Rapid adaptation to unfamiliar lexical tone systems:

the effects of *dialect* and *explicit exposure*

Liang Zhao[1] and Eleanor Chodroff[2]

University of York[1], University of Zurich[2]

June 3, 2023

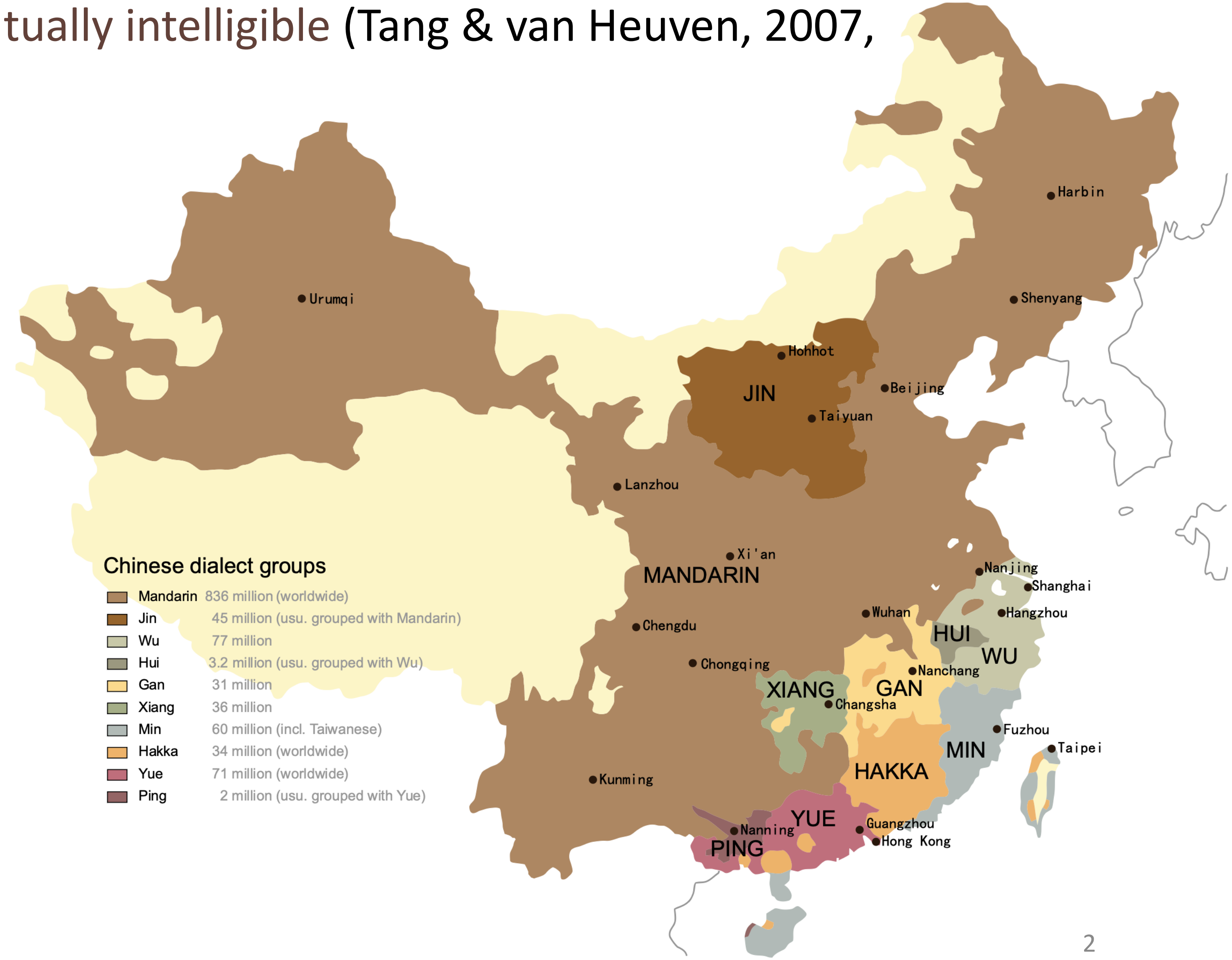
Mandarin regional dialects

Mandarin branch dialects: relatively mutually intelligible (Tang & van Heuven, 2007, 2008, 2009)

- **Mandarin**

Non-Mandarin branch dialects:

- **Wu**
- **Hui**
- **Gan**
- **Xiang**
- **Min**
- **Hakka**
- **Yue**
- **Ping**



Language Atlas of China (1989)

Mandarin regional dialects

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

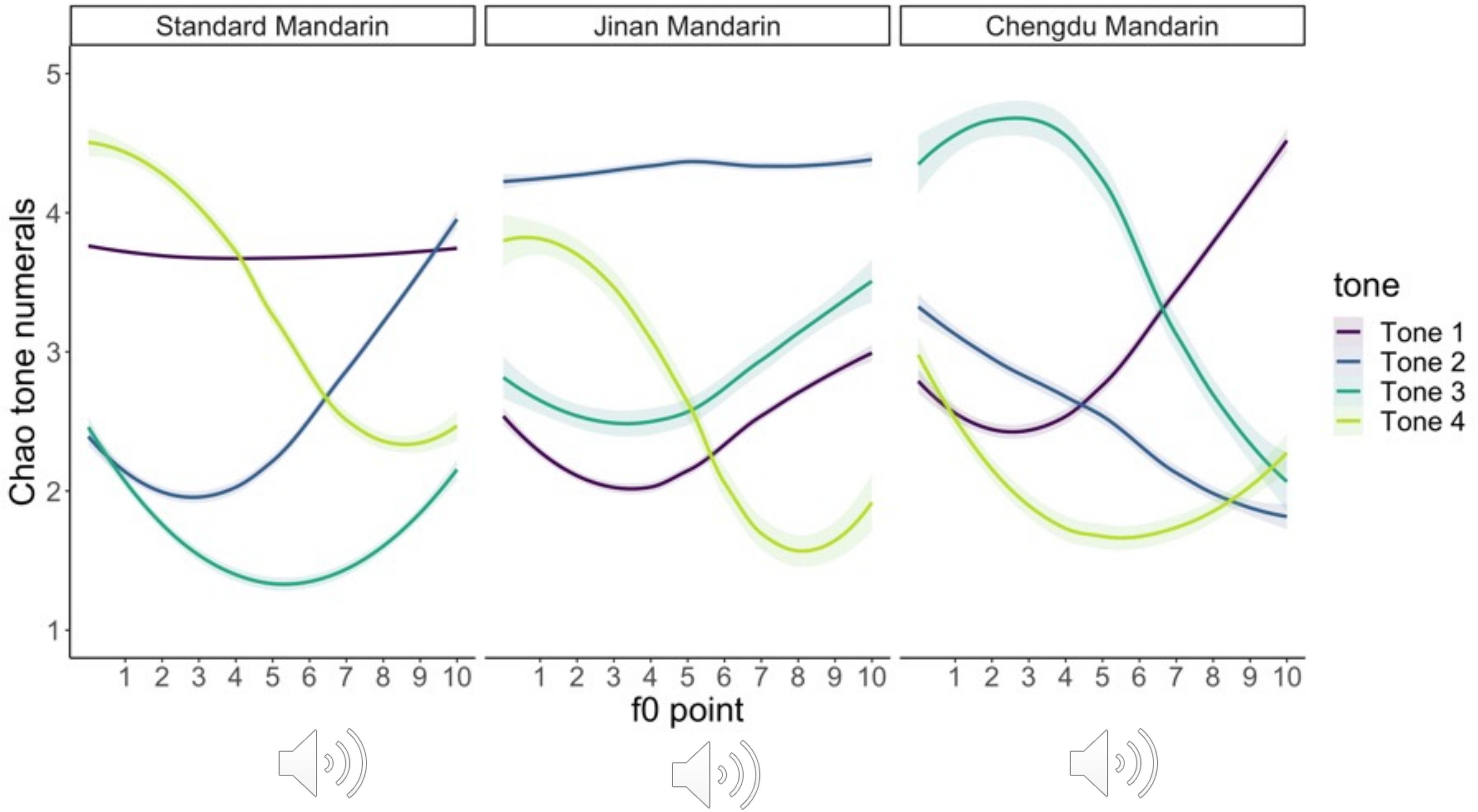
Evidence from:

