



# Integrated top-down and bottom-up processing of non-native tone systems

---

Liang Zhao and Eleanor Chodroff  
Department of Language and Linguistic Science  
University of York  
September 29, 2021



# Outline

---

## 1. Introduction

## 2. Production experiment

- Task overview: Mandarin dialect production experiment (remote audio collection)
- Goal: Establish the acoustic-phonetic tonal realisations in Beijing Mandarin (~Standard Mandarin) and Chengdu Mandarin

## 3. Perception experiment

- Task overview: Surprisal sentence perception (accuracy and reaction time)
- Goal: Determine the relative weighting of top-down and bottom-up information in the perception of lexical tone

## 4. Discussion



# Speech perception revisited

---

speech signal ————— intended utterance

Path from acoustic signal to units of perception: *many-to-many*

- “*The fact that speech soundwaves are not a one-to-one/linear encoding of a speaker’s utterance plan*” (Tatham & Morton, 2006)

The main problem: to decode the intended utterance from the speech signal



# Speech perception revisited

---

speech signal



intended utterance

Path from acoustic signal to units of perception: *many-to-many*

- “*The fact that speech soundwaves are not a one-to-one/linear encoding of a speaker’s utterance plan*” (Tatham & Morton, 2006)

The main problem: to decode the intended utterance from the speech signal



# Speech perception: bottom-up and/or top-down

---

Many early models assumed bottom-up processing as a first attempt

e.g. The Cohort Model (Marslen-Wilson 1978, 1987), Direct Perception (Gibson 1954) & Direct Realism (Fowler 1986)

Further development of the theory did take top-down influence into consideration

- Advocate: TRACE (McClelland & Elman 1986), Acoustic landmarks & distinctive features (Stevens 2002, 2008)
- Opponent: Shortlist (Norris 1994), Merge (Norris, Cutler and McQueen 2000)

**But overall top-down influence remains unclear/brief in many theories and their applications**



# Speech perception: segmental (and suprasegmental)

---

Models of speech perception have been segment-oriented

What about suprasegmental processing?



# Speech perception: lexical tone processing

---

## Debate on the relative role of lexical tone and segmental information for lexical access

- Segmental information >> lexical tone in sub-lexical processing  
(Cutler & Chen 1999, Ye & Connine 2010, Li et al. 2013)
- Lexical tone >> or  $\approx$  segmental information with top-down feedback  
(Schirmer et al. 2005, Liu & Samuel, 2007, Malins & Joanisse, 2010)

## Extension of TRACE: Reverse Accessing Model (RAM, Gao et al. 2019)

- Report on “distinctive advantage of word and **atonal syllable** judgments over phoneme and lexical tone judgments”



# Speech perception: non-native lexical tone processing

---

## Lexical tone processing

- Context may be crucial
- Previous studies were most often with a native lexical tone system

Would **non-native lexical tone** processing be any different? And how?

What happens when **the tone system is non-native, but the segmental system is familiar?**

