side-by-side supervision from the researcher

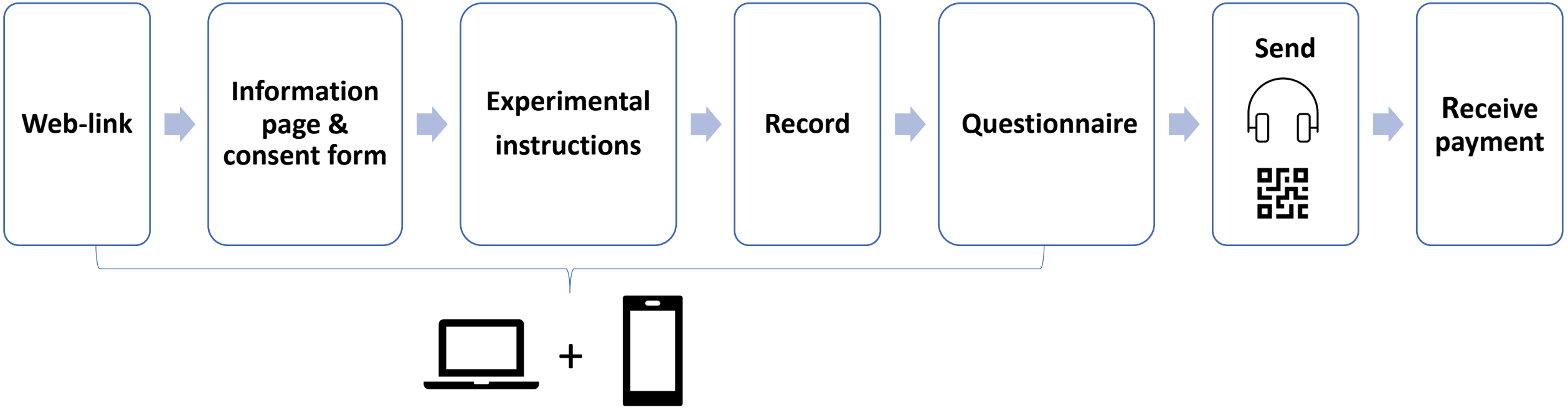
2. Methods

Stimuli (reading material) & Tasks

- Each task completed first in standard Mandarin and then in participant's regional dialect
- Word lists were counterbalanced for segmental structure, tone and position (for disyllabic words)

