Prosodic entrainment in Mandarin and English :
A cross-linguistic comparison

Zhihua Xia, Rivka Levitan, Julia Hirschberg

15th, March, 2019

Main contents

Greetings

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

      1.1 Research questions

      1.2 Tongji Games Corpus

      1.3 Crosss-linguistic comparison

      1.4 Some questions for you about entrainment over tone units

2. The work we are doing--- AutoToBI for Chinese corpus

Greetings

江苏师范大学
JIANGSU NORMAL UNIVERSITY

Associate Professor

同济大学
TONGJI UNIVERSITY

PhD in linguistics

Greetings

My first visiting during 2013-2014

My second visiting during 2018-2019

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.1 Research question:

Do the people from different language groups show similar or different pattern of prosodic entrainment?

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison
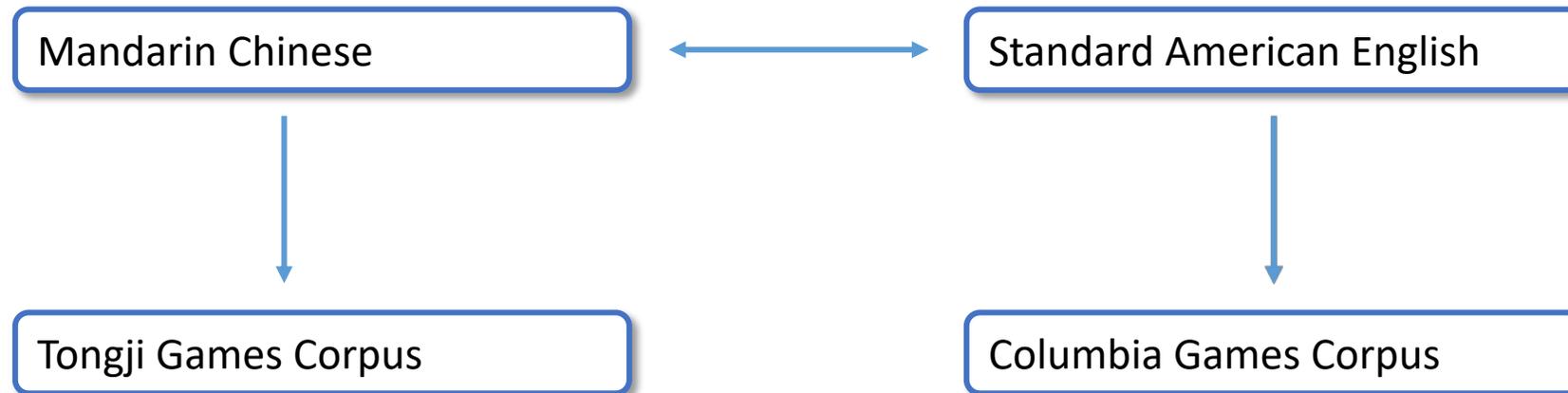
1.1 Research question:

Do the people from different language groups show similar or different pattern of prosodic entrainment?

| Mandarin Chinese | ←——————→ | Standard American English |

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.1 Research question:

Do the people from different language groups show similar or different pattern of prosodic entrainment?

```
Mandarin Chinese  <----->  Standard American English
       |                              |
       v                              v
Tongji Games Corpus        Columbia Games Corpus
```
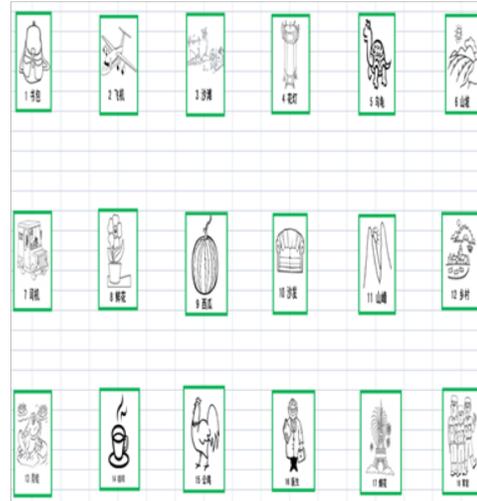
1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison
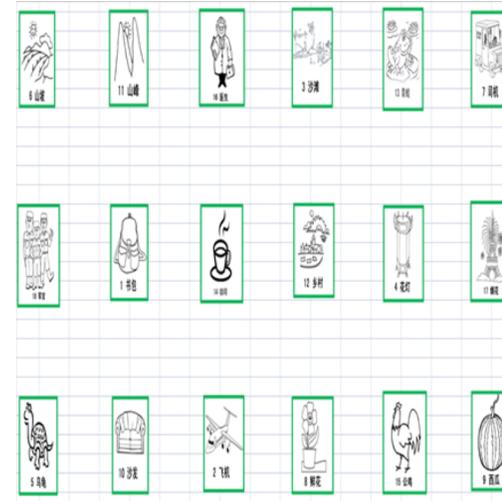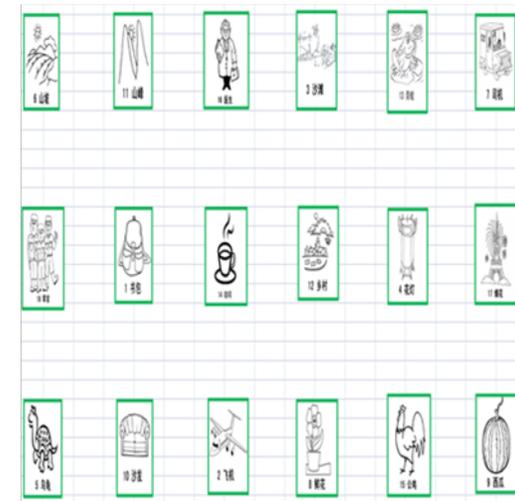
1.2 Tongji Games Corpus:

Picture ordering game
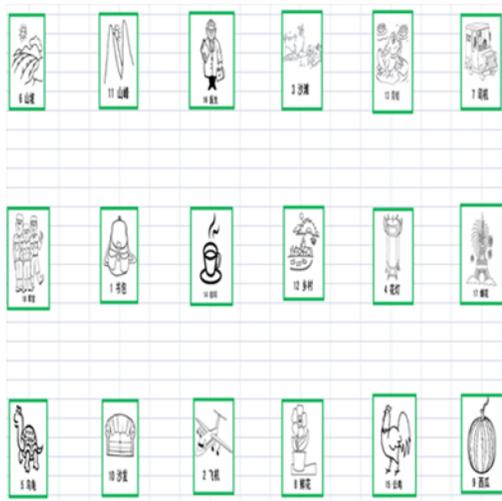
Picture classifying game

giver

follower

speaker A

speaker B

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.2 Tongji Games Corpus:

- 70 pairs (23 female and female pairs

    30 female and male pairs

    17 male and male pairs)
- 118 conversations (60 from Picture ordering games

    58 from Picture classifying games)
- The average duration 6 minutes

    the whole corpus about 12 hours spontaneous,

    task- oriented conversations
- Carrier sentences for target tone units

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.2 Tongji Games Corpus:

Gender group control

Role control

Tone units control

- 70 pairs (23 female and female pairs

    30 female and male pairs

    17 male and male pairs)

- 118 conversations (60 from Picture ordering games

    58 from Picture classifying games)

- The average duration 6 minutes

    the whole corpus about 12 hours spontaneo

    task- oriented conversations

- Carrier sentences for target tone units

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.2 Tongji Games Corpus: (method)

- IPU by SPPAS (Bigi & Hirst 2012)

- Prosodic parameters:

  F0min, F0 mean, F0max

  Intensitymin, Intensitymean, Intensitymax

  Speaking rate

- Three aspects of entrainment (Levitan & Hirschberg 2011; Levitan et al. 2012)