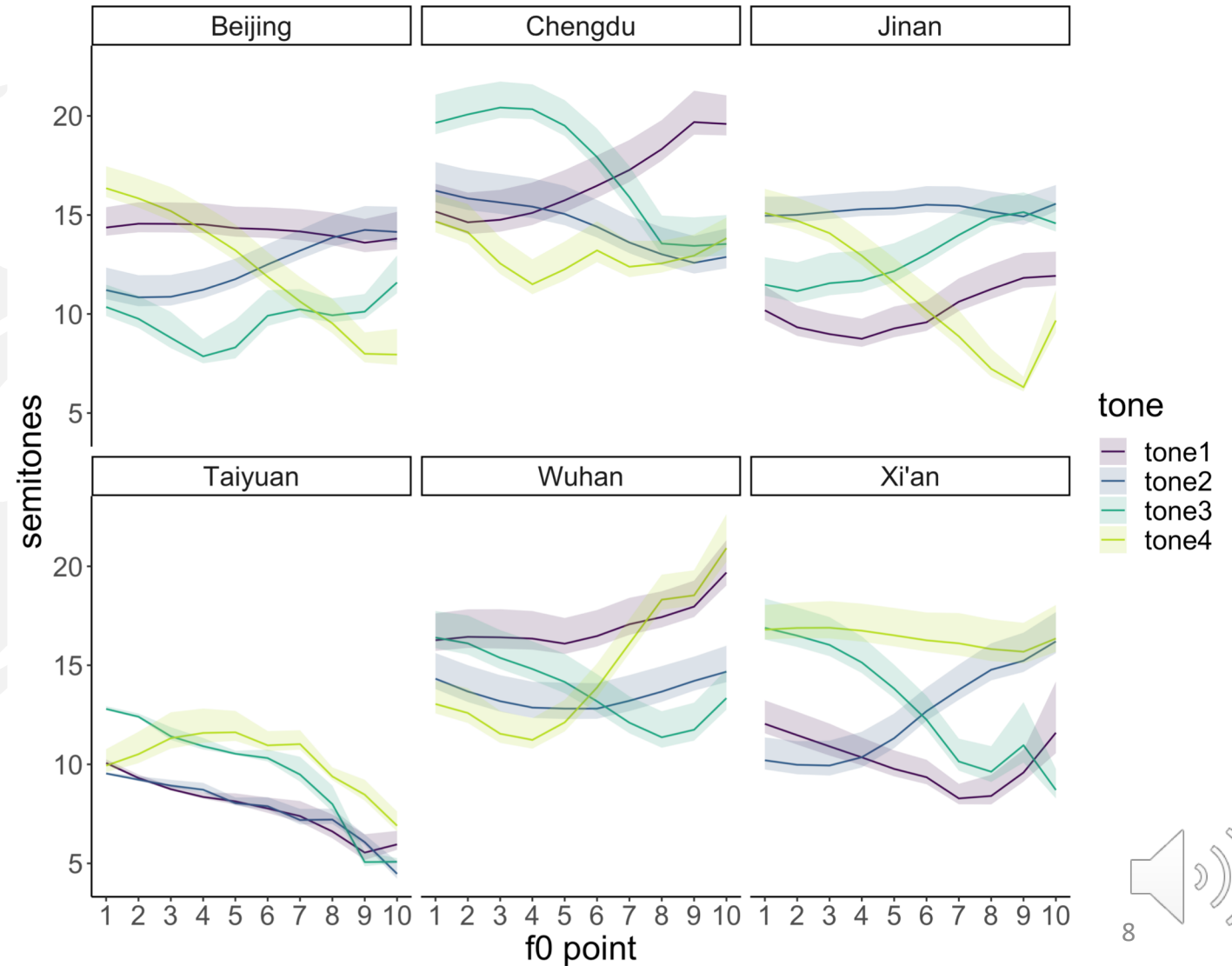
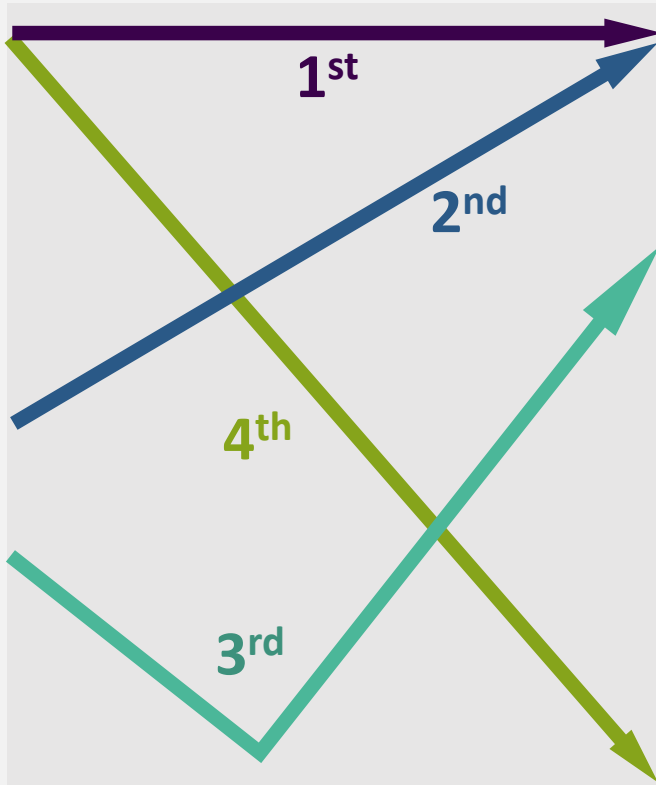
First, let's get to know about the tone system of **Mandarin dialects**





# Comparable segmental inventories, but **distinct** tone inventories

Standard Mandarin Lexical Tones (Beijing)



# Speech perception: non-native Mandarin dialects

---

## Mutual intelligibility (Tang & van Heuven, 2007, 2008, 2009)

- Relatively high mutual intelligibility in semantic decision task with carrier sentences
- Intelligibility dropped with words presented in isolation

We hypothesize an essential role of **top-down information** in perception of non-native speech

BUT!

- Fully top-down?
- Hybrid processing with both top-down and bottom-up processing?