2. Methods

Procedures for participants

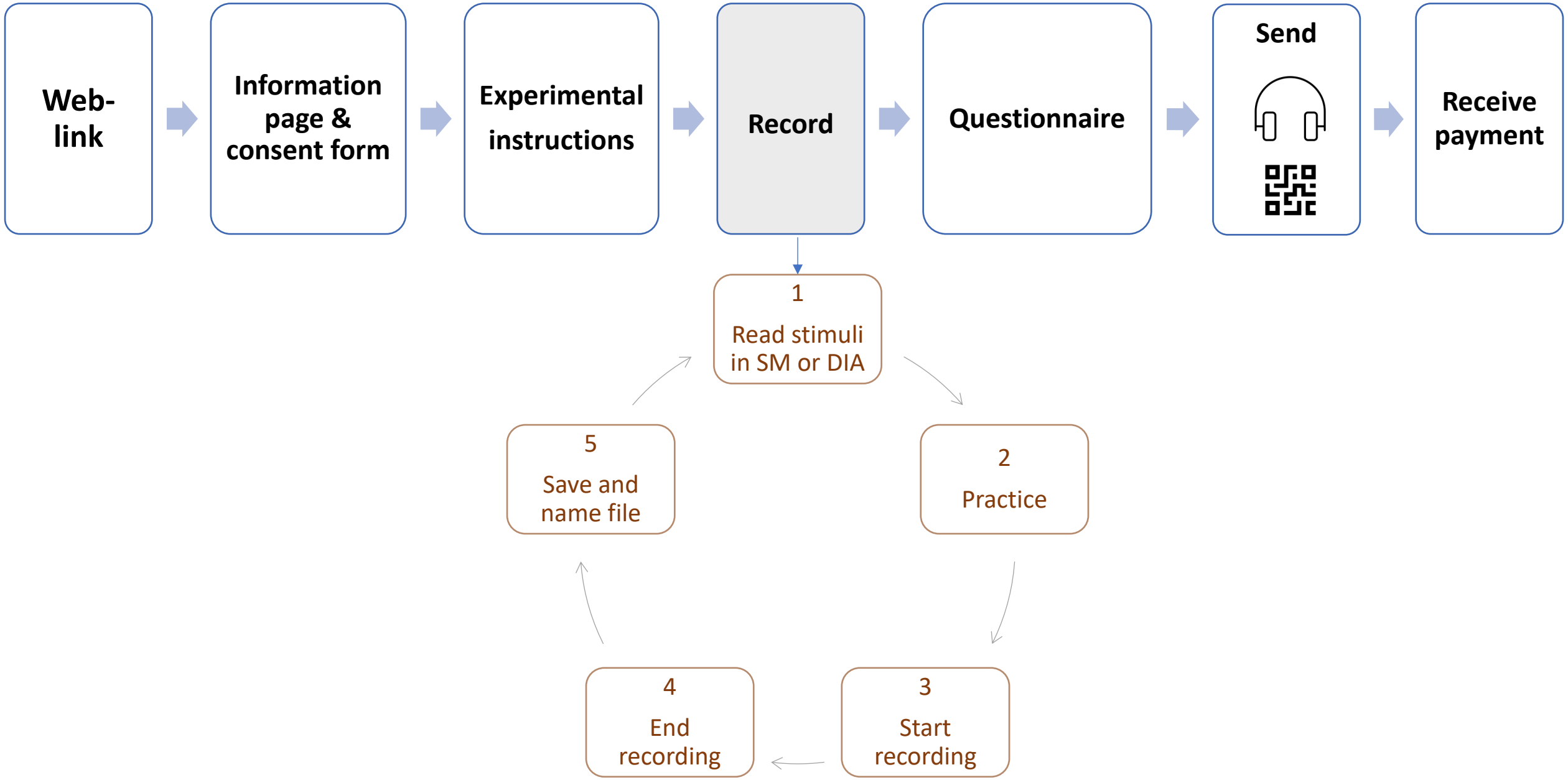


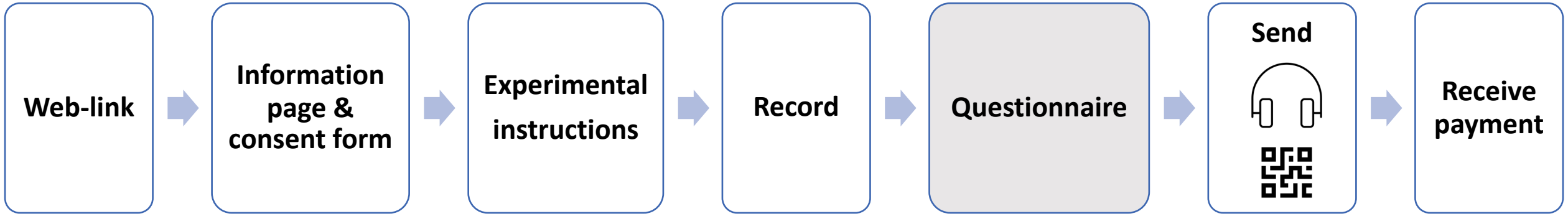
(Duration: 20 to 25 minutes)

Web-link → **Information page & consent form** → **Experimental instructions** → **Record** → **Questionnaire** → **Send** → **Receive payment**

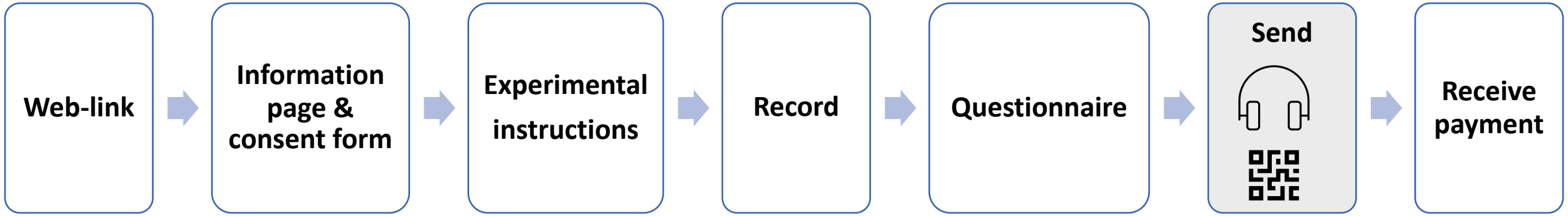
Experimental instructions

- **Prepare recording environment**
  - Close doors, windows, curtains
  - Room with soft furnishing
- **Prepare devices**
- **Install recording apps**
  - iPhone: Awesome Voice Recorder by Newkline Co., Ltd.
  - Android: ASR Recorder by NLL APPS.
- **Set app specifications**
  - File Format: WAV
  - Encode quality: Medium
  - Sample rate: 44100Hz
  - Channels: Stereo
- **Generate experiment ID**


Web-link
Information page & consent form
Experimental instructions
Record
Questionnaire
Send
Receive payment
1
Read stimuli in SM or DIA
2
Practice
3
Start recording
4
End recording
5
Save and name file



- Participants were given a 5-point Likert scale to rate aspects of the experiment like clarity, difficulty, and duration (20 to 25 min)
- All responses so far have been in the positive range (3 to 5)



In this step, participants sent to the researcher through email or upload to a cloud drive folder:

