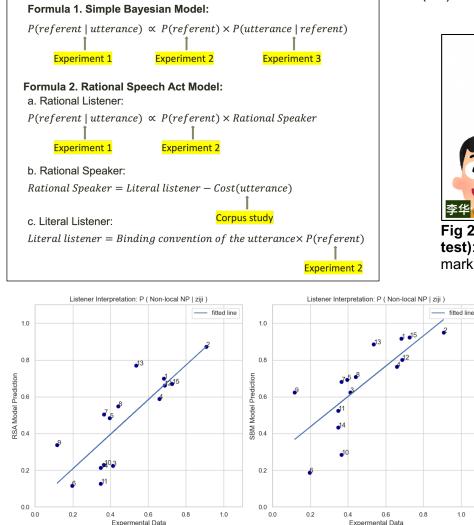**Probabilistic Listener: A Case of Reflexive *ziji "self"* Ambiguity Resolution in Mandarin**
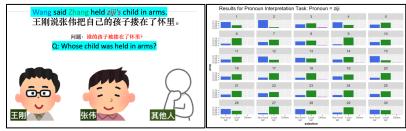
Fengyue Zhao[1,2], Brian Dillon[2], Ming Xiang[3]
[1] Cornell University, [2] University of Massachusetts Amherst, [3] University of Chicago

What mechanisms do people use to resolve ambiguous pronouns? Prior studies have tried to identify factors that contribute to the resolution mechanisms, such as first-mention biases [1], verb semantics [2], and world knowledge [3]. Probabilistic models are another promising attempt [4][5][6][7]. This study aims to test the generality of two Bayesian models — the Simple Bayesian model (SBM) and the Rational Speech Act model (RSA) — on pronoun ambiguity resolution in Mandarin Chinese in the case of the reflexive pronoun *ziji* "self". SBM has been examined to make good quantitative predictions for Chinese, but only for across-sentence relations in discourses with personal pronouns [7]. RSA has been examined in pronoun resolution in French and English [6]. This study provides cross-linguistic evidence to support both Bayesian models, investigating within-sentence relations which are often considered to be regulated by a distinct set of grammatical principles than discourse anaphors.

The resolution of reflexive *ziji*: Reflexive in Mandarin may take a simplex form — *ziji* "self" which permits both local (e.g. Ming in 1) and non-local (e.g. Hong in 1) referent interpretation [8]. In **Experiment 1** (N_{subj}=135, N_{item}=30, Fig.1), we measured Mandarin speakers' preferences for resolving *ziji* in ambiguous sentences with an antecedent selection task. Stimuli have the sentence structure in (1), differing only in the verb. We found that comprehenders preferred local antecedents 59% of the time on average. Moreover, the preference of non-local antecedents ranged between 9% and 88% across items. This non-uniform result showed that preferences and extent of preferences differ across stimuli items. Therefore, item-by-item quantitative analysis is necessary to test if two models can capture this item variation.
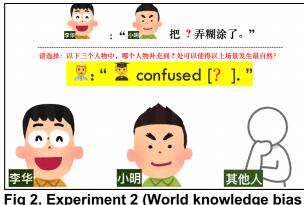
Two Bayesian Models: The **Simple Bayesian model (SBM) models** the listener's probability of selecting a specific referent as proportional to their prior that this referent will be mentioned next, and the likelihood that a speaker will produce this pronoun when signaling a specific referent (see Formula 1) [5]. The **Rational Speech Act model (RSA)** suggests that listeners assume speakers are rational agents who have already chosen the best utterance among all possible options to convey intended information. Listeners combine this recursive thinking with their prior world knowledge to interpret ambiguous pronouns (see Formula 2) [4]. To evaluate these models' fit against our data, we estimated: P(referent) in both models in **Experiment 2**, a world knowledge bias test (N_{subj}=28, N_{item}=30, Fig.2); P(utterance | referent) in SBM in **Experiment 3**, a pronoun production task (N_{subj}=65, N_{item}=30, Fig.3); Cost(utterance) in RSA in a **corpus study** (N_{token}=16.5 billion [9]), which is the logarithm of the frequency of each pronoun in a certain sentence structure, penalizes the speaker from producing low frequency pronoun. An interesting contrast is that *ziji* "self" is ambiguous but has lower cost than the unambiguous *taziji "himself"*.