- Chao tone descriptions from the canonical Chinese dialect dictionaries (Li, 1998; Hou, 2002)

  ---- “tang1”: tang55 (Standard Mandarin), tang35 (Jinan), tang 213 (Chengdu)

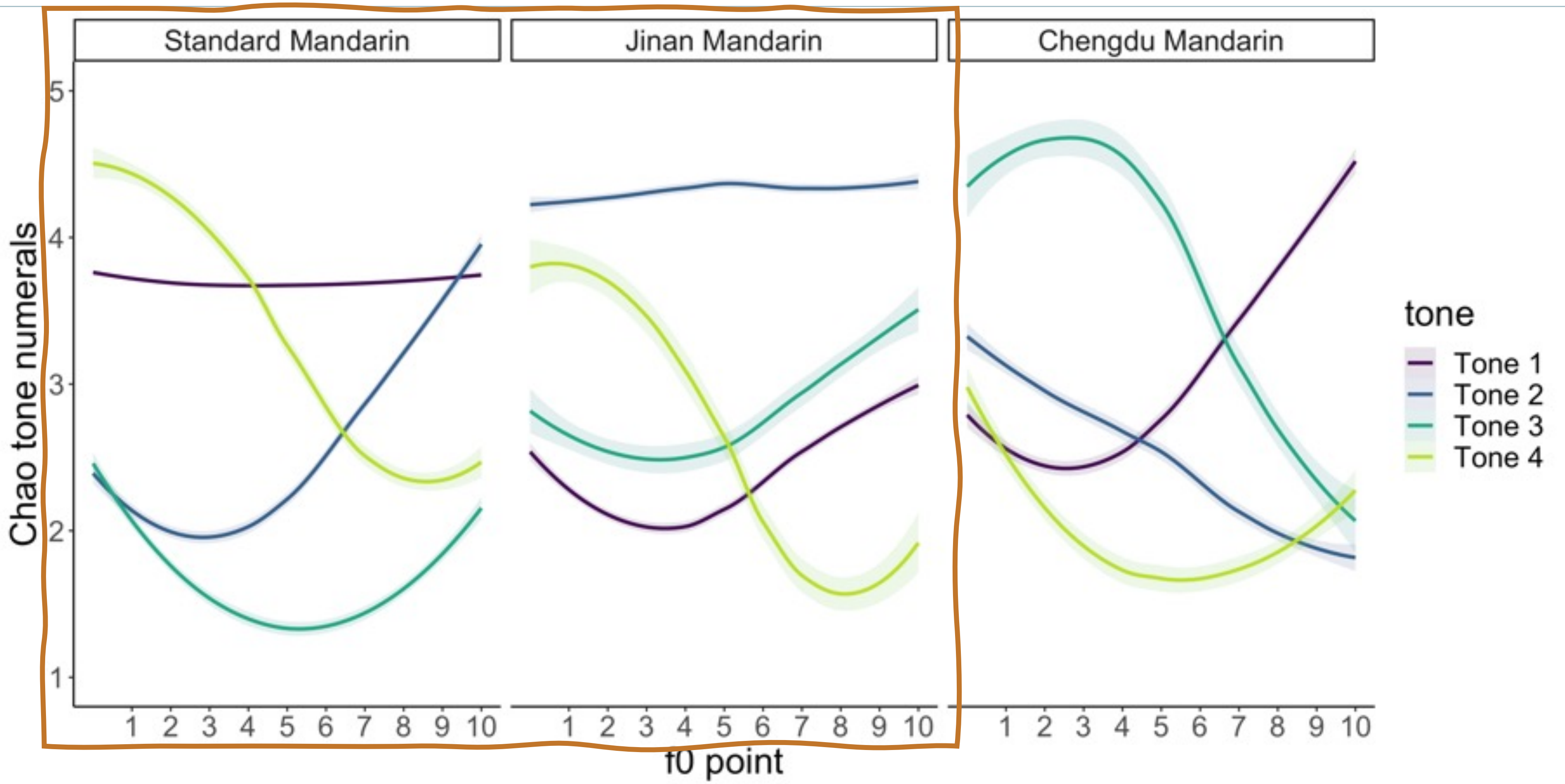
- Corpus-phonetic analysis of the tone systems with the spoken data (ManDi Corpus, Zhao & Chodroff, 2022)

Mandarin dialects:
Comparable segmental inventories, but **distinct** tone realizations



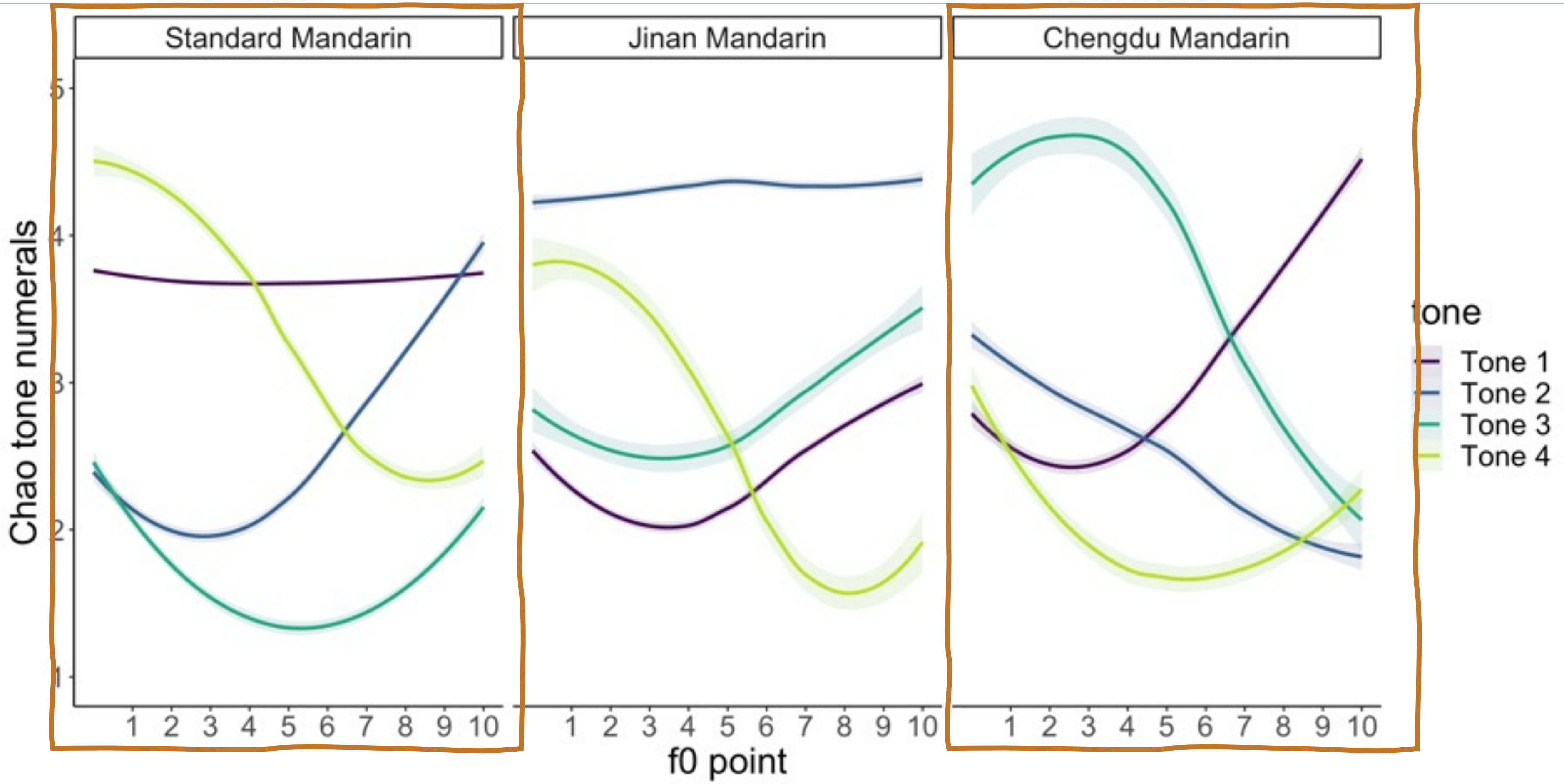
Example word: **tang1 /taŋ/** (“soup”; 汤)

Standard Mandarin vs. Jinan & Chengdu



Jinan Mandarin: **phonetically and perceptually similar** contours compared to Standard Mandarin

Standard Mandarin vs. Jinan & Chengdu



Chengdu Mandarin: **phonetically and perceptually less similar** compared to Standard Mandarin

Task for the listener ?