- Ten audio files
- WeChat payment QR code
- Unique experiment ID

Overview

1. Introduction
2. Methods
   - Participants
   - Tasks and stimuli
   - Procedures
3. Results and preliminary analysis
   - Recording quality
   - Vowel space
   - F0 measurements—tone categories of Mandarin dialects
4. Conclusion

3. Results and preliminary analysis

Recording quality

- Good but not perfect quality of the recordings
- Various acoustic conditions and background noise
  - Echoey room----prefer room with soft furnishing
  - Mouse clicking overlapped with speech----insert an inter-stimulus interval of 600ms
  - Phone ringtone/alert
  - Background noise from cricket chirping
  - Clock ticking
- Unwanted noise, errors, repetitions from participants

3. Results and preliminary analysis

Goals of this sections:

- To provide a qualitative review on the overall recording quality

- To what extent are the collected data suitable for acoustic analysis?
  - Vowel space: F1 and F2
  - Tone categories: f0 measurements

Acoustic analysis

Alignment

- Forced alignment using Montreal Forced Aligner
- Manual adjustment of vowel boundaries

Vowel space: F1 and F2

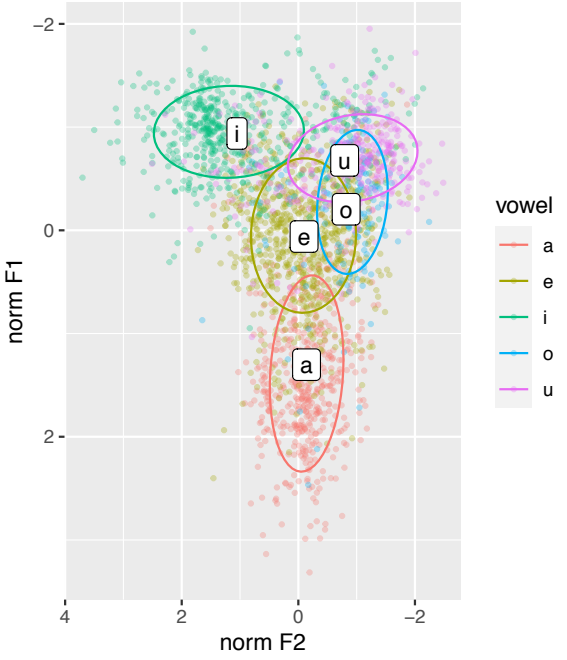
- Midpoint F1 and F2 from /i ə a o u/ measured automatically in Hz using Praat
- Removed tokens 2.5 standard deviations away from talker- and vowel-specific means
- Lobanov normalization

Tone categories: f0 contours

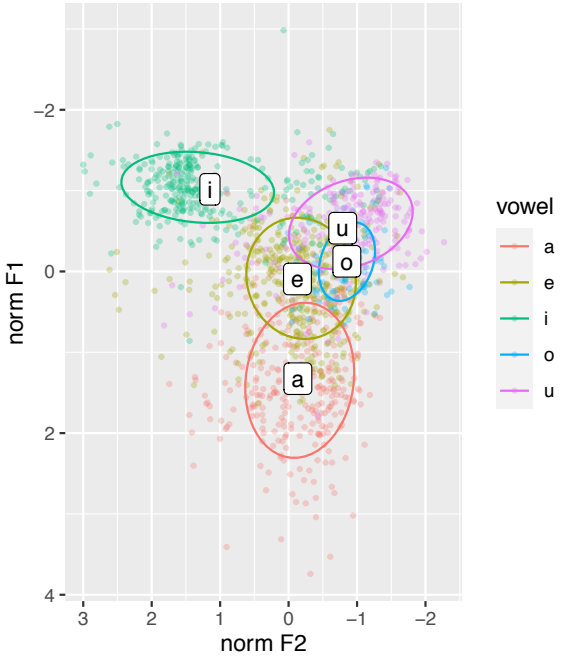
- 10 equally spaced f0 values across all vocalic intervals
- Converted to semitones