Results and Discussions: If listeners processing ambiguous pronouns follows SBM/RSA, there should be a strong correlation between the Experiment 1 results and SBM/RSA predictions. Item-by-item quantitative analysis in Fig.4 shows that both SBM and RSA can make accurate predictions for the resolution of the ambiguous pronoun *ziji* (R^2>0.59, p<0.001), providing a case study to support that within-sentence reflexive binding obeys principles of Bayesian inference too. Meanwhile, although both SBM and RSA are Bayesian models, the statistical results in Table 1 show that RSA performs a little better while SBM overestimates the non-local antecedent choices and underestimates the local antecedent choices. The difference between two models is that listeners in SBM reason about the production of the pronoun directly using their own experience without encoding explicit the frequency of pronouns, while listeners in RSA are rational and reason indirectly by thinking about how a rational speaker would choose pronouns, and this rational speaker would explicitly take the frequency of pronouns into account (the Cost term in Formula 2). One possible explanation for this is that a multilevel recursive reasoning between listeners and speakers coupled with the explicitly encoded frequency of pronouns could enhance model predictions, leading to a better fit between RSA and the experimental data.

(1) [小红]ᵢ 说 [小明]ⱼ 总 把 [自己]ᵢ/ⱼ 弄糊涂。

[Hong]ᵢ says [Ming]ⱼ always BA [ziji]ᵢ/ⱼ confuses.

'[Hong]ᵢ says that [Ming]ⱼ always confuses [self]ᵢ/ⱼ.'

Fig 1. Experiment 1 (antecedent selection task): Sample item (left) and item-by-item results (right).

Formula 1. Simple Bayesian Model:

$$P(referent \mid utterance) \propto P(referent) \times P(utterance \mid referent)$$

Experiment 1   Experiment 2   Experiment 3

Formula 2. Rational Speech Act Model:

a. Rational Listener:

$$P(referent \mid utterance) \propto P(referent) \times Rational\ Speaker$$

Experiment 1   Experiment 2

b. Rational Speaker:

$$Rational\ Speaker = Literal\ listener - Cost(utterance)$$

Corpus study

c. Literal Listener:

$$Literal\ listener = Binding\ convention\ of\ the\ utterance \times P(referent)$$

Experiment 2

Fig 2. Experiment 2 (World knowledge bias test): choose a character to fill the question mark based on world knowledge.

Fig 4. Fit RSA model predictions (left) and SBM model predictions (right) with experimental data from Experiment 1 on an item-by-item basis.

Fig 3. Experiment 3 (pronoun production task): select a pronoun from a drop-down menu to convey given information.

	R-squared	MSE	MAE	P-value
SBM	0.598	0.025	0.138	0.0007
RSA	**0.674**	**0.006**	**0.048**	**0.0002**

MSE = Mean squared error, the lower the better
MAE = Mean absolute error, the lower the better

Table 1. R-squared, MSE, MAE, and P-value of the two models for model evaluations.

References:

[1] Järvikivi, J., van Gompel, R. P., Hyönä, J., & Bertram, R. (2005). Psychological Science. [2] Koornneef A. W., & van Berkum J. J. A. (2006). Journal of Memory and Language. [3] Hobbs, J. R. (1979). Cognitive Science. [4] Frank, M. C., & Goodman, N. D. (2012). Science. [5] Kehler, A, & Rohde, H. (2013). Theoretical Linguistics. [6] Schulz, M., Burnett, H., & Hemforth, B. (2021). A Journal of General Linguistics. [7] Zhan, M., Levy, R., Kehler, A. (2020). PLOS ONE. [8] Huang, C.-T. J., Li, Y.-h. A., & Li, Y. (2009). Cambridge University Press. [9] Kilgarriff, A., Baisa, V., Bušta, J. et al. (2014). Lexicography

Non–local NP   **Local NP**   ***Pronoun***

- [张伟]$_i$ 说   [小明]$_j$ 把   *[自己]$_{i/j}$*   弄糊涂了。

[Zhang Wei]$_i$ says [Xiao Ming]$_j$ BA   *[ziji]$_{i/j}$* confused.

[Zhang Wei]$_i$ says that [Xiao Ming]$_j$ confused *[self]$_{i/j}$*.

RSA: A Probabilistic model

The mathematical formulars:

(r: Referent, u: Utterance)

Rational Listener Interpretation

$$P_L(r|u) \propto P(r) \times P_s(u|r)$$

Speaker Probability (Likelihood)

$$P_s(u|r) \propto exp(\alpha \times U_s(u, r))$$

Speaker Utility

$$U_s(u, r) = ln(P_{LL}(r|u)) - Cost(u)$$

Literal Listener (normalize)

Let R be set of all grammatical referents, {r1, r2, ...}, for utterence u,

$$P_{LL}(r|u) = \frac{P(r)}{\sum_{i \in R} P(i)}$$

Prior probability for each world state

$$P(r)$$

Message Cost

$$Cost(u) = -log(P(u))$$

Online Experiments

- Stimuli:
  - ([Non-local NP] says that [Local NP] [VERB] [PRONOUN].)
  - **30 root stimuli** with similar structure were designed which were used in both experiments.
  - 15 are in **co-argument** condition (ex. self, himself, him).
  - The other 15 are in **possessor** condition (ex. self's, himself's, him's).