# Outline

---

## 1. Introduction

## 2. Production experiment

- Task overview: Mandarin dialect production experiment
- Goal: Establish the acoustic-phonetic tonal realisations in Beijing Mandarin (~Standard Mandarin) and Chengdu Mandarin

## 3. Perception experiment

- Task overview: Surprisal sentence perception (accuracy and reaction time)
- Goal: Determine the relative weighting of top-down and bottom-up information in the perception of lexical tone

## 4. Discussion



# Production experiment

---

## Participants

- 8 native Beijing speakers (represent Standard Mandarin)
- 7 native Chengdu speakers

## Stimuli (reading material) & Tasks

- **80 monosyllabic words** (20 forms \* four tones)
- **48 sentences** (potential perceptual stimuli)
- Production in Standard Mandarin and their regional dialect



# Production experiment

---

## Acoustic analysis

- Alignment
  - Transcripts generated from *Gorilla* and processed as text files for each speaker
  - Forced alignment using Montreal Forced Aligner
  - Manual adjustment of vowel boundaries
- Tone categories: f0 contours
  - 10 equally spaced f0 values across all vocalic intervals
  - Converted to semitones



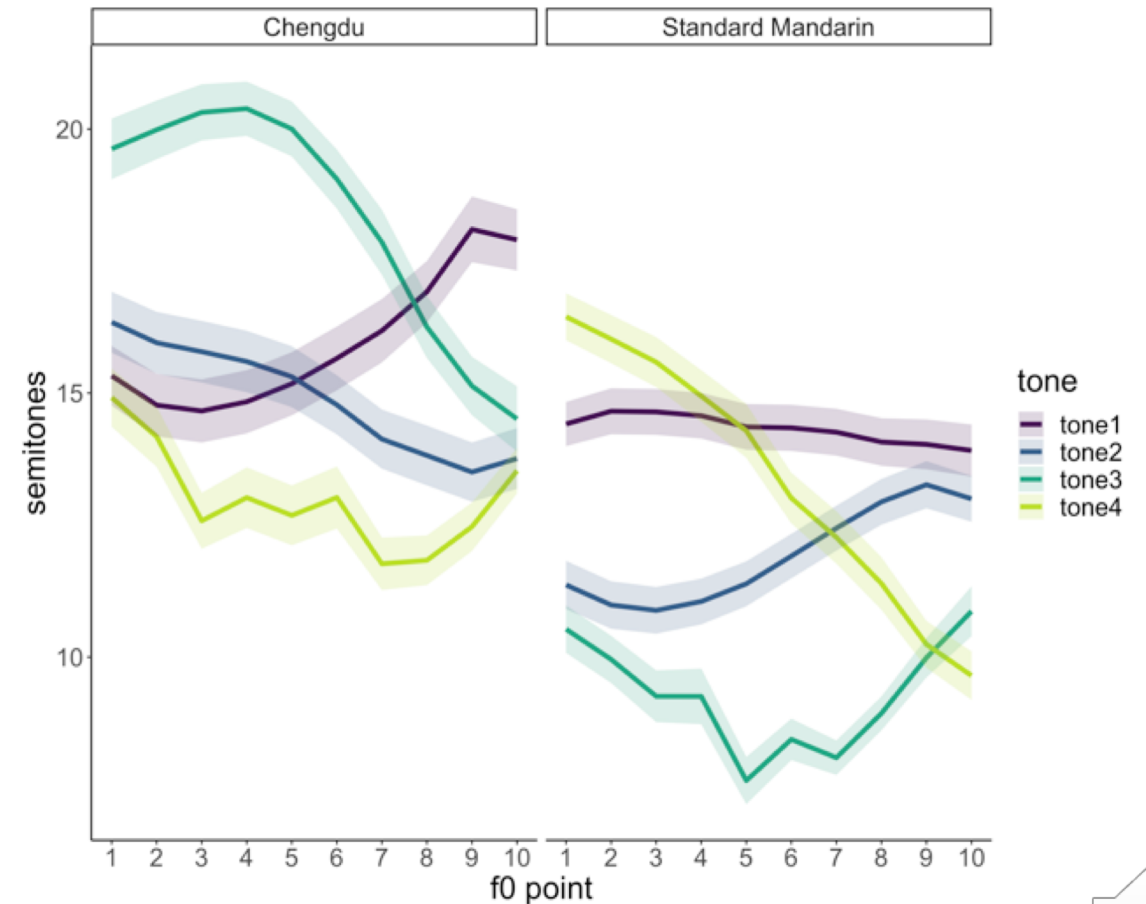
# Production experiment

## Results

- Highly disparate realisations of lexical tone between Standard Mandarin and Chengdu Mandarin
- Contour similarity observed for some tone categories