- Vowel space of six dialects based on F1 & F2 from **all the tasks**

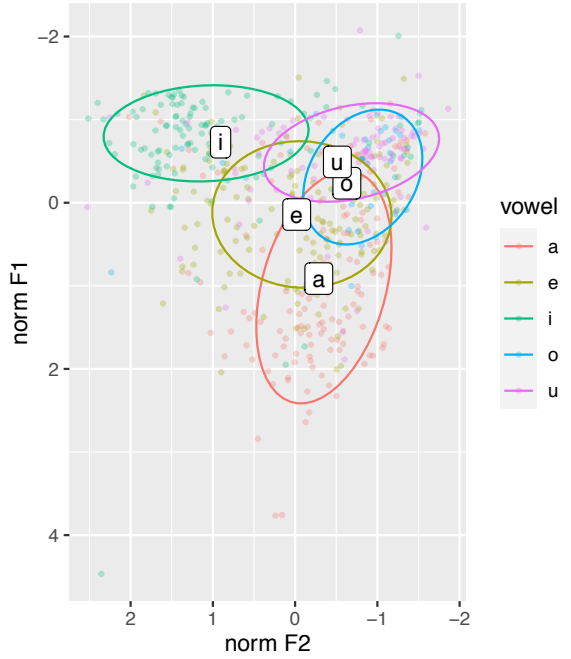
Beijing



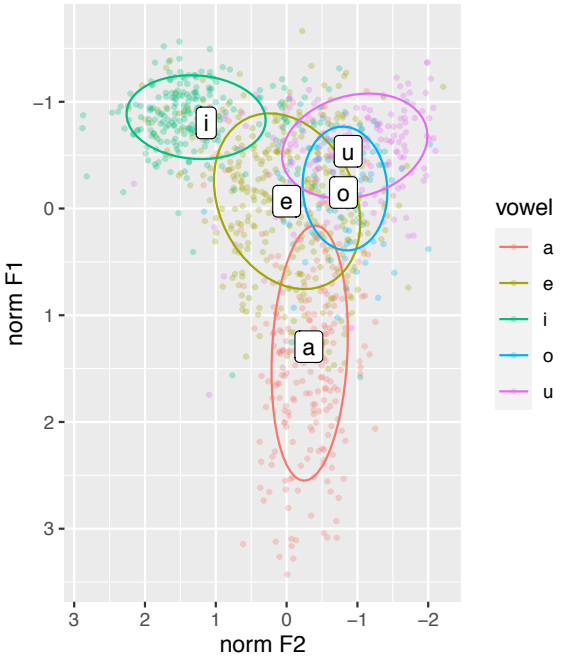
Taiyuan



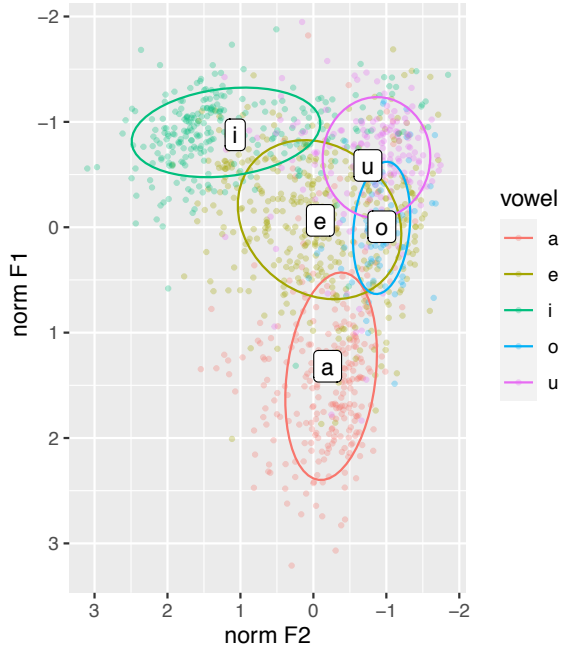
Jinan