  Proximity

  convergence

  synchrony
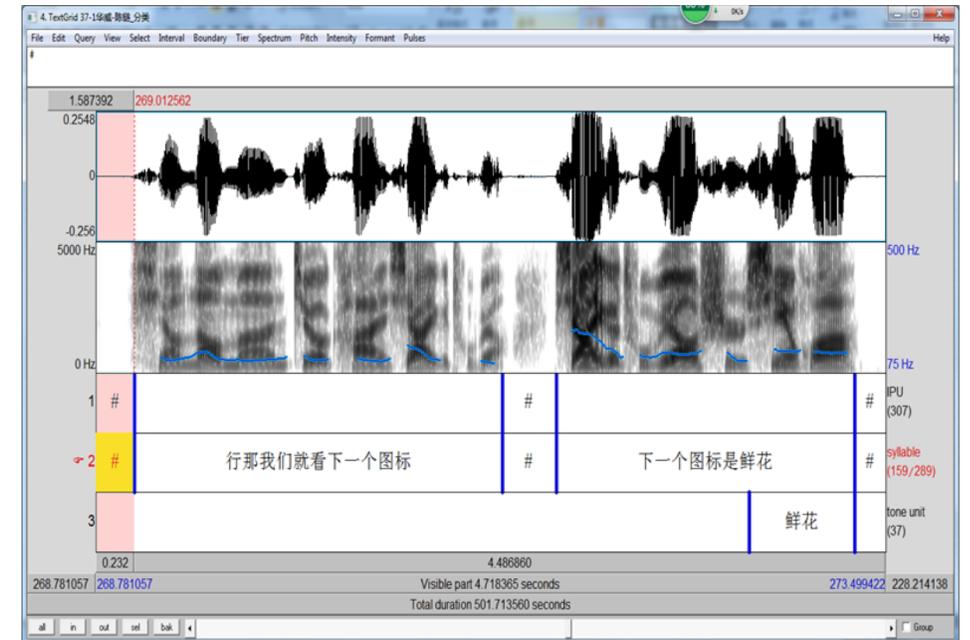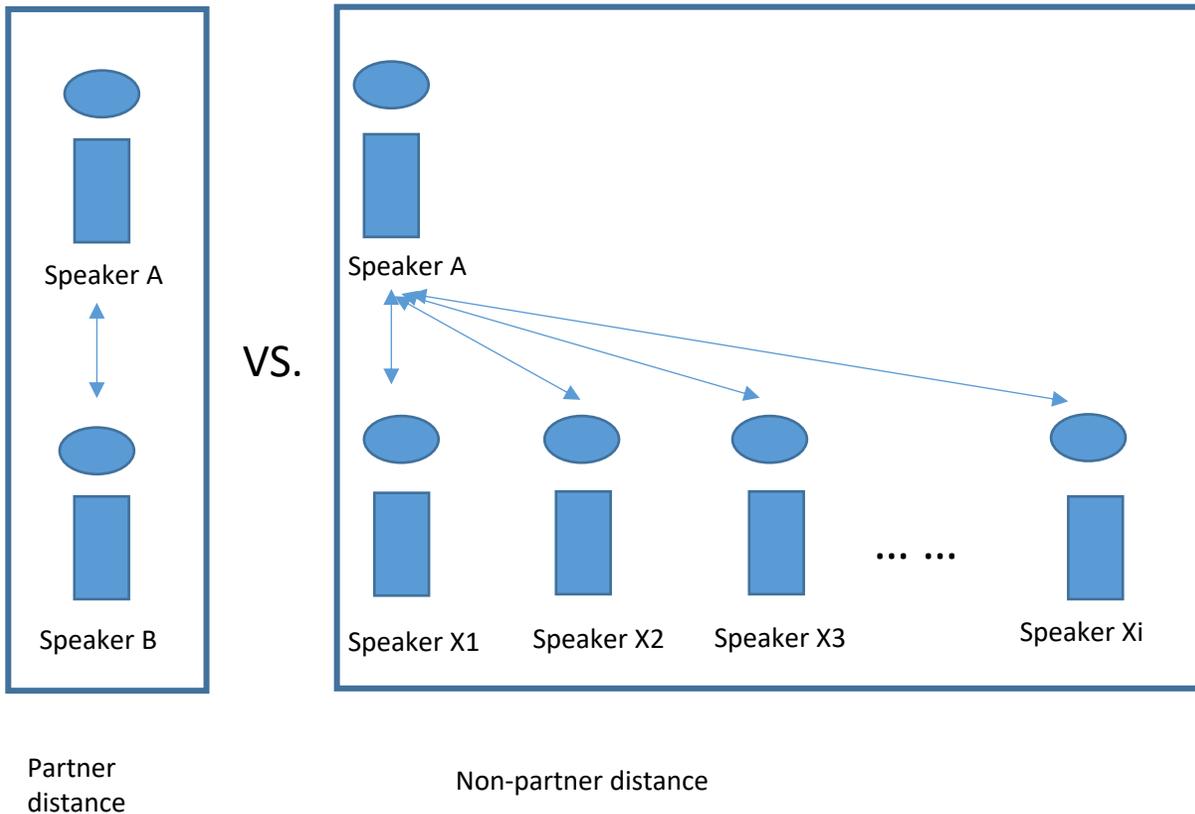
1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.2 Tongji Games Corpus: (method)

Speaker A

Speaker B

Partner distance

VS.