Processing unfamiliar Mandarin dialectal speech:

- Segmental information: familiar
- Tonal information: unfamiliar

How do listeners process novel tones?

General questions:

- What **perceptual mechanisms** are used in processing phonetic tone variation in these dialects?

- What might be the **potential factors** affecting perceptual adaptation to an unfamiliar tone system?

Background: Novel tone processing

- Joint top-down & bottom-up processing for lexical access with tone information:
    - Predominant use of sentential context and segmental information
    - Evidence that listeners are sensitive to new tone acoustics and phonological associations

- Listeners start to adapt to an unfamiliar tone system from experimental trials directly

(Zhao, Sloggett & Chodroff, 2022, 2023)

Primary questions

- Does adaptation differ **between dialects**?

---- Is adaptation easier for a tone system that is more **dissimilar** to the listener’s native tone system? (e.g., Perceptual Assimilation Model, Best & Tyler, 2007; So & Best, 2011)

Chengdu >> Jinan

- What about **explicit exposure**?

---- Can explicit exposure to the target dialect further facilitate adaptation?

(post-exposure >> pre-exposure)

Perception experiment: surprisal-based processing

24 sentence pairs were created manipulating the phonetic tone of one target word to be:

- **low-surprisal** (semantically plausible sentence)
- **high-surprisal** (semantically implausible sentence)

low-surprisal condition	high-surprisal condition
a) 有 一只 鹰 在 天上 飞 *You3 yi4 zhi1 ying1 zai4 tian1 shang4 fei1* *There is an eagle in the sky flying* "There is an eagle flying in the sky"	b)* 有 一只 鹰 在 天上 肥 *You3 yi4 zhi1 ying1 zai4 tian1 shang4 fei2* *There is an eagle in the sky gaining weight* "There is an eagle gaining weight in the sky"

Sentence plausibility judgment task ("yes"/"no")

"Does this sentence make sense to you?"

Perception experiment

Participants: Standard Mandarin speakers from China

- 14 in the Chengdu Condition
- 13 in the Jinan Condition

Procedure (with *Gorilla* Experiment builder):

Familiarization → Pre-exposure task → Explicit exposure: *North Wind and the Sun* → Explicit exposure: *North Wind and the Sun with fill-in-the-blank task* → Post-exposure task

Measures

Accuracy (Bayesian logistic mixed-effects regression)

Response time (Bayesian linear mixed-effects regression (Bürkner, 2018)

- calculated as the interval between the end of the audio file and the click registering a judgment

Factors

- Surprisal (high vs low)
- Dialect (Chengdu vs Jinan)
- Exposure (pre vs post)
- (Tone)

Accuracy

Credible main effects of surprisal and exposure

- *Surprisal*: low-surprisal >> high-surprisal

❖ Listeners strongly report the sentences as plausible in both surprisal conditions

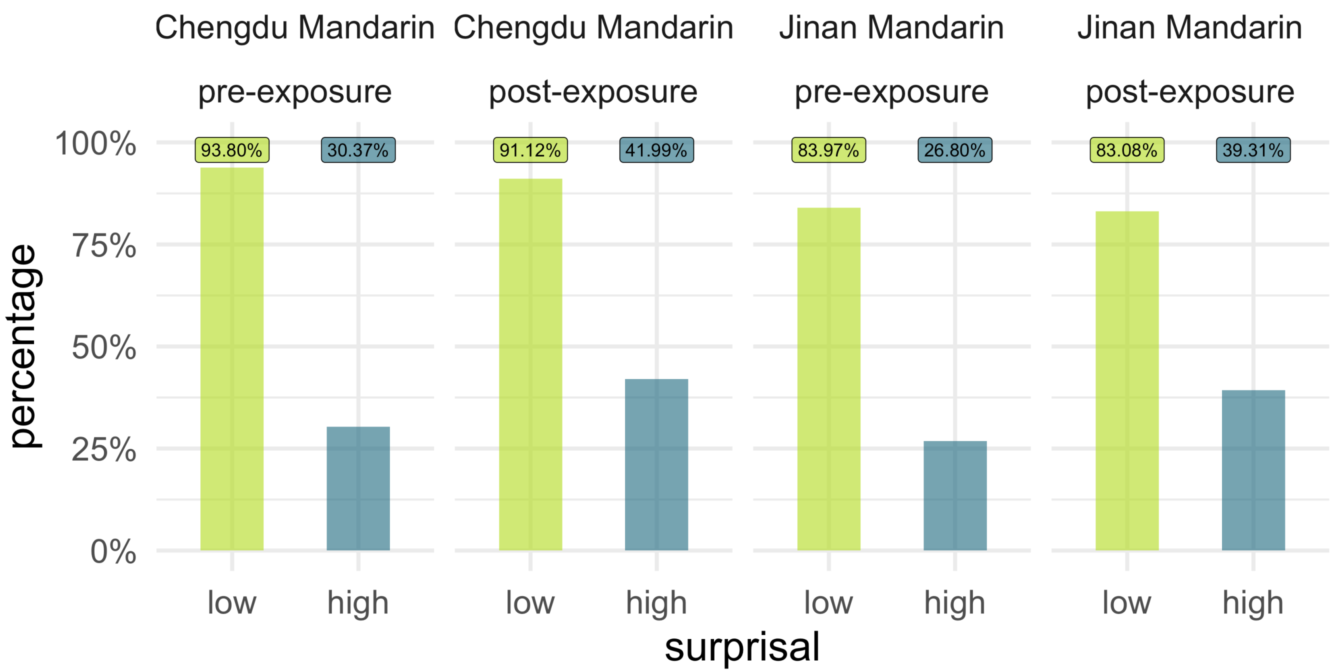


Figure: Percentage of "*correct*" responses across dialect, surprisal, and exposure conditions.

Accuracy

Credible main effects of surprisal and exposure

- *Exposure*: post-exposure >> pre-exposure

❖ Explicit exposure reliably improved accuracy of the sentence plausibility judgment task for both dialects

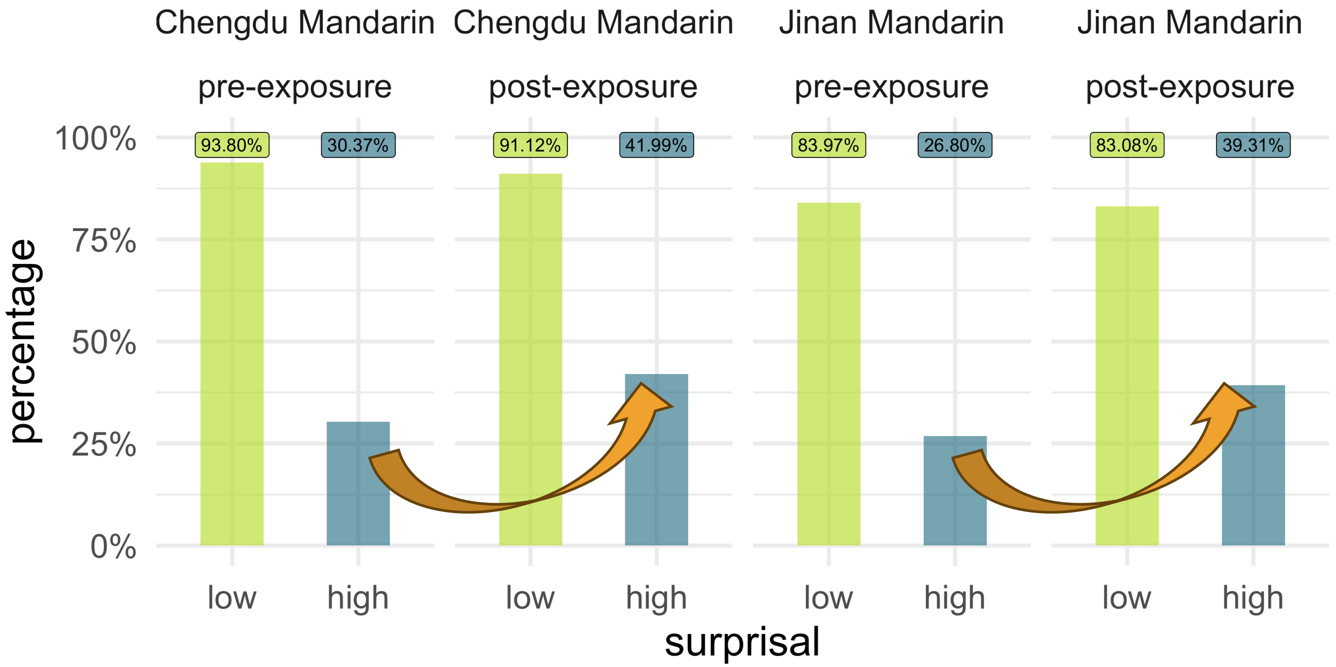


Figure: Percentage of "*correct*" responses across dialect, surprisal, and exposure conditions.