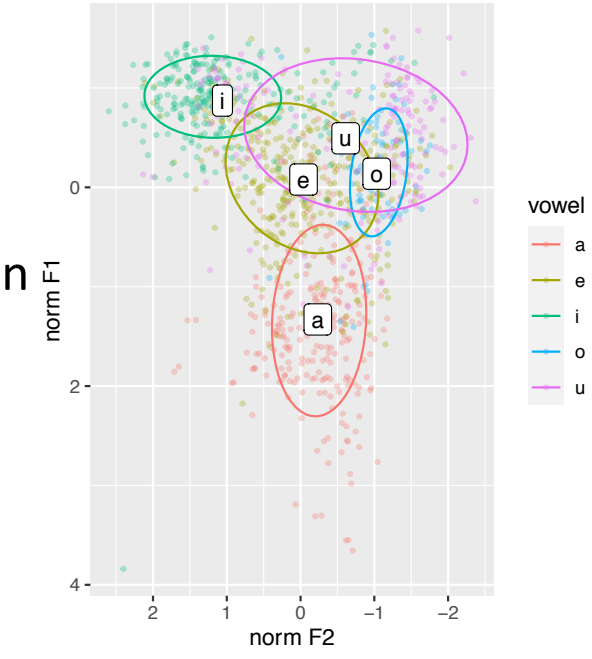
Xi'an



Chengdu

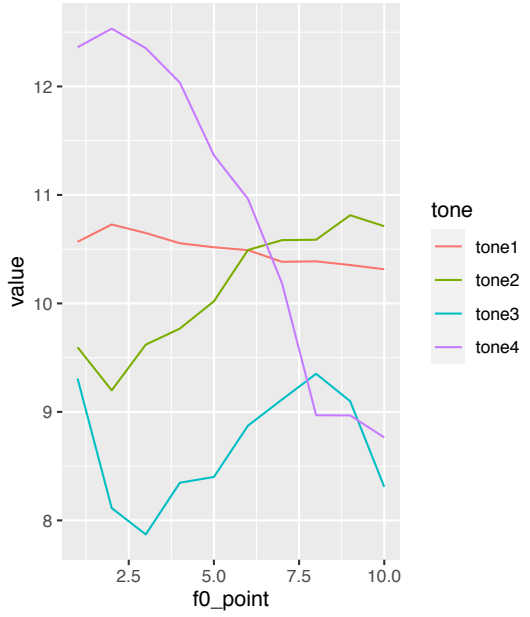


Wuhan

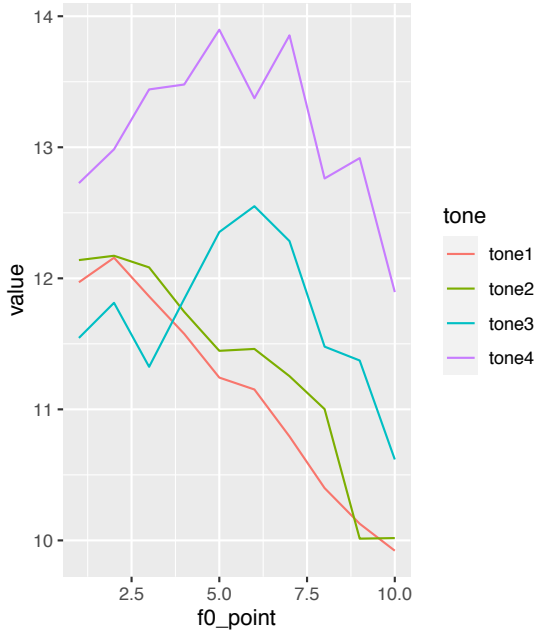


- Tone categories of six dialects based on mean f0 from **WL1** (monosyllabic characters)