	Chengdu	Beijing
Tone 1	25 (55*)	55
Tone 2	21	35
Tone 3	53	214
Tone 4	213	51

- (55\*) is what has been previously recorded in dialect dictionaries, but our result shows this tone is more likely a rising tone (25)



# Outline

---

## 1. Introduction

## 2. Production experiment

- Task overview: Mandarin dialect production experiment (remote audio collection)
- Goal: Establish the acoustic-phonetic tonal realisations in Beijing Mandarin (~Standard Mandarin) and Chengdu Mandarin

## 3. Perception experiment

- Task overview: Surprisal sentence perception (accuracy and reaction time)
- Goal: Determine the relative weighting of top-down and bottom-up information in the perception of lexical tone

## 4. Discussion



# Perception experiment

---

## Participants

- 21 native Mandarin, non-Chengdu speakers

## Task & Procedure

- Semantic plausibility judgment task of a spoken sentence
  - “Does this sentence make sense?” and clicked **“yes”** or **“no”** on the screen after hearing the whole sentence
- Accuracy and reaction time were measured as independent variables
  - If an answer matched the expected plausibility, then it was considered correct





# Perception experiment

---

Experimental manipulation (48 trials = 12 sentences \* 4 conditions)

- **Dialect conditions:** Standard Mandarin vs. Chengdu Mandarin
- **Surprisal conditions:** low surprisal vs. high surprisal



# Experimental manipulation (48 trials = 12 items \* 4 conditions)

- An example item in the four conditions:

	Low-surprisal sentence	High-surprisal sentence
Standard Mandarin	<p>有 一只 鹰 在 天上 飞</p> <p>You3 yi4 zhi1 ying1 zai4 tian1 shang4 <u>fei1</u></p> <p><i>There is an eagle in the sky <u>flying</u></i></p> <p>“There is an eagle flying in the sky”</p>	<p>有 一只 鹰 在 天上 肥*</p> <p>You3 yi4 zhi1 ying1 zai4 tian1 shang4 <u>fei2*</u></p> <p><i>There is an eagle in the sky <u>gaining weight*</u></i></p> <p>“There is an eagle gaining weight in the sky”</p>
Chengdu Mandarin	<p>有 一只 鹰 在 天上 飞</p> <p>You yi zhi ying zai tian shang <u>fei</u></p> <p><i>There is an eagle in the sky <u>flying</u></i></p> <p>“There is an eagle flying in the sky”</p>	<p>有 一只 鹰 在 天上 肥*</p> <p>You yi zhi ying zai tian shang <u>fei*</u></p> <p><i>There is an eagle in the sky <u>gaining weight*</u></i></p> <p>“There is an eagle gaining weight in the sky”</p>



# Perception experiment

---

## More on stimuli

All 48 sentences were

- produced by a native, male speaker of Standard Mandarin and Chengdu Mandarin
- randomly presented with no repetitions or indication of dialect

Listeners were made aware of the two possible dialects in the task instructions



# Perception experiment

---

## Analysis

Assessing **accuracy** and **reaction time** for the effects of semantic plausibility (**surprisal**) and dialect familiarity (**dialect**)