Experiment 1 - Pronoun Interpretation Task

- Goal: To find out how people interpret different pronouns in Mandarin.

- Task: Given a **complete sentence** with a pronoun, participants were asked to choose who the pronoun refers to by clicking the picture of the character.

- 30 items for 3 conditions
  [*ziji*, *taziji*, *ta*]

- 135 valid participants.

- (About 45 participants per condition)

Wang said Zhang held *ziji's* child in arms.

王刚说张伟把自己的孩子搂在了怀里。

问题：谁的孩子被搂在了怀里?

Q: Whose child was held in arms?

王刚 张伟 其他人

Experiment 2 - World Knowledge Bias Test

- Goal: To find out people's knowledge about the world. Who is more likely to be the theme/experiencer of an action, Non-local NP or Local NP or Others?

- Task: Given a **real-world situation** with a **question mark** in the target position, participants need to choose a character to fill the question mark based on their world knowledge.

- 30 item same as Exp. 1

- 28 valid participants

Experiment 3 - Pronoun Production Task

- Goal: To find out how people choose pronoun to refer to a given referent

- Task: Given a sentence with a gap in the target position and given a character, participants were asked to select the most natural pronoun from a drop-down menu to convey the character information.

- 30 root stimuli over 2 conditions

[Non-local NP, Local NP]

(since if we want to refer to Others,

the only possible pronoun is ta(him))

- 65 valid participants.

(Around 32 participants per condition)

图中的人物代表句子缺少的人物。

请根据您的直觉及日常习惯，选择与图中人物相符的代词来完整句子。

Choose the most natural pronoun to convey the given referent information

小明说王刚把 ✓自己 的照片发到了朋友圈。

Ming said Wang posted ziji/taziji/ta/Ming 's picture.

小明
他自己

小明

Given: Ming

Corpus Study

- Source: **Corpus of the Chinese Web** 2017 (zhTenTen17)
  - is a Chinese corpus made up of texts collected from the Internet.
  - 16,593,146,196 (16.5 billion tokens)
  - 13,531,331,169 (13.5 billion words)

Results:

Pronoun	Coargument	Possessor
ziji	29.54%	**71.29%**
ta	**70.31%**	28.45%
taziji	0.15%	0.26%

Stimuli: Co-argument

1. 把自己弄糊涂了。
2. 把自己出卖了。
3. 把自己照顾得很好。
4. 把自己当成小孩子了。
5. 把自己弄伤了。
6. 把自己想像成侦探了。
7. 把自己灌醉了。
8. 把自己绊倒了。
9. 把自己锁在教室里了。
10. 把自己逼得太紧了。
11. 把自己置于危险中。
12. 把自己气哭了。
13. 把自己吓坏了。
14. 把自己弄骨折了。
15. 把自己逗笑了。

1. confused ziji
2. betrayed ziji
3. took good care of ziji
4. treated ziji as a child
5. hurted ziji
6. imagined ziji as a policeman
7. got ziji drunk
8. tripped over ziji
9. locked ziji in the classroom
10. pushed ziji too hard
11. put ziji in danger
12. made ziji cry
13. scared ziji
14. broken ziji's bones
15. made ziji laugh

Stimuli: Possessor

16. 把自己的照片发到了朋友圈。
17. 把自己的演讲稿修改了。
18. 把自己的想法忘记。
19. 把自己的作业遗漏了。
20. 把自己的遭遇讲给了别人听。
21. 把自己的工作做完了。
22. 把自己的想法表达出来了。
23. 把自己的孩子搂在了怀里。
24. 把自己的证件弄丢了。
25. 把自己的小狗找回来了。
26. 把自己的玩具模型弄坏了。
27. 把自己的房间打扫干净了。
28. 把自己的秘密告诉了别人。
29. 把自己的衣服捐掉了。
30. 把自己的文件删除了。

16. posted ziji's photos to social media
17. revised ziji's speech
18. forgot ziji's ideas
19. left out ziji's homework
20. told others ziji's encounters
21. finished ziji's work
22. expressed ziji's idea
23. held ziji's child in arms
24. lost ziji's documents
25. found ziji's dog
26. broke ziji's toy model
27. cleaned ziji's room
28. told ziji's secret to others
29. donated all ziji's clothes
30. deleted ziji's documents