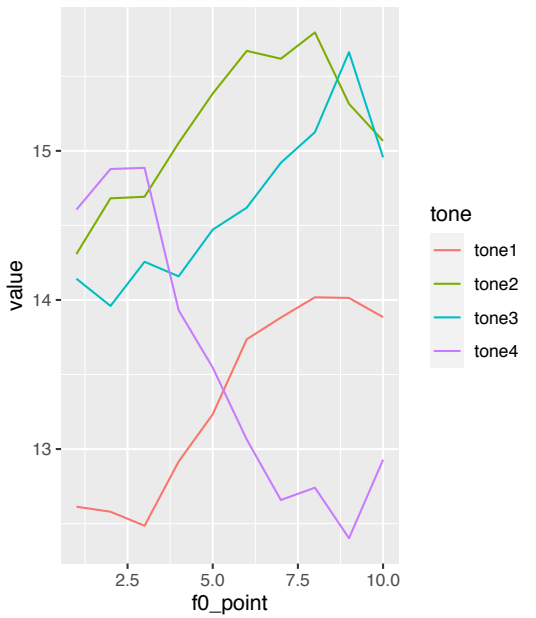
Beijing



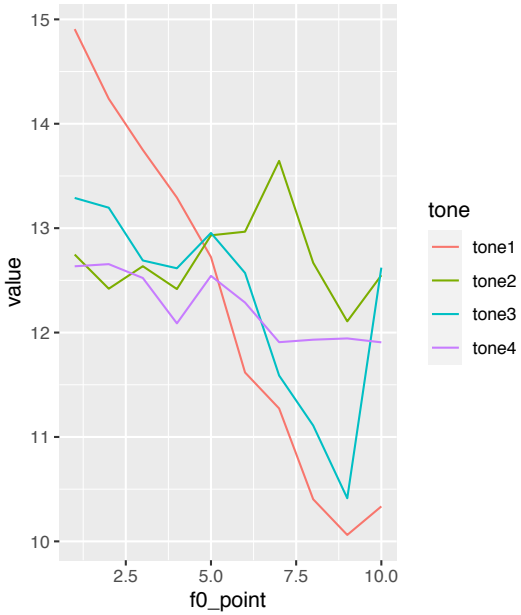
Taiyuan



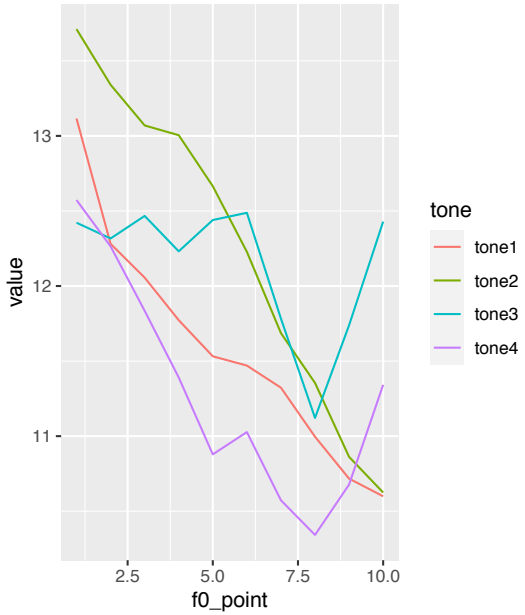
Jinan



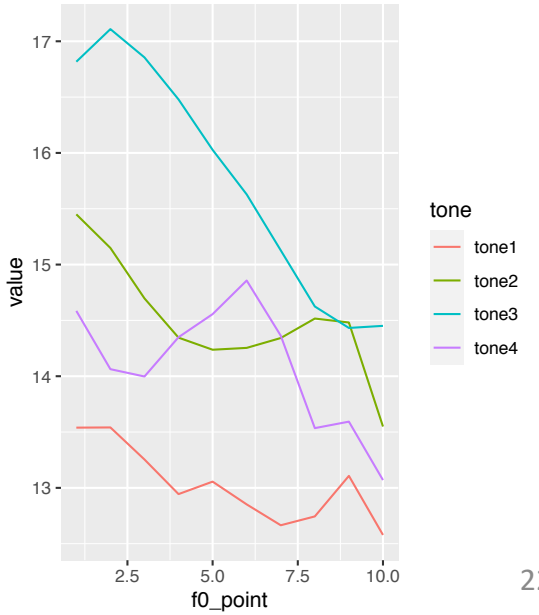
Xi'an



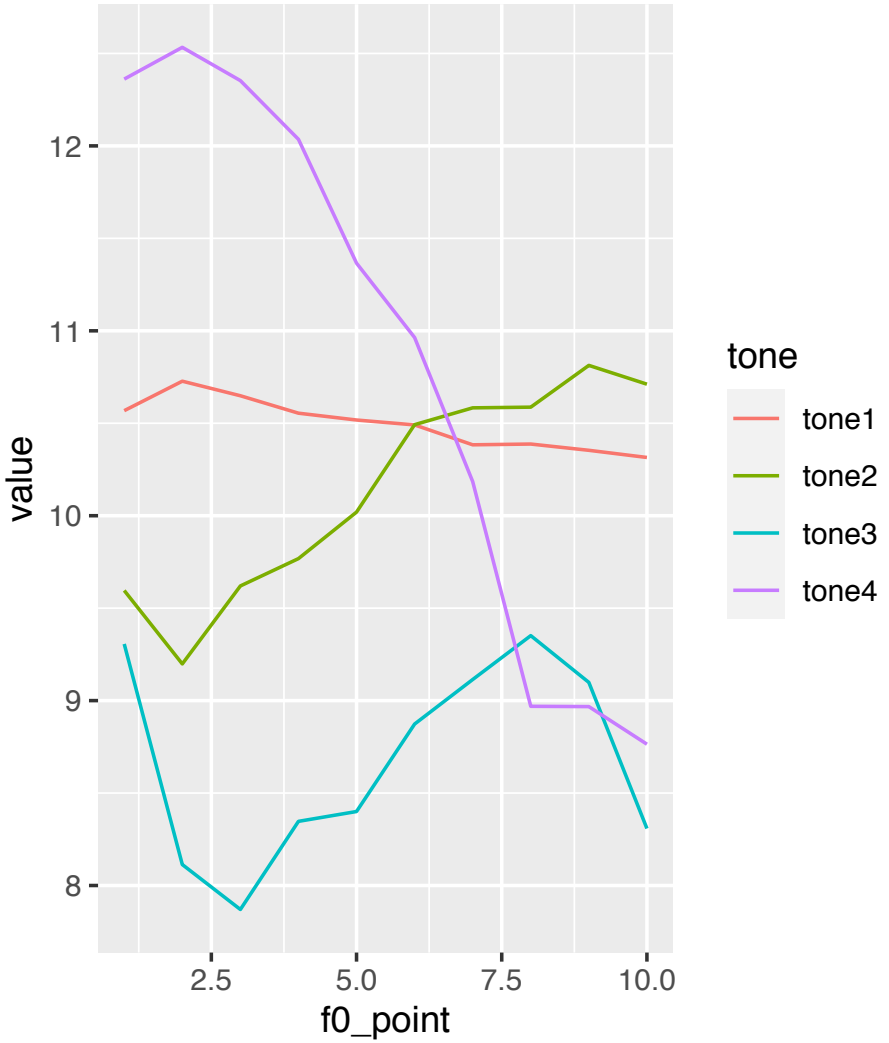
Chengdu