Speaker A

Speaker X1   Speaker X2   Speaker X3   ... ...   Speaker Xi

Non-partner distance

（Levitan & Hirschberg 2011）

- disp = Partner distance of A and B
- disnp = Non-partner distance of A and X(X1, X2 , X3……Xi)

X = set of speakers of the same gender and role as A's partner B who are never paired with A

- Hypothesis:

$$disp \quad < \quad disnp$$

$$disp = \left| A_f - B_f \right| \quad < \quad disnp \ = \frac{\sum_i \left| A_f - X_{i\,f} \right|}{|X|}$$

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.2 Tongji Games Corpus: (method)

Speaker A

Speaker B

Partner distance

VS.

Speaker A

Speaker X1　Speaker X2　Speaker X3　　　　Speaker Xi

… …

Non-partner distance

（Levitan & Hirschberg 2011）

Global
Local
Gender

- disp ... A and B
- dis ... tance of A and X(X1, ..., X3......Xi)

X = set of speakers of the same gender and role as A's partner B who are never paired with A

- Hypothesis:
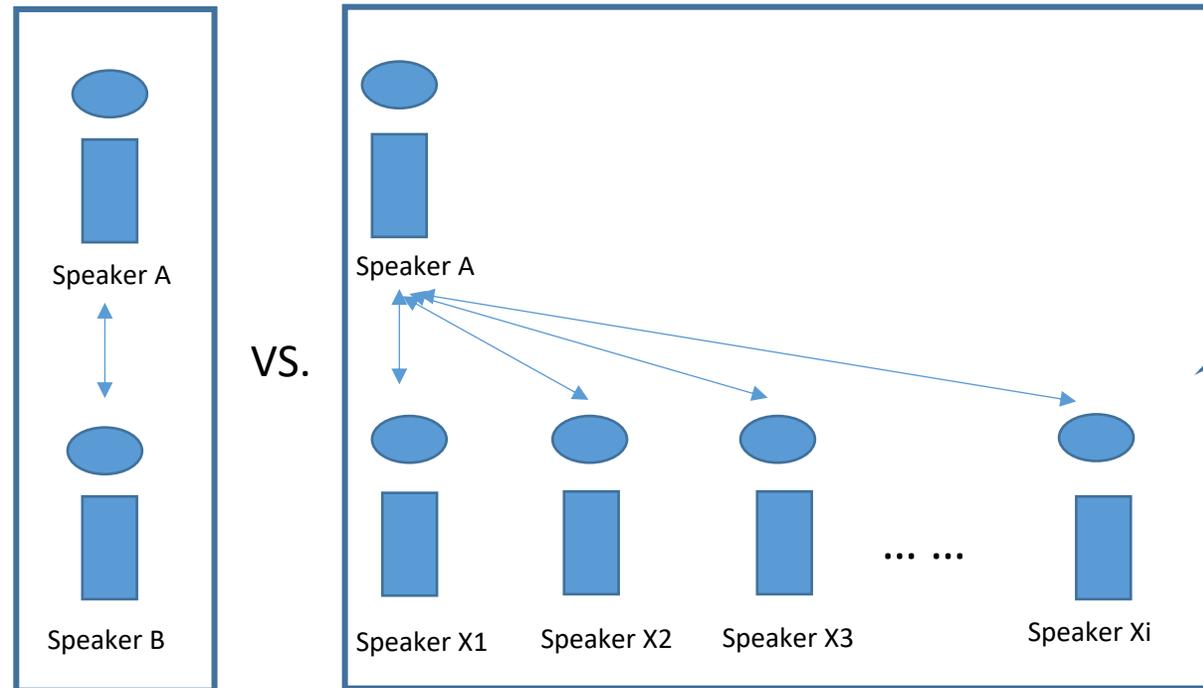
  disp < disnp

$$disp = \left| A_f - B_f \right| \quad < \quad disnp = \frac{\sum_i \left| A_f - Xi_f \right|}{|X|}$$

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese

- Global or local?

- Exact or relative?

- Convergent or constant?

Feature	Global proximity	Local proximity	Synchrony	Global convergence	Local convergence
Intensity mean	✓	✓	0.63		
Intensity max	✓	✓	0.55 *(0.47)*		
Intensity min			0.31 --		0.09
F0 mean			0.66 *(0.28)*		0.22
F0 max	✓		0.61 *(0.18)*		0.24
F0 min			0.63 --		0.19
Speaking rate	✓	✓	*(0.15)*		

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?

- Exact or relative?

- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)			(0.02)
Intensity min		--		--	0.31	--	--	0.09	--
F0 mean				(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)		0.24	(0.05)
F0 min		--		--	0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?
- Exact or relative?
- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)			(0.02)
Intensity min		--		--	0.31	--	--	0.09	--
F0 mean				(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)		0.24	(0.05)
F0 min		--		--	0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?
- Exact or relative?
- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)			(0.02)
Intensity min		--		--	0.31	--	--	0.09	--
F0 mean				(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)		0.24	(0.05)
F0 min		--		--	0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?
- Exact or relative?
- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)			(0.02)
Intensity min		--		--	0.31	--	--	0.09	--
F0 mean				(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)		0.24	(0.05)
F0 min		--		--	0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?
- **Exact or relative?**
- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)			(0.02)
Intensity min		--		--	0.31	--	--	0.09	--
F0 mean				(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)		0.24	(0.05)
F0 min		--		--	0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?
- **Exact or relative?**
- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)			(0.02)
Intensity min		--		--	0.31	--	--	0.09	--
F0 mean				(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)		0.24	(0.05)
F0 min		--		--	0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?
- Exact or relative?
- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)			(0.02)
Intensity min	--		--		0.31	--	--	0.09	--
F0 mean				(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)		0.24	(0.05)
F0 min			--		0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?

- **Exact or relative?**

- Convergent or constant?

Feature	Global proximity		Local proximity	Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓ (✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓ (✓)	0.55	(0.47)			(0.02)
Intensity min		--	--	0.31	--	--	0.09	--
F0 mean			(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)	(✓)	0.61	(0.18)		0.24	(0.05)
F0 min		--	--	0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓ (✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?
- Exact or relative?
- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)			(0.02)
Intensity min		--		--	0.31	--	--	0.09	--
F0 mean				(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)		0.24	(0.05)
F0 min		--		--	0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?

- Exact or relative?

- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence		Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)		(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)				(0.02)
Intensity min		--		--	0.31	--		--	0.09	--
F0 mean				(✓)	0.66	(0.28)		(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)			0.24	(0.05)
F0 min		--		--	0.63	--		--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)				(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?
- Exact or relative?
- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence		Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)		(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)				(0.02)
Intensity min		--		--	0.31	--		--	0.09	--
F0 mean				(✓)	0.66	(0.28)		(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)			0.24	(0.05)
F0 min		--		--	0.63	--		--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)				(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?
- Exact or relative?
- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)			(0.02)
Intensity min		--		--	0.31	--	--	0.09	--
F0 mean				(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)		0.24	(0.05)
F0 min		--		--	0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?
- Exact or relative?
- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)			(0.02)
Intensity min		--		--	0.31	--	--	0.09	--
F0 mean				(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)		0.24	(0.05)
F0 min		--		--	0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)

- Global or local?

- Exact or relative?

- Convergent or constant?

Feature	Global proximity		Local proximity		Synchrony		Global convergence	Local convergence	
Intensity mean	✓	(✓)	✓	(✓)	0.63	(0.50)	(✓)		(0.03)
Intensity max	✓	(✓)	✓	(✓)	0.55	(0.47)			(0.02)
Intensity min		--		--	0.31	--	--	0.09	--
F0 mean				(✓)	0.66	(0.28)	(✓)	0.22	(0.06)
F0 max	✓	(✓)		(✓)	0.61	(0.18)		0.24	(0.05)
F0 min		--		--	0.63	--	--	0.19	--
Speaking rate	✓	(✓)	✓	(✓)		(0.15)			(0.03)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese & *SAE*

Mandarin Chinese

- Evidence of entrainment on three main aspects of prosody

- Local entrainment is more evident than the global entrainment.

- Different prosodic features exhibit entrainment at different level.

- Strong evidence of synchrony, especially F0

- Some local convergence

Comparisons with SAE

- Proximity
  Global proximity: same as SAE
  Local proximity: not F0
- Convergence
  No global convergence
  Stronger local convergence.
- Synchrony
  Stronger synchrony

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese & *SAE*

Mandarin Chinese

- Evidence of entrainment on three main aspects of prosody
- Local entrainment is more evident than the global entrainment.
- Different prosodic features exhibit entrainment at different level.
- Strong evidence of synchrony, especially F0
- Some local convergence

Comparisons with SAE

- Proximity
  Global proximity: same as SAE
  Local proximity: not F0
- Convergence
  No global convergence
  Stronger local convergence.
- Synchrony
  Stronger synchrony

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)------ entrainment and gender

- The male dominance hypothesis emphasizes men's dominant position in society.
- CAT states that when there is an imbalance of power between speakers, the less dominant speaker will converge more.
- Females are known to be more sensitive to perceived differences in speech.
- Perception behavior link states perceiving something makes people more likely to mimic it.

- Predictions made by these theories: in mf conversations, females should entrain more. There should be more entrainment in FF conversations than MM conversations.