# Accuracy

Expected results in the semantic plausibility task

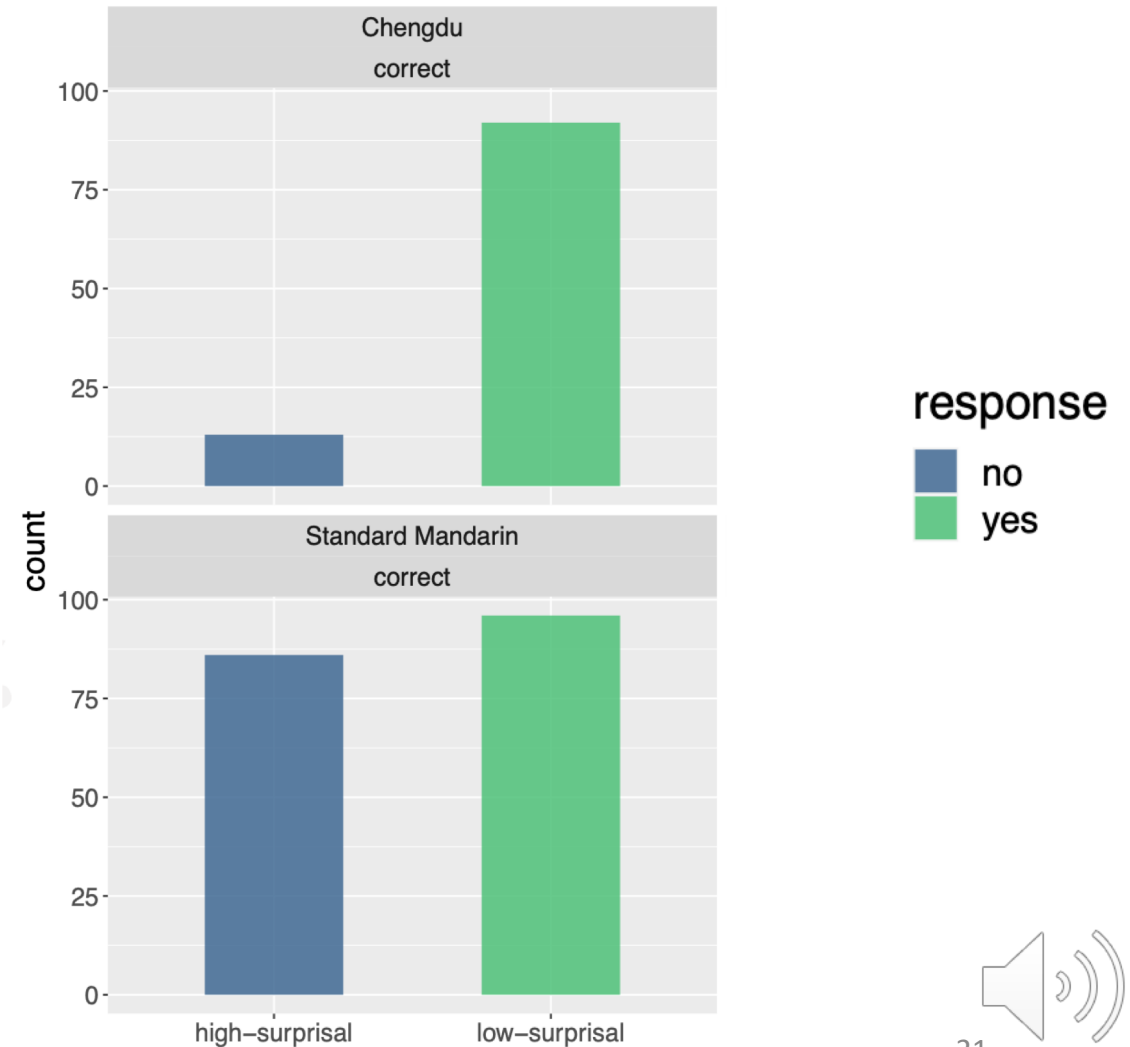
	Low-surprisal sentence	High-surprisal sentence
Native speech	Accuracy ↑ “yes”	Accuracy ↑ “no”
Non-native speech	Accuracy ↓ “yes”	Accuracy ↓ *“yes”



# Accuracy

Significant main effect of *surprisal*, *dialect* and an *interaction between surprisal and dialect* (each  $p < 0.01$ )