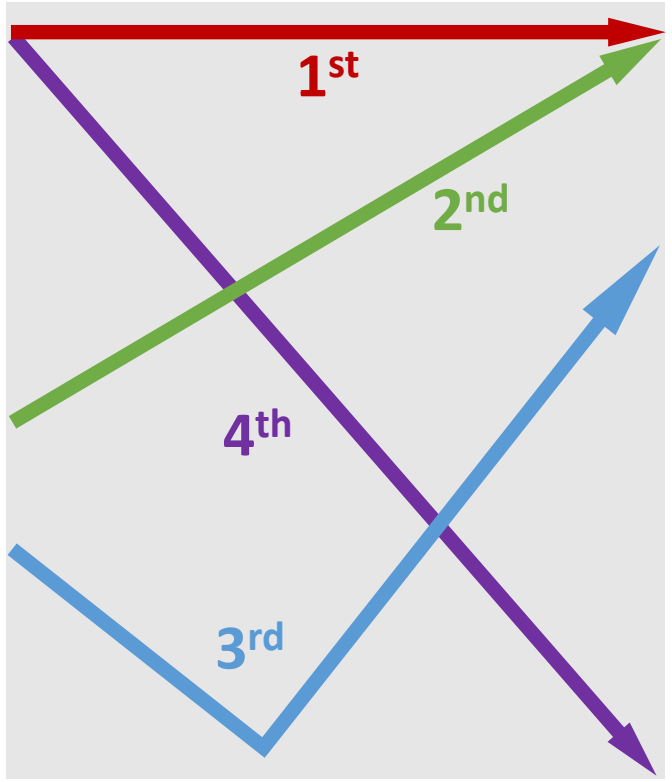
Wuhan



Beijing

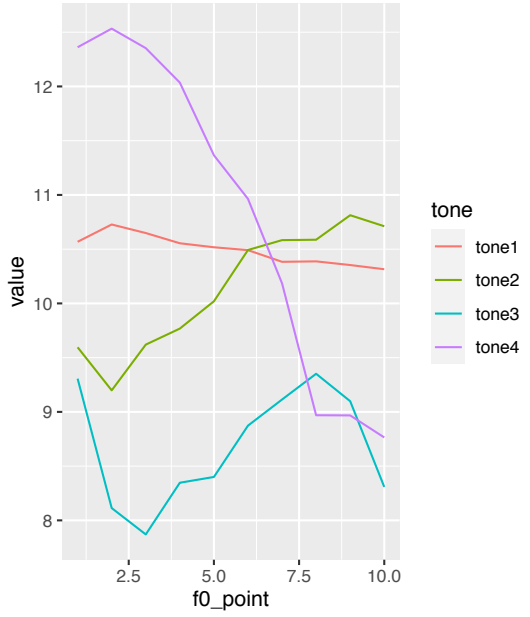


Standard Mandarin

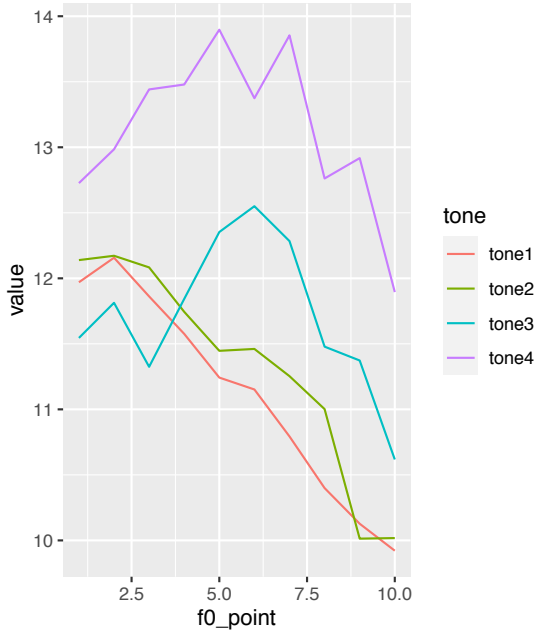


- Tone categories of six dialects based on mean f0 from **WL1** (monosyllabic characters)