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)------ entrainment and gender

$$ENT_p = -\left|\text{speaker1}_f - \text{speaker2}_f\right|$$

$$ENT_x = -\frac{\sum_i \left|\text{speaker1}_f - X_{i,f}\right|}{|X|}$$

X = set of speakers of the same gender and role as speaker1's partner who are never paired with speaker1.

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)------ entrainment and gender

Feature	FF		MM		FM	
	MC	*SAE*	*MC*	*SAE*	*MC*	*SAE*
Intensity mean	✓	✓		✓	✓	✓
Intensity max	✓	✓		✓	✓	✓
Intensity min		--		--		--
F0 mean					✓	✓
F0 max					✓	✓
F0 min		--		--		--
Speaking rate	✓	✓	✓	✓	✓	✓

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)------ entrainment and gender

Feature	FF		MM		FM	
	MC	*SAE*	*MC*	*SAE*	*MC*	*SAE*
Intensity mean	✓	✓		✓	✓	✓
Intensity max	✓	✓		✓	✓	✓
Intensity min		--		--		--
F0 mean					✓	✓
F0 max					✓	✓
F0 min		--		--		--
Speaking rate	✓	✓	✓	✓	✓	✓

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)------ entrainment and gender

Feature	*FF*		*MM*		*FM*	
	MC	*SAE*	*MC*	*SAE*	*MC*	*SAE*
Intensity mean	✓	✓		✓	✓	✓
Intensity max	✓	✓		✓	✓	✓
Intensity min		--		--		--
F0 mean					✓	✓
F0 max					✓	✓
F0 min		--		--		--
Speaking rate	✓	✓	✓	✓	✓	✓

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)------ entrainment and gender

- MM << FF << MF

- Similar patterns for Mandarin Chinese (MC) and Standard American English (SAE)

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.3 Cross-linguistic comparison: Chinese (*SAE*)------ conclusions

- Entrainment in MC on three main aspects of prosody

- Strong evidence of dynamic entrainment, especially on F0

- Gender differences: MM << FF << FM

- Consistencies in pattern with SAE

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.4 More work on entrainment over tone units

Chinese →   tone language
Tones interact with intonation in speaking

How does the prosodic entrainment realize over tone units?

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison
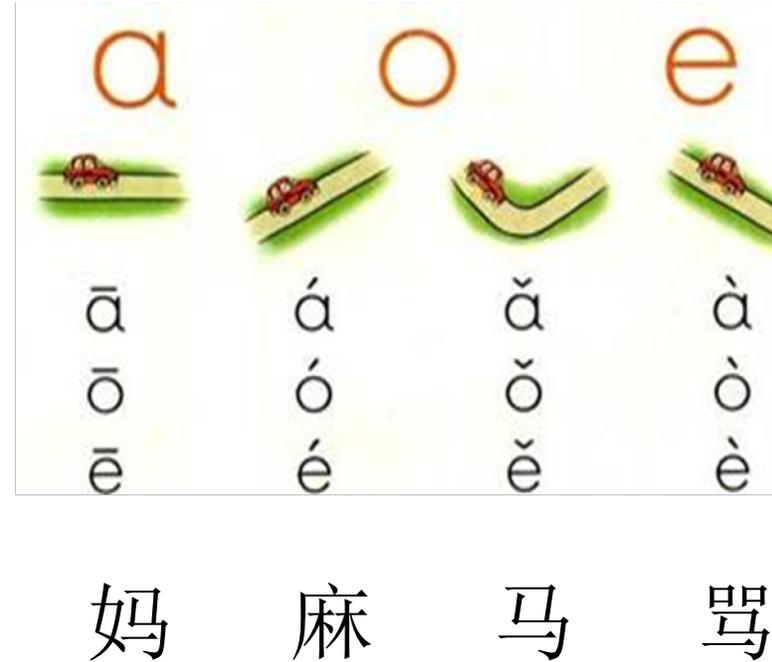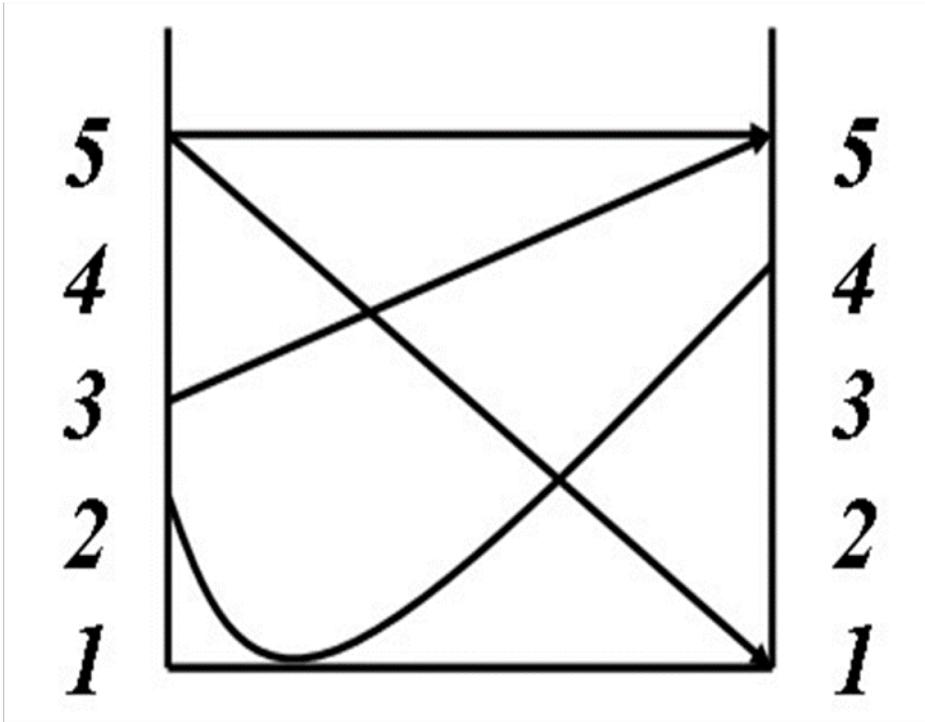
1.4 More work on entrainment over tone units