Accuracy

No effect of dialect, tone or any interaction:

- *Dialect*: Chengdu ≈ Jinan

❖ For accuracy, adaptation to the novel tone system did not reliably differ between Chengdu and Jinan Mandarin

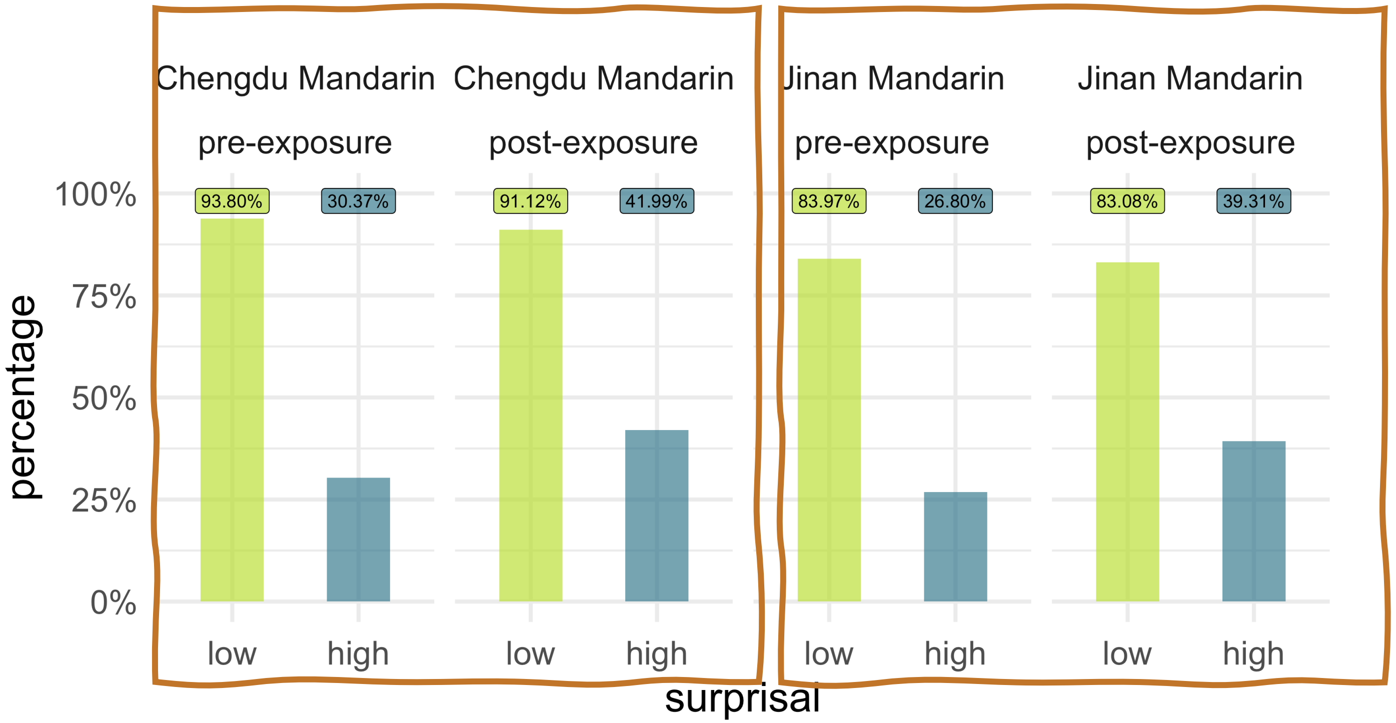


Figure: Percentage of "*correct*" responses across dialect, surprisal, and exposure conditions.

Response time

Credible main effects of surprisal, dialect, interaction between dialect and surprisal, and interaction of surprisal, dialect and exposure

- *Surprisal*: consistent slowdown for high-surprisal sentences

❖ Listeners were sensitive to the unfamiliar tones in both dialects
❖ *Listeners are not completely discarding tone information*

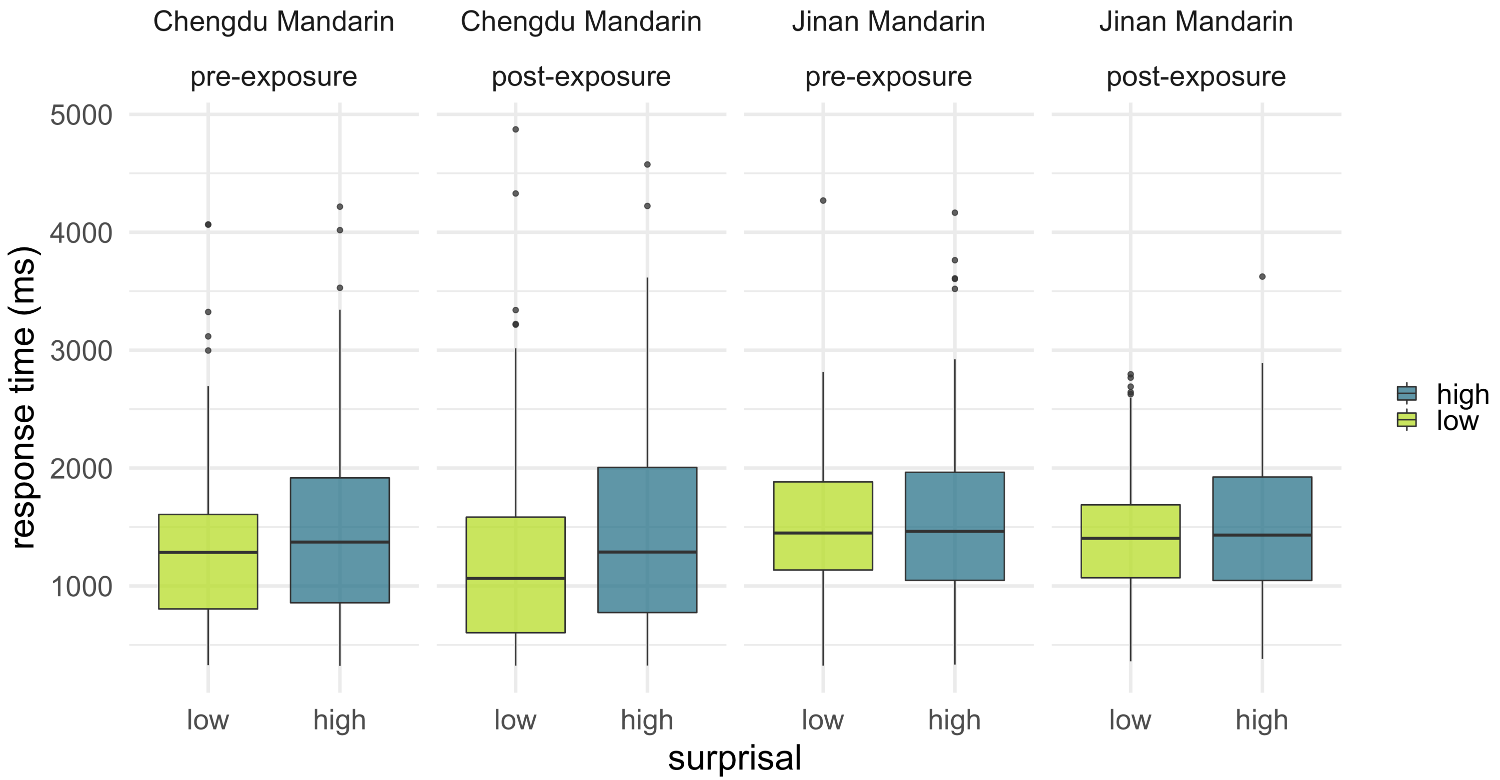


Figure: Response times across dialect, surprisal, and exposure conditions.