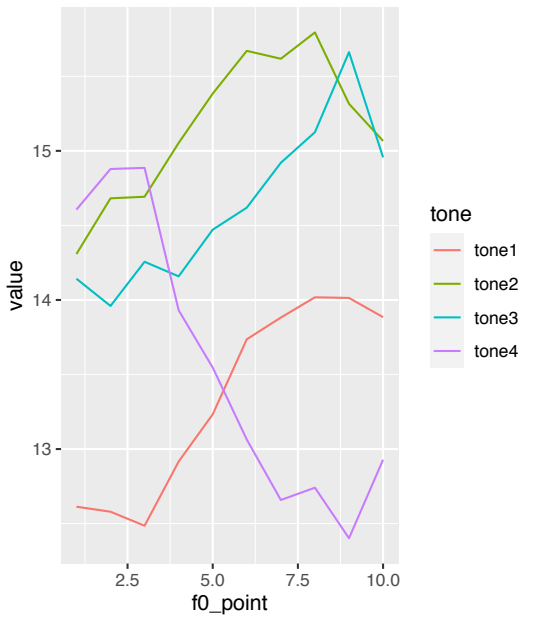
Beijing



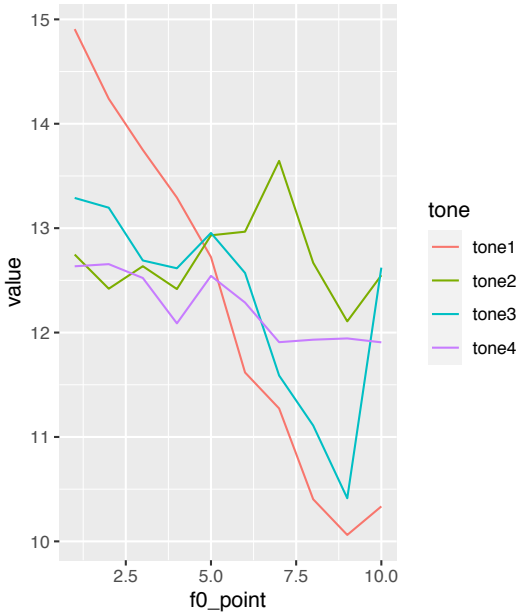
Taiyuan



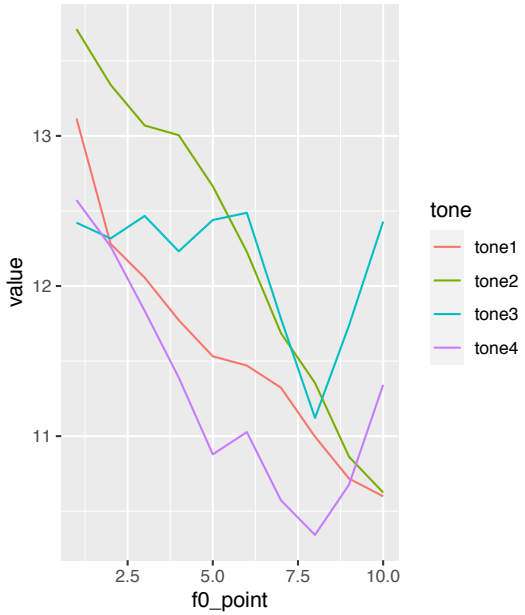
Jinan



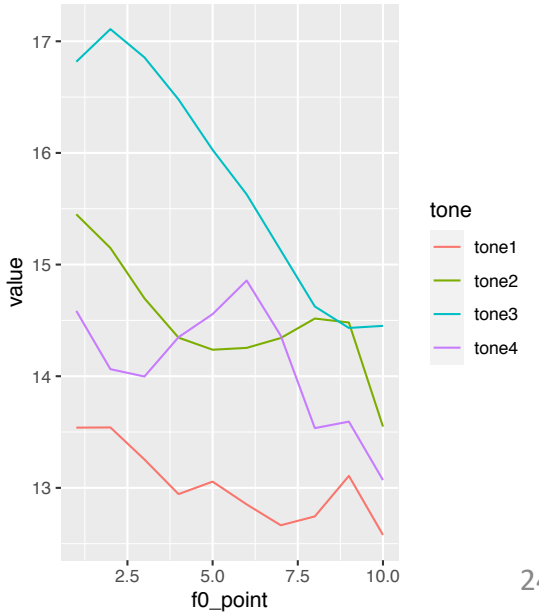
Xi'an



Chengdu



Wuhan



4. Conclusion

Remotely collected speech data can be suitable for acoustic analysis focusing on relative patterns in talker- and dialect-specific phonetic systems

- Decent recording quality
- Practicable design of remote delivery of the experiment
- Sufficiently reliable measurements for understanding acoustic-phonetic variation in the vowel space and tone system

- Further research necessary to understand overall impact of device and environment noise on the absolute values of phonetic measurements

4. Conclusion

- Directions for further analysis

Specific effects of different sources of variability on acoustics measurements

- Between-subject tests on lab recordings and smartphone recordings
  - Stimuli: recordings of *North Wind and the Sun* from ALLSSTAR corpus
- Within-subject tests on lab recordings and smartphone recordings
  - Comparison between environments and smartphone devices

4. Conclusion

- Suggestions on future remote audio collection

  - Design
    - Experimental instructions: video demonstration plus written clarifications
    - Communication: pre-informed time & date on when to record
    - Duration of the experiment

  - Recordings
    - Separate recording clips for multiple production tasks
    - Completing all production tasks with randomized stimuli in one attempt, instead of several attempts on different days to avoid data missing

  - Audio file uploading
    - Multiple options for participants to share, upload or send the files

14-15 December 2020
(Sydney time)

Thank you for your attention!

One million km^2: Broad coverage of Mandarin Chinese dialects using smartphone technology

Liang Zhao, Philip Harrison, Paul Foulkes and Eleanor Chodroff

University of York, December 2, 2020

Liang Zhao: lz1136@york.ac.uk