ā ó ǎ à
ō ó ǒ ò
ē é ě è
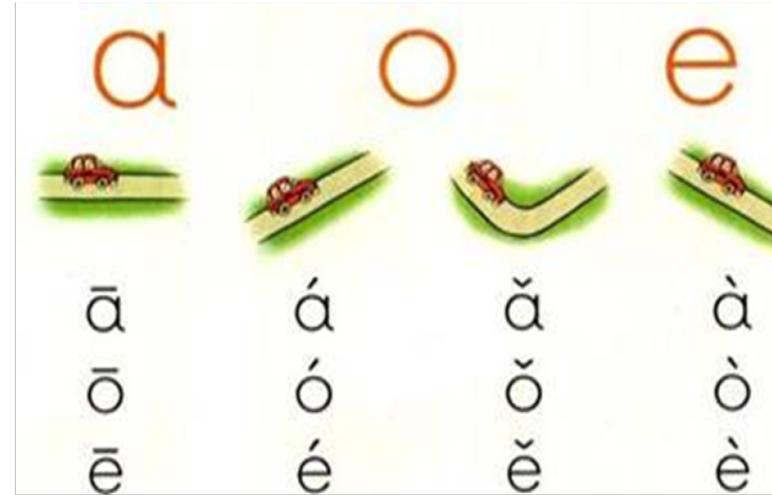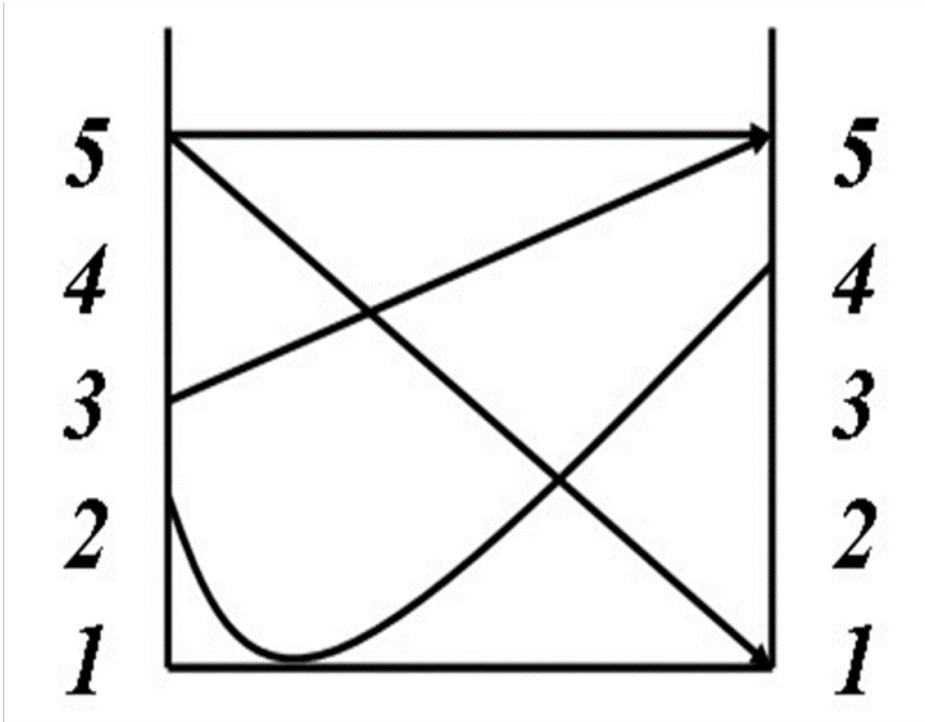
妈　麻　马　骂