Response time

Credible main effects of surprisal, dialect, interaction between dialect and surprisal, and interaction of surprisal, dialect and exposure

- *Dialect*: faster responses for Chengdu sentences

❖ Chengdu sentences might be easier to understand than Jinan sentences

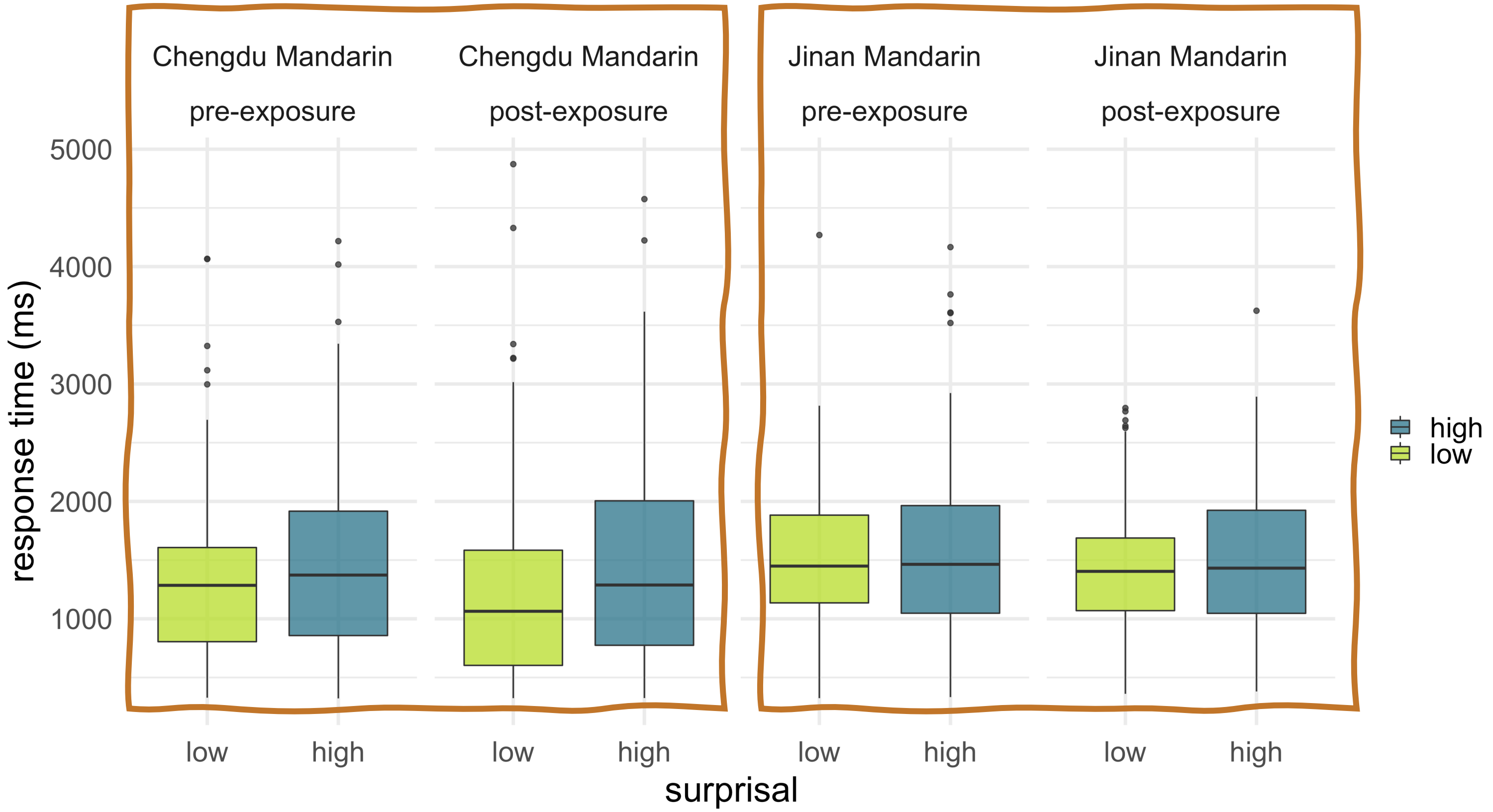


Figure: Response times across dialect, surprisal, and exposure conditions.

Response time

Credible main effects of surprisal, dialect, interaction between dialect and surprisal, and interaction of surprisal, dialect and exposure

- *Dialect x surprisal*: slower responses to the high-surprisal condition for Chengdu sentences than Jinan

❖ Increased sensitivity to the surprisal difference for Chengdu sentences than Jinan sentences

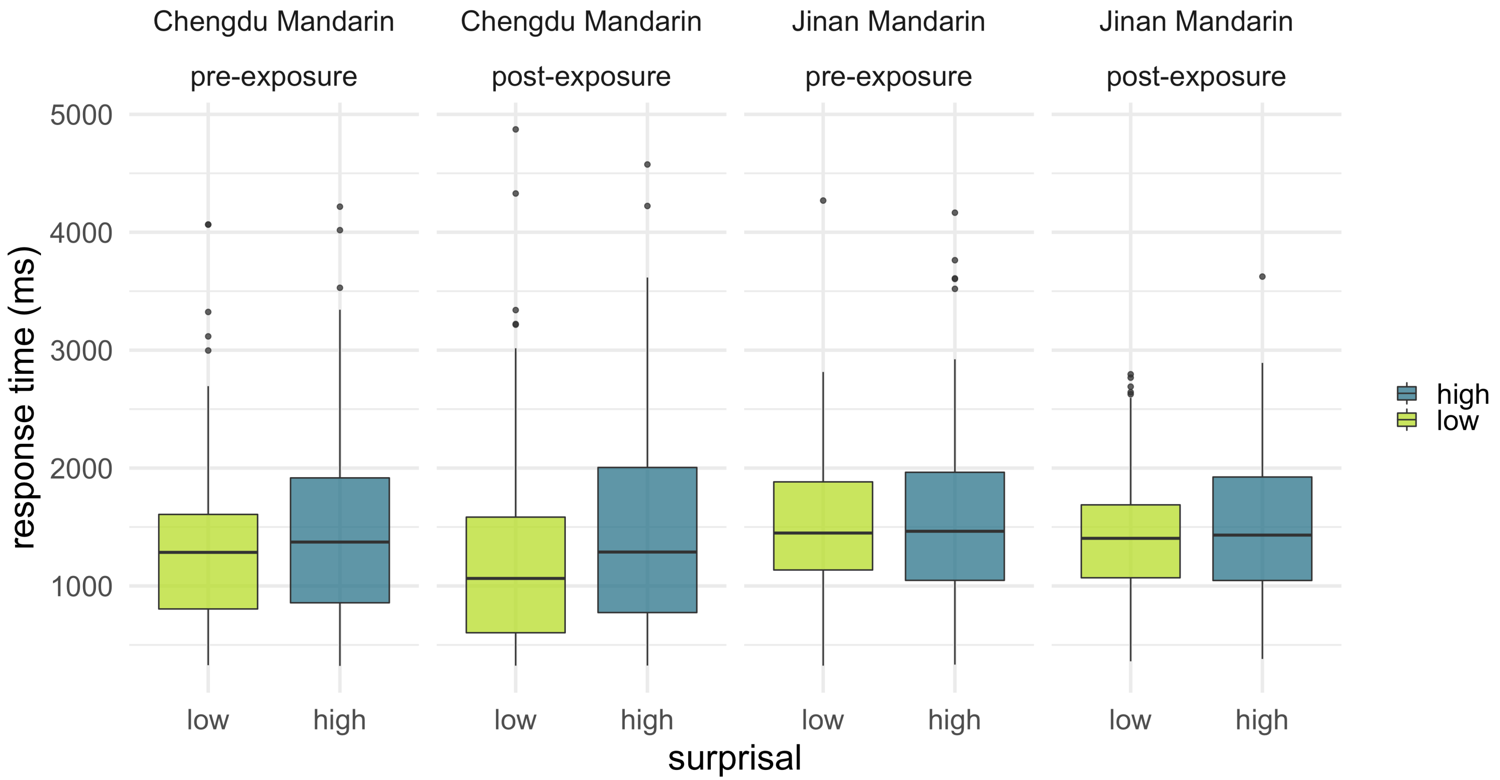


Figure: Response times across dialect, surprisal, and exposure conditions.