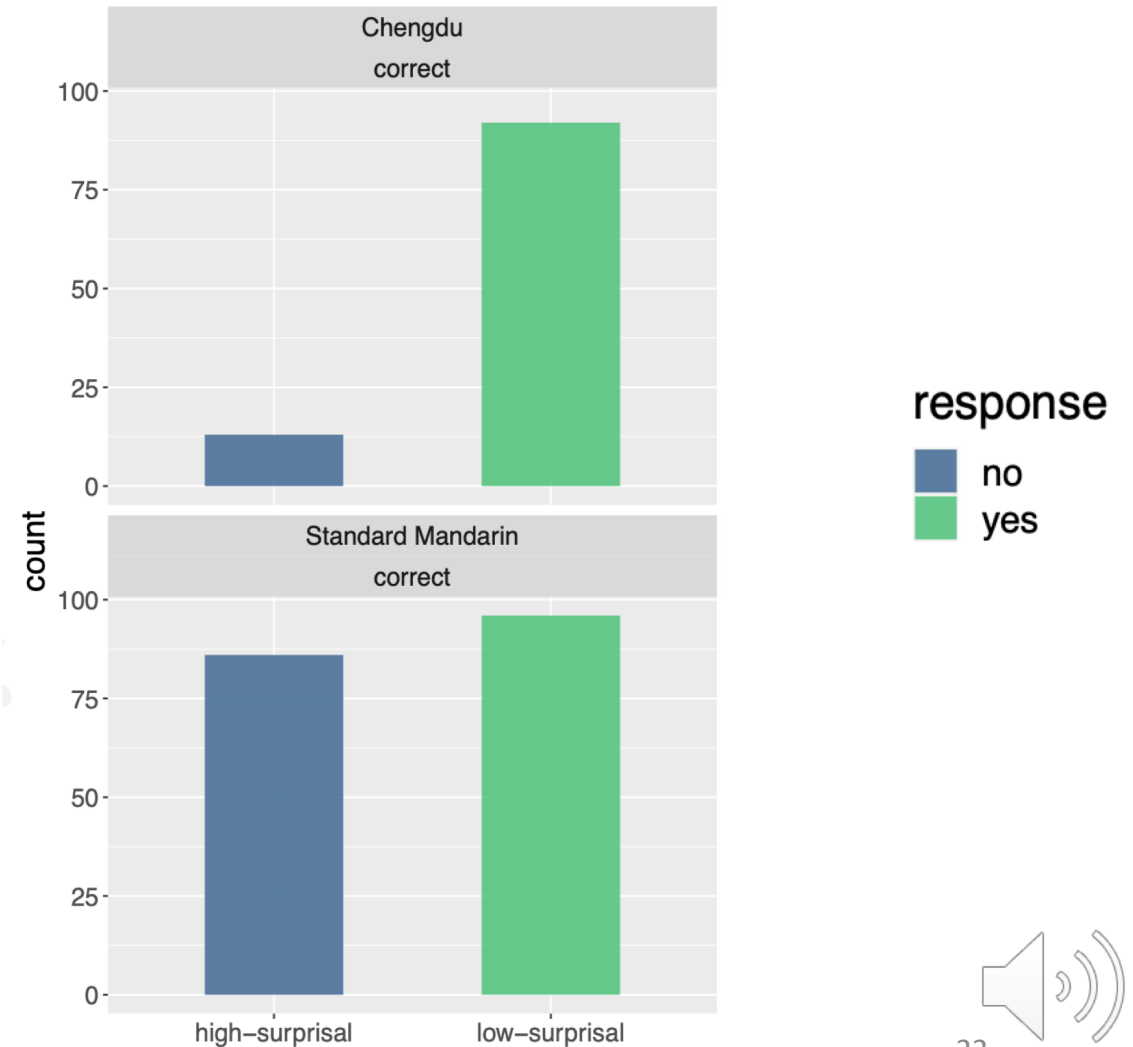
- *Surprisal*: low-surprisal >> high-surprisal condition
- *Dialect*: Standard Mandarin >> Chengdu Mandarin
- *Interaction*: significantly less accurate in the high-surprisal Chengdu condition relative to average