4 Mandarin tones
Tones carry meaning

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.4 More work on entrainment over tone units

妈　麻　马　骂

How does the prosodic entrainment realize over tone units?

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.4 More work on entrainment over tone units

第一个图标是书包(shū bāo)。

(the first picture is bag.)

下一个图标是飞机(fēi jī)。

(the next picture is bag.)

下一个图标是山峰(shān fēng)。

(the first picture is mountain.)

......

How does the prosodic entrainment realize over tone units?

1. Prosodic entrainment in Mandarin and English : A cross-linguistic comparison

1.4 More work on entrainment over tone units
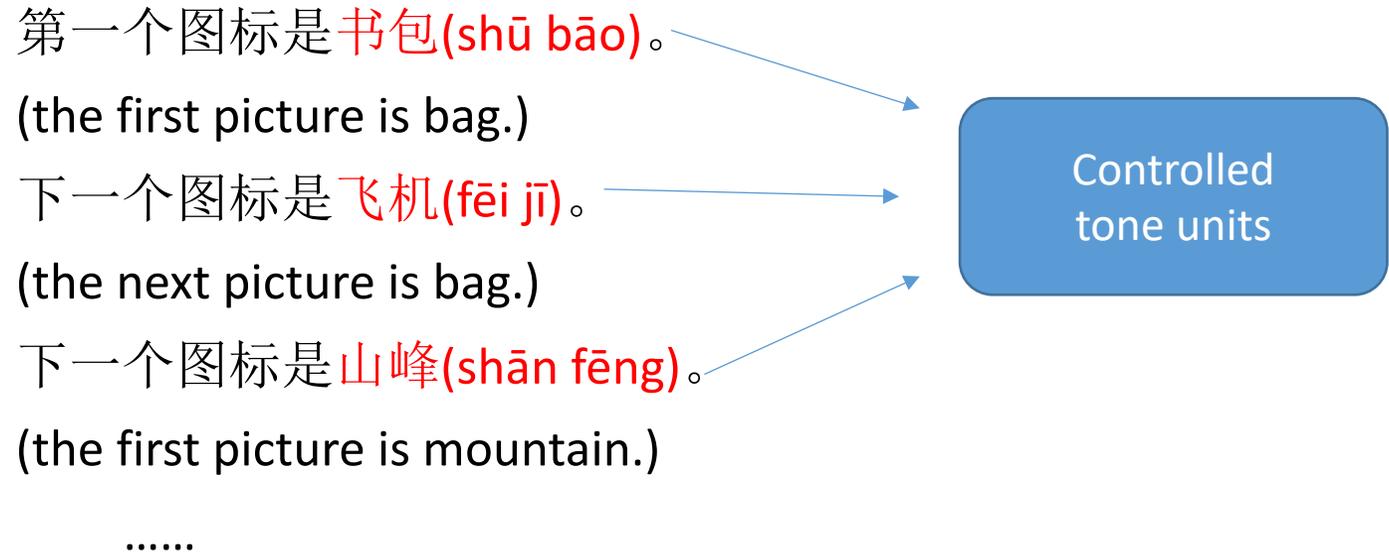
第一个图标是书包(shū bāo)。

(the first picture is bag.)

下一个图标是飞机(fēi jī)。

(the next picture is bag.)

下一个图标是山峰(shān fēng)。

(the first picture is mountain.)

……

Controlled tone units

How does the prosodic entrainment realize over tone units?

2. AutoToBI for Chinese corpus

The importance of labeling in corpus in TTS & linguistics

AutoToBI for English (Columbia Games Corpus) (Rosenberg 2009)

AutoToBI for Mandarin in phrase and stress prediction
(trained by DUR & COSPRO

applied in Tongji Games Corpus)

Thanks for your attention!

谢谢聆听！

xzhlf@163.com