Response time

Credible main effects of surprisal, dialect, interaction between dialect and surprisal, and interaction of surprisal, dialect and exposure

- *Dialect x surprisal x exposure*: even slower responses for Chengdu high-surprisal sentences after exposure

❖ Explicit exposure facilitates discrimination between surprisal conditions more for Chengdu sentences than Jinan sentences

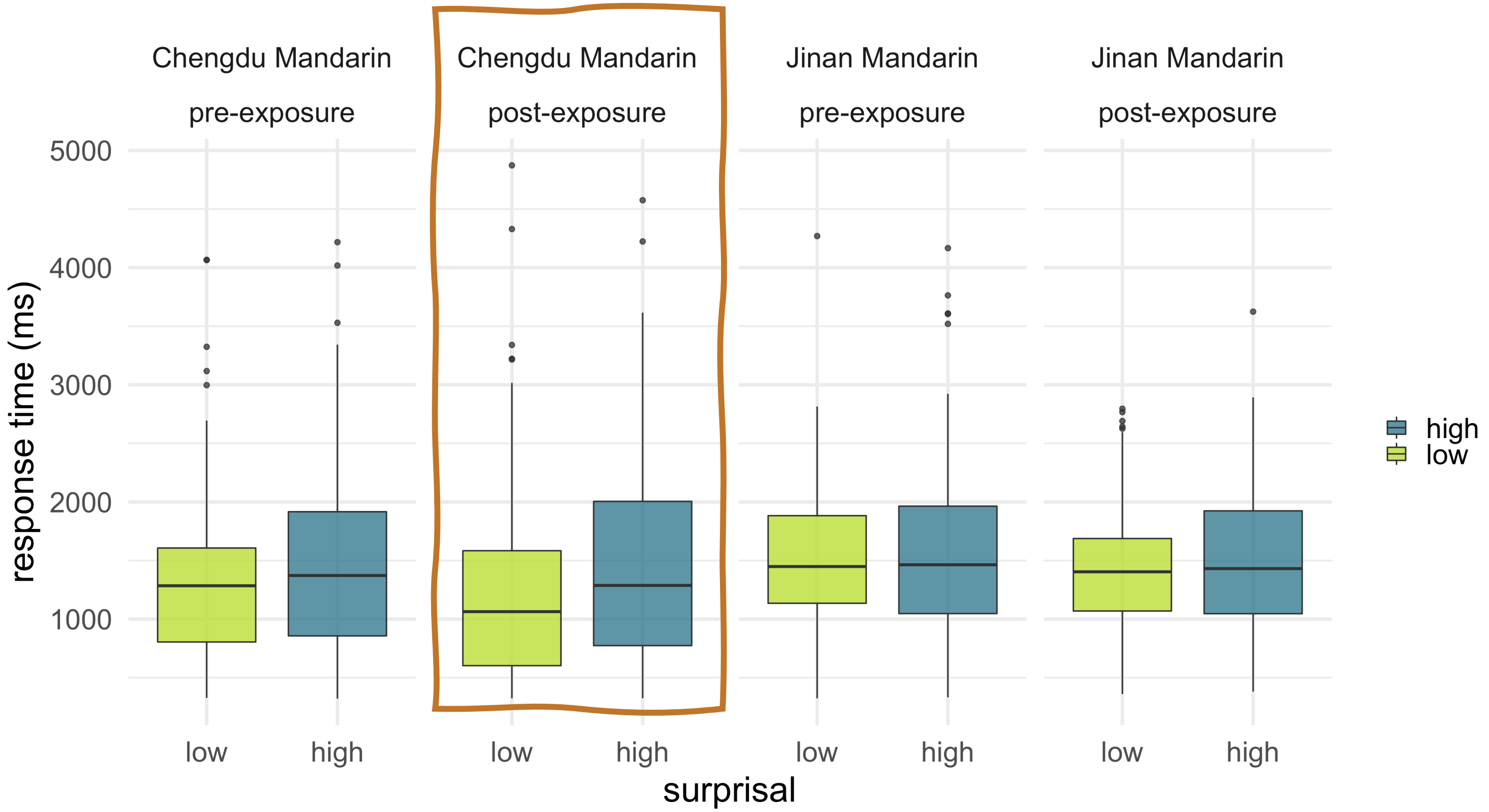


Figure: Response times across dialect, surprisal, and exposure conditions.

Response time

Credible main effect of tone comparison between tone 2 and tone 4

- Tone 2 vs. tone 4: faster responses to tone 2 in general

❖ Response times did not vary considerably across tone categories, except for faster responses for Tone 2 than Tone 4

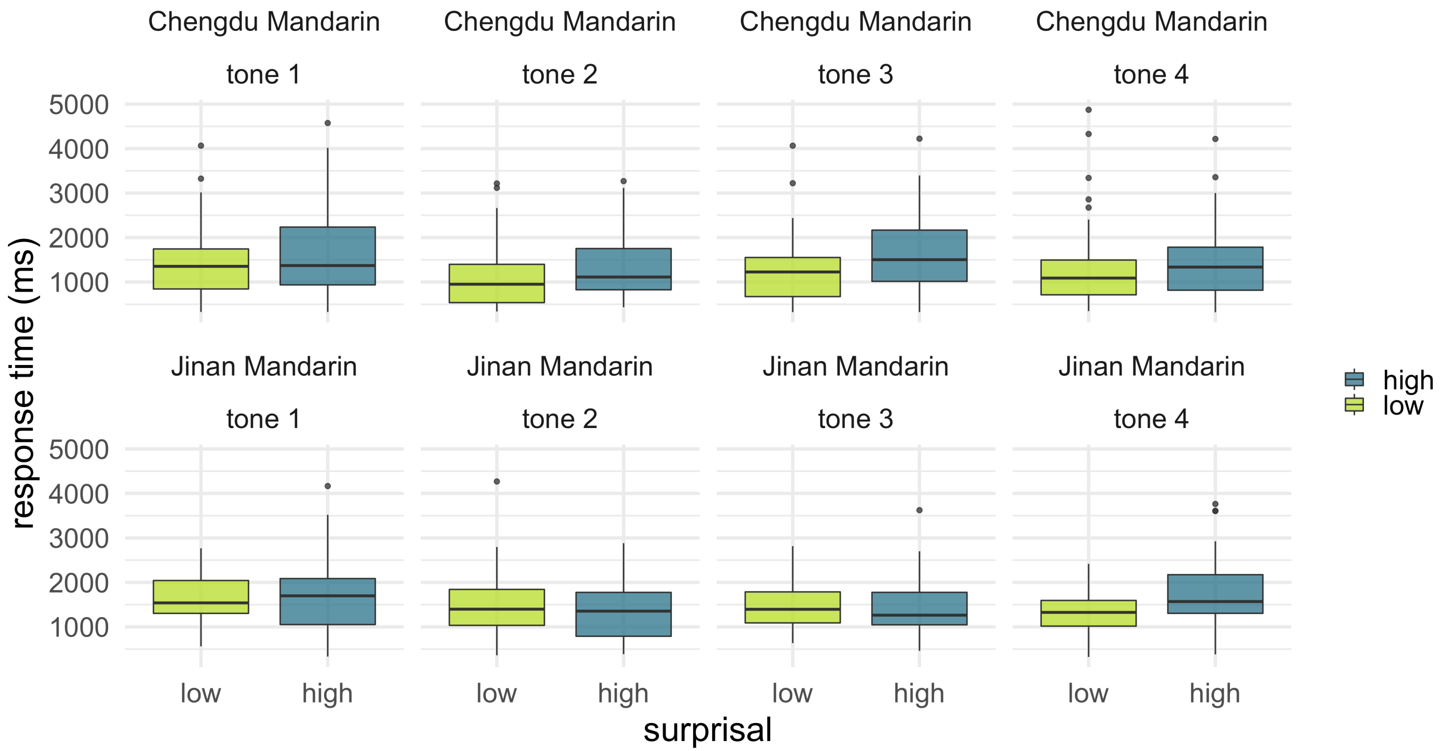


Figure: Response times across dialect, surprisal, and exposure conditions.