# Accuracy

## Native speech (Standard Mandarin)

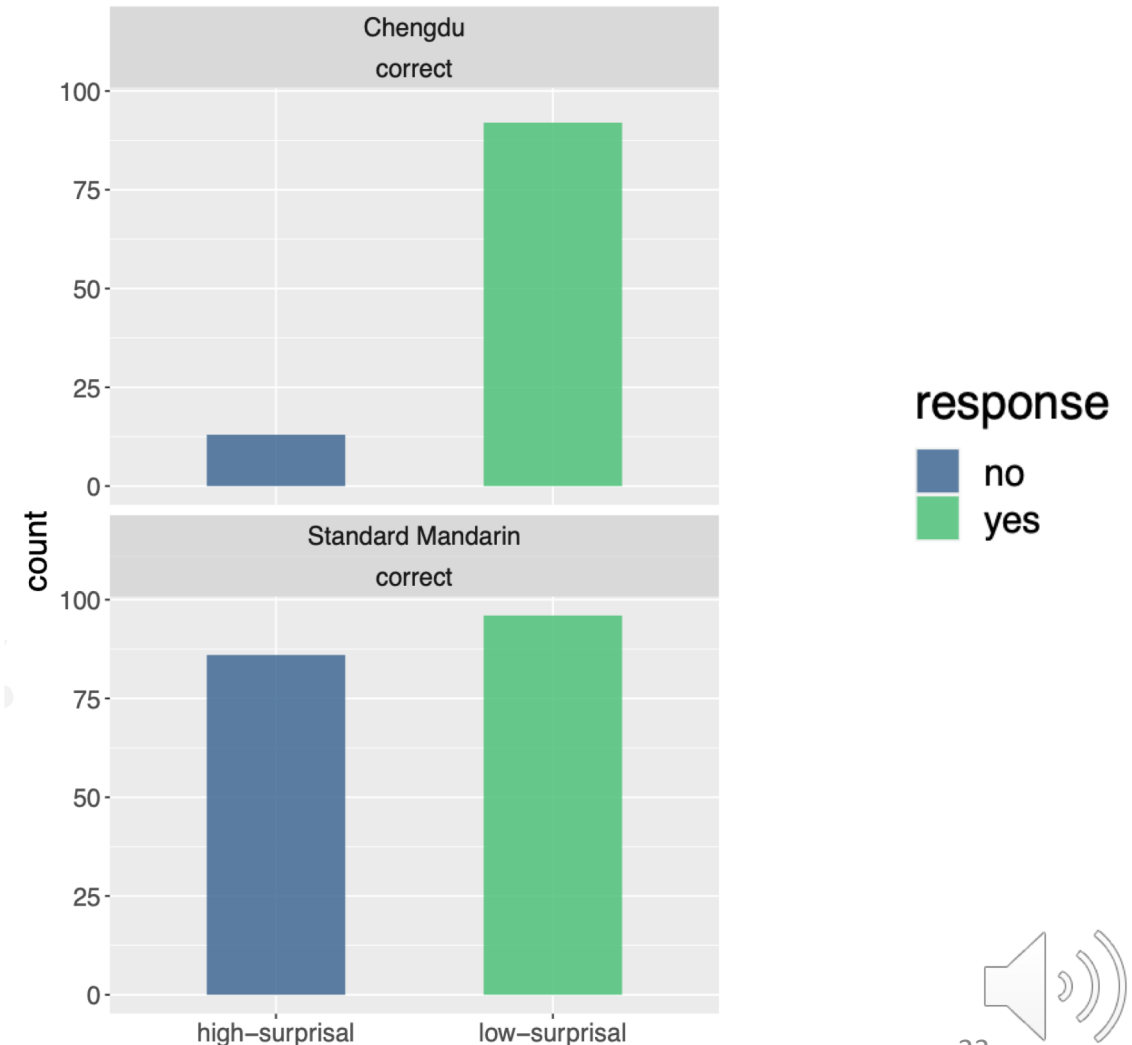
- Overall accuracy at 92%  
(Saying implausible when they should say implausible)
- Listeners do understand the task
- Use both top-down (sentential context) and bottom-up information (word-tone knowledge)



# Accuracy

## Non-native speech (Chengdu Mandarin)

- Bias to respond “semantically plausible”/ “yes” (86%) for all sentences
  - Suggests relatively **high intelligibility**
- Unable to use tone information in high-surprisal sentences
  - Major bottom-up **\*failure\*** in identifying tone mismatch in high-surprisal condition
- **Top-down influence overriding bottom-up tone acoustics**





# Reaction time

Expected results in the semantic plausibility task

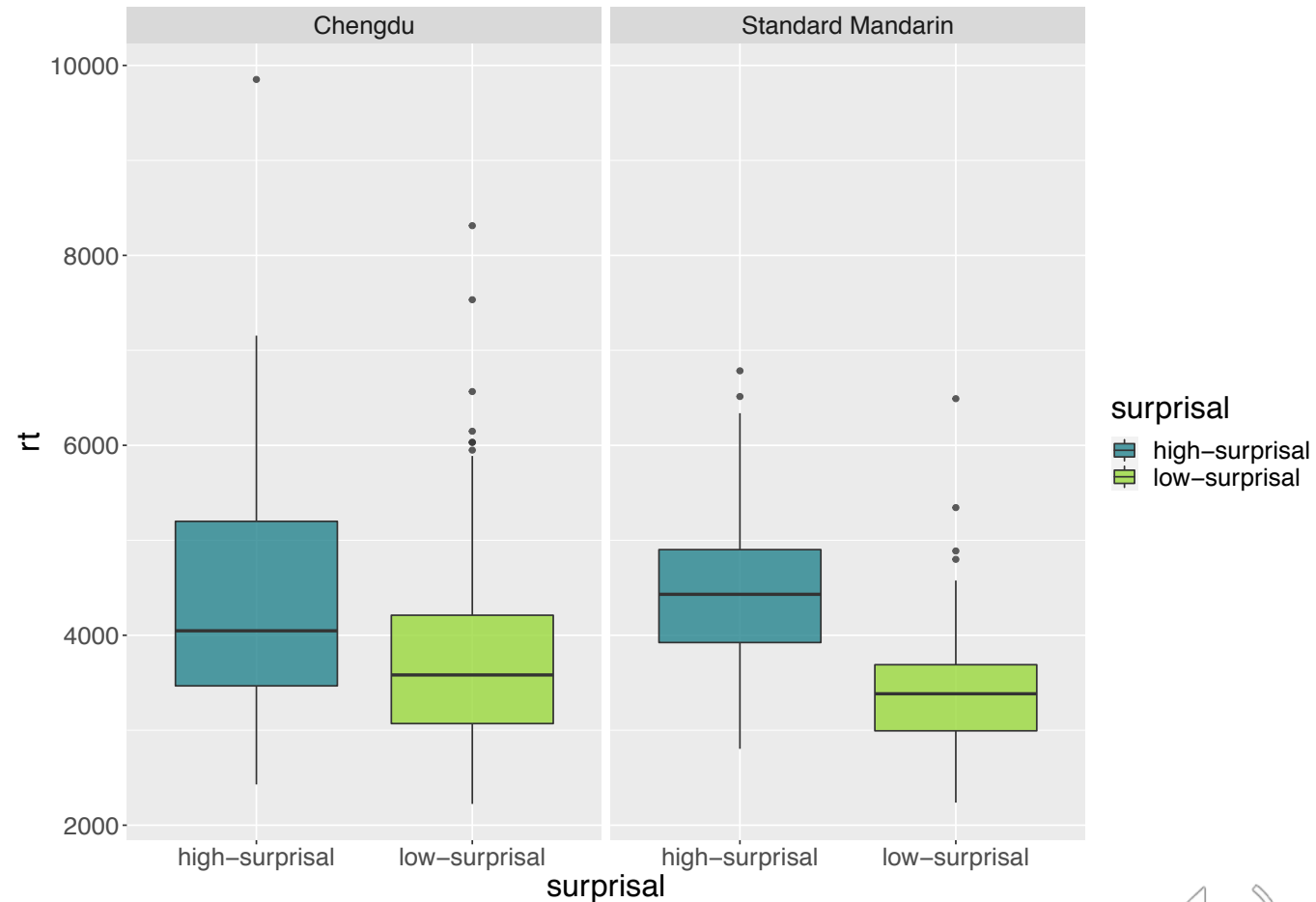
	Low-surprisal sentence	High-surprisal sentence
Native speech	baseline	slowdown ↓
Non-native speech	==	==



# Reaction time

## Significant main effect of *surprisal*

- Significant slowdown in high-surprisal condition for **both** native and non-native speech
  - For native speech: expected
  - For non-native speech: **unexpected!**
- Both bottom-up and top-down processing are present
  - **Processing tone information in non-native speech!**



# Summary

---

## Findings

- Highly **disparate tone systems** between Chengdu and Standard Mandarin
- **High intelligibility** of Chengdu dialect by native speakers of Standard Mandarin
- Accuracy results suggest **dominance of higher-level top-down information** in perception of non-native Mandarin dialect (Chengdu)
- Reaction time results suggest sensitivity and potential learning of the non-native tone system indicating **the presence of bottom-up processing**



# Discussion

---

Bottom-up and top-down processing are both actively involved in perception of native and non-native tone systems

- Contrary to RAM (Gao et al. 2019) which proposed that tone information is “processed *if necessary*”, we argue that tone information is *always* processed



# Discussion

---

Bottom-up and top-down processing are both actively involved in native and non-native speech perception

## **Native speech:**

- Accuracy results suggest that listeners have strong representations of tone and segments (bottom-up)
- Reaction time results suggest they are sensitive to the sentence surprisal (top-down)



# Discussion

---

Bottom-up and top-down processing are both actively involved in native and non-native conditions

## Non-native speech:

- Accuracy results suggest listeners use context alone (top-down)
- Differences in reaction time suggests that they are constructing *impoverished tone representations* during online processing (bottom-up)
  - We don't know if they are building long-lasting representations (learning) or temporary representations
  - We do know that top-down information typically overrides the output of the tone level (low confidence)



# Discussion

---

## Follow-up experiment with Chengdu Mandarin exposure

- Accuracy improves after two minutes of uninterrupted exposure
- Listeners may indeed be learning tone representations or gaining confidence of them



# Future directions

---

Current study could be extended:

- To experiment on a broad range of Mandarin dialects to test the hypotheses
- To investigate whether perception varies across dialects and explain potential patterns using production data
- To examine the relative temporal relations between top-down and bottom-up processing (Shuai & Gong, 2014)





# Thank you!

---

Special thanks to Dr. Shayne Sloggett! (shayne.sloggett@york.ac.uk )

## Contact

Liang Zhao: [lz1136@york.ac.uk](mailto:lz1136@york.ac.uk)

Eleanor Chodroff: [eleanor.chodroff@york.ac.uk](mailto:eleanor.chodroff@york.ac.uk)