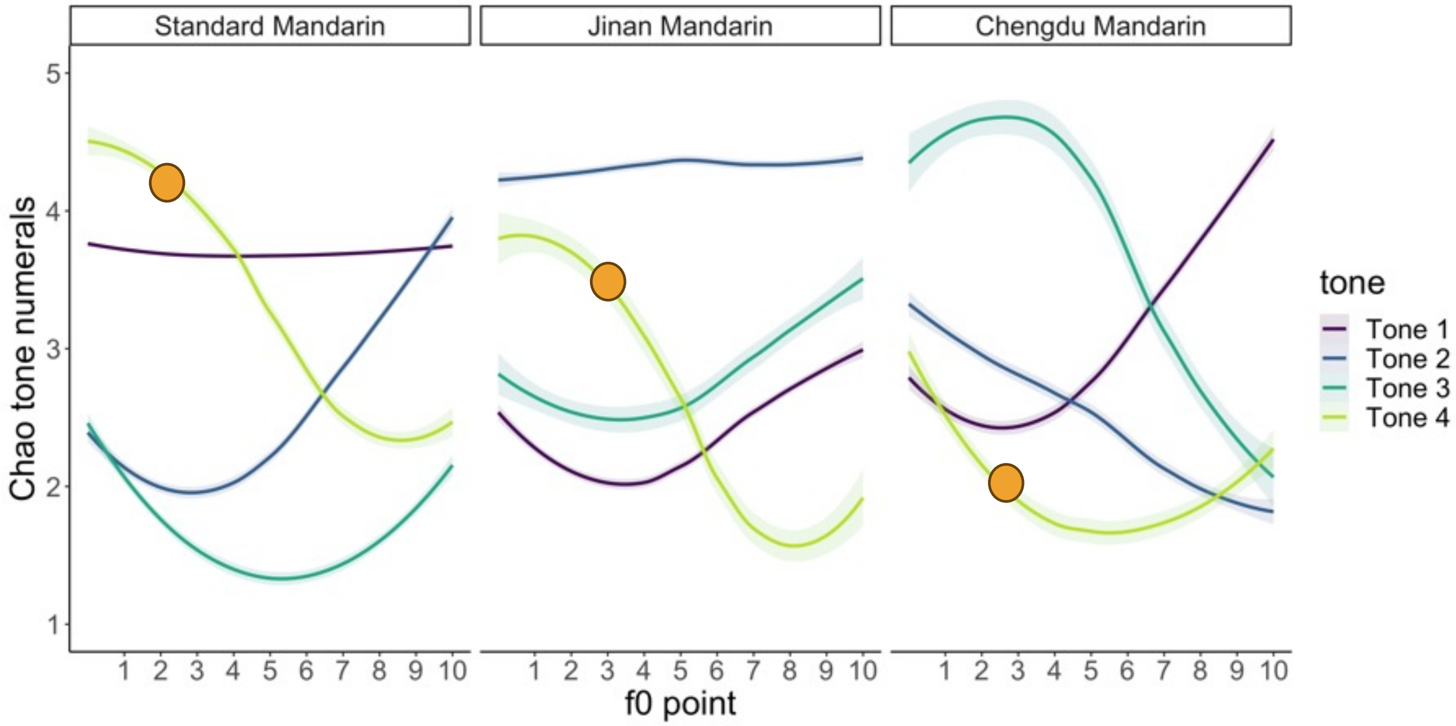
Response time

Credible main effect of tone comparison between tone 2 and tone 4

- Tone 2 vs. tone 4: slower responses to tone 4 in general

❖ Tone 4 has similar falling components across all three dialects

❖ Tone 2 has disparate contour types in both dialects compared to Standard Mandarin



Conclusions

Accuracy

- Strong use of top-down information for lexical access
- Possible use of a “benefit of the doubt” strategy (“I don’t think you meant to say something that strange”)
- Explicit exposure reliably improved adaptation to a novel tone system

Response time

- Sensitivity to tonal contrasts in the new tone system
- Listeners were more sensitive to the surprisal manipulation in Chengdu Mandarin (phonetically less similar) than in Jinan Mandarin (phonetically more similar)
- Tones with similar contour–category mappings across dialects slow down response times

Conclusions

Tone information always processed in lexical access even if it's not directly used

Implications for models of lexical access in tone languages (e.g., Gao et al. 2009)

More sensitivity to the less similar tone systems

Greater dissimilarity leads to better discrimination (Perceptual Assimilation Model, Best & Tyler, 2007; So & Best, 2011)

Thank you!

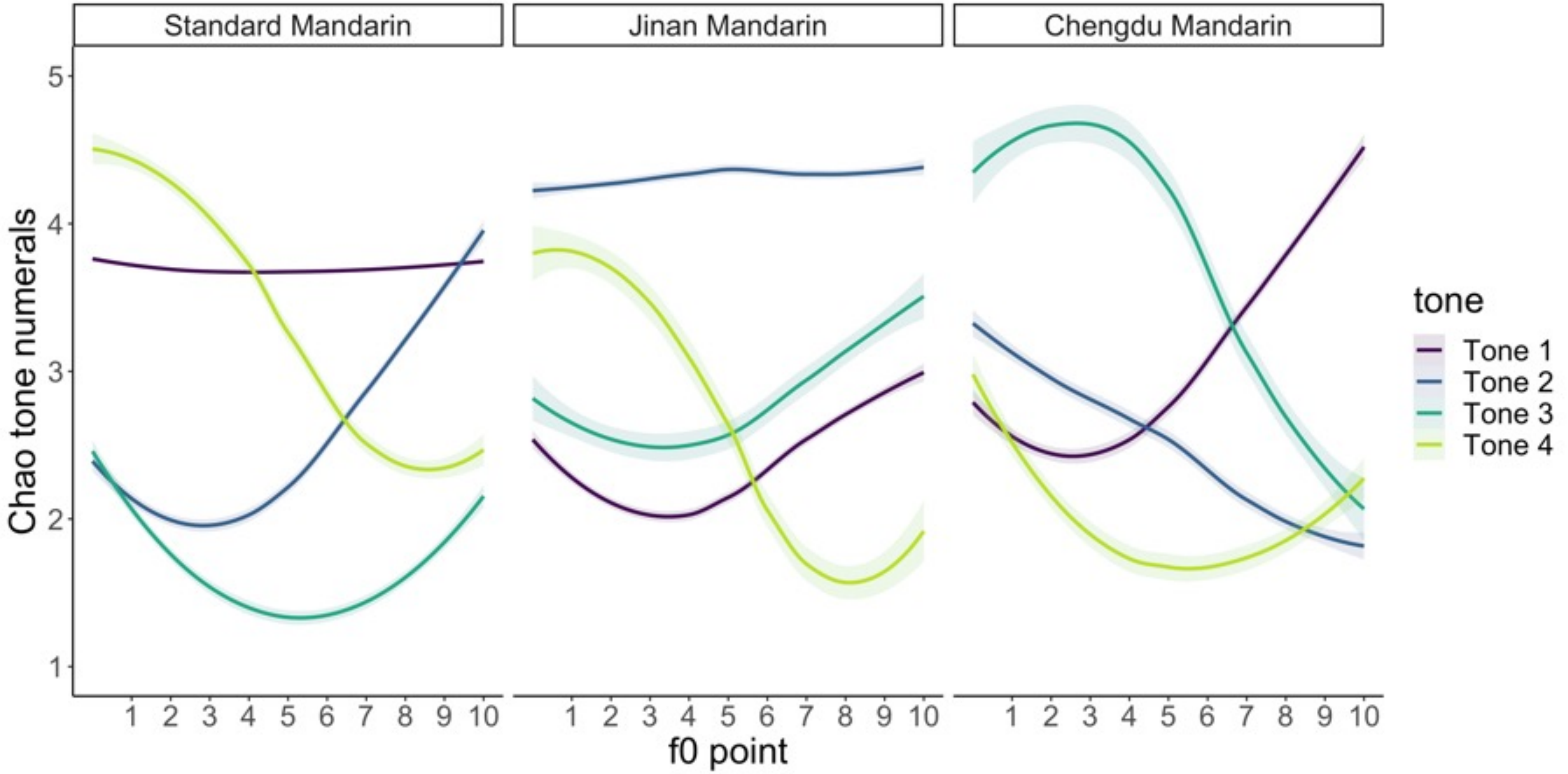
Thanks also to:

Shayne Sloggett, University of York

SNSF Grant 208460

No thanks to DHL UK

The word yi (/iː/, "one, single"; 一) in four tone categories



Surprisal sentences

24 sentence pairs were created manipulating the phonetic tone of one target word to be:

- **low-surprisal** (semantically implausible sentences)
- **high-surprisal** (semantically plausible sentence)*

	low-surprisal condition	**high-surprisal condition**
	a) 有 一只 鹰 在 天上 飞 *You3 yi4 zhi1 ying1 zai4 tian1 shang4 fei1* *There is an eagle in the sky flying* "There is an eagle flying in the sky"	b)* 有 一只 鹰 在 天上 肥* *You3 yi4 zhi1 ying1 zai4 tian1 shang4 fei2** *There is an eagle in the sky gaining weight** "There is an eagle gaining weight in the sky"
Chengdu	Chengdu low-surprisal	Chengdu high-surprisal
Jinan	Jinan low-surprisal	Jinan high-surprisal

Data analysis (models)

Accuracy: Bayesian logistic mixed-effects regression

Response time: Bayesian linear mixed-effects regression (Bürkner, 2018)

Each model included

- The fixed effects: surprisal, dialect, exposure and the interactions of the three factors, trial number, tone comparisons and their interaction with dialect
- The random effects:
    - For participant: an intercept for participant, slopes for surprisal, dialect, trial number, the interaction between surprisal and dialect
    - For sentence frame, an intercept and random slope for dialect

Weakly informed priors

- Accuracy model: $\mathcal{N}(0, 20)$ for the intercept and the fixed factors, $\mathcal{N}(0, 0.05)$ for the random effects
- RT model: $\mathcal{N}(0, 7)$ for the intercept, $\mathcal{N}(0, 1)$ for the fixed factors, $\mathcal{N}(0, 0.01)$ for the random effects

R codes: accuracy model

```
fit_acc_bayes <- brm(acc2 ~ nDialect + nSurprisal + nExposure +
                   nDialect:nSurprisal + nDialect:nExposure + nSurprisal:nExposure +
                   nDialect:nSurprisal:nExposure + nTrial + nTone1 + nTone2 + nTone3 +
                   nTone1:nDialect + nTone2:nDialect + nTone3:nDialect +
                   (1 + nSurprisal + nExposure + nTrial + nSurprisal:nExposure | subj)
                  + (1 + nSurprisal + nDialect | frame), prior = c(prior(normal(0, 20), class = Intercept),
                   prior(normal(0, 20), class = b, coef = nDialect), prior(normal(0, 20), class = b, coef = nSurprisal), prior(normal(0, 20), class = b, coef = nExposure),
                   prior(normal(0, 20), class = b, coef = nTone1), prior(normal(0, 20), class = b, coef = nTone2), prior(normal(0, 20), class = b, coef = nTone3),
                   prior(normal(0, 20), class = b, coef = nDialect:nSurprisal),
                   prior(normal(0, 20), class = b, coef = nDialect:nExposure),
                   prior(normal(0, 20), class = b, coef = nSurprisal:nExposure),
                   prior(normal(0, 20), class = b, coef = nDialect:nTone1), prior(normal(0, 20), class = b, coef = nDialect:nTone2), prior(normal(0, 20), class = b, coef = nDialect:nTone3),
                   prior(normal(0, 20), class = b, coef = nDialect:nSurprisal:nExposure),
                   prior(normal(0, 0.05), class = sd, group = subj, coef = Intercept), prior(normal(0, 0.05), class = sd, group = subj, coef = nSurprisal),
                   prior(normal(0, 0.05), class = sd, group = subj, coef = nExposure), prior(normal(0, 0.05), class = sd, group = subj, coef = nTrial),
                   prior(normal(0, 0.05), class = sd, group = subj, coef = nSurprisal:nExposure),
                   prior(normal(0, 0.05), class = sd, group = frame, coef = Intercept), prior(normal(0, 0.05), class = sd, group = frame, coef = nSurprisal),
                   prior(normal(0, 0.05), class = sd, group = frame, coef = nDialect)
                  ), family = "bernoulli", d)
```

R codes: RT model

```
fit_rt_bayes <- brm(rt ~ nDialect + nSurprisal + nExposure +
                    nDialect:nSurprisal + nDialect:nExposure + nSurprisal:nExposure +
                    nDialect:nSurprisal:nExposure + nTrial +
                    nTone1 + nTone2 + nTone3 +
                    nTone1:nDialect + nTone2:nDialect + nTone3:nDialect +
                    (1 + nSurprisal + nExposure + nTrial +
                      nSurprisal:nExposure | subj)
                  + (1 + nSurprisal + nDialect | frame), prior = c(
                    prior(normal(7, 1), class = Intercept),
                    prior(normal(0, 1), class = b, coef = nDialect),  prior(normal(0, 1), class = b, coef = nSurprisal),
                    prior(normal(0, 1), class = b, coef = nExposure), prior(normal(0, 1), class = b, coef = nTrial),
                    prior(normal(0, 1), class = b, coef = nTone1), prior(normal(0, 1), class = b, coef = nTone2),  prior(normal(0, 1), class = b, coef = nTone3),
                    prior(normal(0, 1), class = b, coef = nDialect:nSurprisal),
                    prior(normal(0, 1), class = b, coef = nDialect:nExposure),
                    prior(normal(0, 1), class = b, coef = nSurprisal:nExposure),
                    prior(normal(0, 1), class = b, coef = nDialect:nTone1),
                    prior(normal(0, 1), class = b, coef = nDialect:nTone2),
                    prior(normal(0, 1), class = b, coef = nDialect:nTone3),
                    prior(normal(0, 1), class = b, coef = nDialect:nSurprisal:nExposure),
                    prior(normal(0, 0.01), class = sd, group = subj, coef = Intercept),
                    prior(normal(0, 0.01), class = sd, group = subj, coef = nSurprisal),
                    prior(normal(0, 0.01), class = sd, group = subj, coef = nExposure),
                    prior(normal(0, 0.01), class = sd, group = subj, coef = nTrial),
                    prior(normal(0, 0.01), class = sd, group = subj, coef = nSurprisal:nExposure),
                    prior(normal(0, 0.01), class = sd, group = frame, coef = Intercept),
                    prior(normal(0, 0.01), class = sd, group = frame, coef = nSurprisal),
                    prior(normal(0, 0.01), class = sd, group = frame, coef = nDialect)
                  ), family = lognormal(), subset(d, rt > 0))
```

Data analysis: Does minimal-pair presentation matter?

Our previous study found that minimal-pair presentation of the stimuli did **not** credibly affect listeners' sensitivity to the surprisal manipulation (Zhao, Sloggett & Chodroff, ICPhS 2023)

For this study, we also ran the accuracy and RT models with *presentation* (with-minimal-pair design vs. no-minimal-pair design) as a fixed factor

- Accuracy model: **no difference** in accuracy between the two designs
- Response time model: **faster responses** when minimal pairs were presented.
    - But this might be just because listeners heard more trials (24 trials) in the with-minimal-pair experiment than the no-minimal-pair experiment (12 trials per participant)

Perception experiment: surprisal-based processing

Surprisal sentence pairs in sentence plausibility judgment task ("yes"/"no")

"Does this sentence make sense to you?"

24 sentence pairs were created manipulating the phonetic tone of one target word to be:

- **low-surprisal** (semantically implausible sentences)
- **high-surprisal** (semantically plausible sentence)*

low-surprisal condition	high-surprisal condition
a) 有 一只 鹰 在 天上 飞 *You3 yi4 zhi1 ying1 zai4 tian1 shang4 fei1* *There is an eagle in the sky flying* "There is an eagle flying in the sky"	b)* 有 一只 鹰 在 天上 肥* *You3 yi4 zhi1 ying1 zai4 tian1 shang4 fei2** *There is an eagle in the sky gaining weight** "There is an eagle gaining weight in the sky"