

Speaker choice in language production: Optional passive marker BEI in Mandarin
Yingzhao Zhou* (NEU), Yingtong Liu* (Harvard), Edward Gibson (MIT) [* Equal Contribution]
zhou.yingz@husky.neu.edu

The language production system needs to choose an informative and efficient utterance when multiple forms are available to express the same meaning. We tested whether optional word mention is explained by two separate accounts. One is the interference account, which is hypothesized to stem from the nature of memory representations in sentence production. The interference account appears to affect speakers' implicit structure choice, in that semantically similar words interfere with lexical retrieval and therefore affect sentence production (Hsiao & MacDonald, 2016; H&M). The other is the noisy-channel account, which states there are various of noise in the communication procedure, such as environment noise or humans' inherent bias (Gibson et al., 2013; Levy, 2008). Speakers tend to choose the linguistic form that facilitates robust information transmission and smooth communication.

The two accounts make different predictions on Mandarin production of the passive marker BEI in relative clauses. BEI is optional and does not affect the meaning of the clause, as shown in (1c). Relative clauses involve complex structures and BEI serves as a cue to mark the agent in the event structure, thus potentially assisting information transmission. The noisy-channel account therefore predicts speakers use BEI more often when the event described in the relative clause is less consistent with people's bias in semantic role assignment. In contrast, the interference account predicts speakers produce BEI more often when the animacy conditions of relative clause agent and patient are the same (e.g., both are animate) than when they are different, to reduce interference.

H&M (2016) argued for the interference theory. **Exp1** aims to replicate their results. The interference account and the noisy-channel account make the same prediction here. 48 Mandarin speakers were provided with two simple sentences such as (1ab) and asked to combine them into a complex sentence. Across all trials, N1 was animate and N2 was either animate (1a) or inanimate (1b). To avoid potential garden path effects, the matrix verb in each trial (e.g., 'protect' in (2)) was designed to be pragmatically incompatible with N1 ('the bomb') in the Chinese context. There were six trials in total with Latin square design. Each participant constructed six complex sentences such as (1c) in which they either use BEI or not. Both the SI account and NC model predict more BEI usage in (1a) than (1b). Events including animate agent and inanimate patient (1b) are more consistent with people's semantic-role-assignment bias and have less similarity-interference compared to animate agent and patient (1a). A mixed logit regression with N2 animacy condition as predictor was employed to analyze the presence/absence of BEI in responses. Our results replicated H&M (2016): Participants were more likely to produce BEI if N2 (patient) was animate (1a) ($\beta=-1.53$, $z=-2.5$, $p<0.013$), as Fig.1.

The interference account and the noisy-channel account make different predictions in **Exp2**. 70 Mandarin speakers were provided with two simple sentences such as (2ab) and were asked to combine them into one sentence. Across all trials, N1 is inanimate, while N2 is either animate (2a) or inanimate (2b). As in Exp1, each participant constructed six complex sentences in which they either use BEI or not. The noisy-channel account predicts more BEI usage in (2a) where agent is inanimate and patient is animate, than (2b) where both agent and patient are inanimate. This is because people prefer animate agents to inanimate agents, and semantic role assignment in event (2a) is more reversible and noisy than (2b). However, the interference account predicts more BEI usage in (2b) where both agent and patient are inanimate, than (2a), in that inanimate agent and patient are more semantically similar and may lead to a higher interference. A similar analysis was applied to the results. Participants were more likely to produce BEI if patient is animate (2a) ($\beta=-1.14$, $z=-2.25$, $p<0.025$) as shown in Fig.2, thus supporting the noisy-channel account but not the interference account.

Our findings support the noisy-channel account over the interference account: The usage of optional functional words like BEI is driven by informative and efficient communication.

(1) Experiment I stimuli

- a. 商场 调查 店员(N2)。 顾客(N1) 抱怨 店员(N2)。
The shopping mall investigate staff. Customer criticize staff.
- b. 商场 调查 商品(N2)。 顾客(N1) 抱怨 商品(N2)。
The shopping mall investigate product. Customer criticize product
- c. 商场 调查 (被) 顾客(N1) 抱怨 的 店员(N2)。
The shopping mall investigate (BEI) customer criticize DE staff
'The shopping mall investigated the staff that [the customer criticized/was criticized by the customer]'

(2) Experiment II stimuli

- a. 部队 保卫着 总统(N2)。 炸弹(N1) 袭击 总统(N2)。
The army protects the president. The bomb attacks the president.
- b. 部队 保卫着 村庄(N2)。 炸弹(N1) 袭击 村庄(N2)。
The army protects the village. The bomb attacks the village.

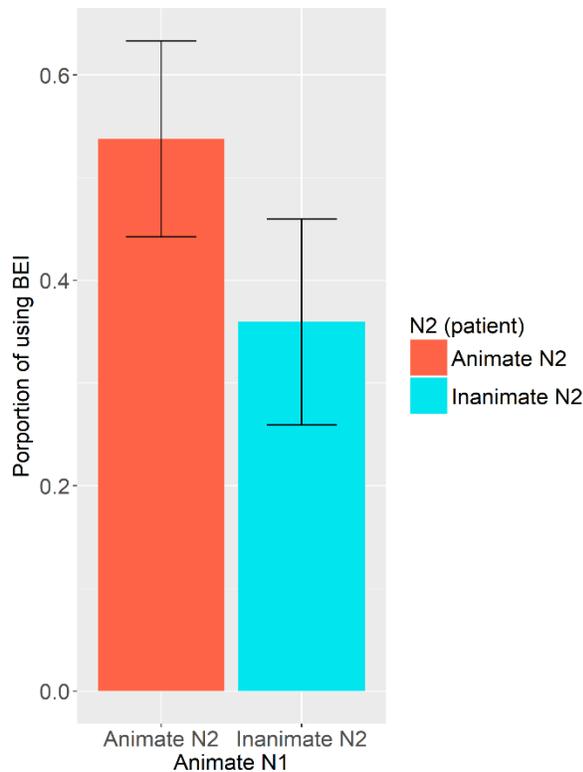


Fig. 1. Proportion of using BEI in Exp1 by patient/N2 animacy (animate vs. inanimate) The error bars show 95% confidence interval.

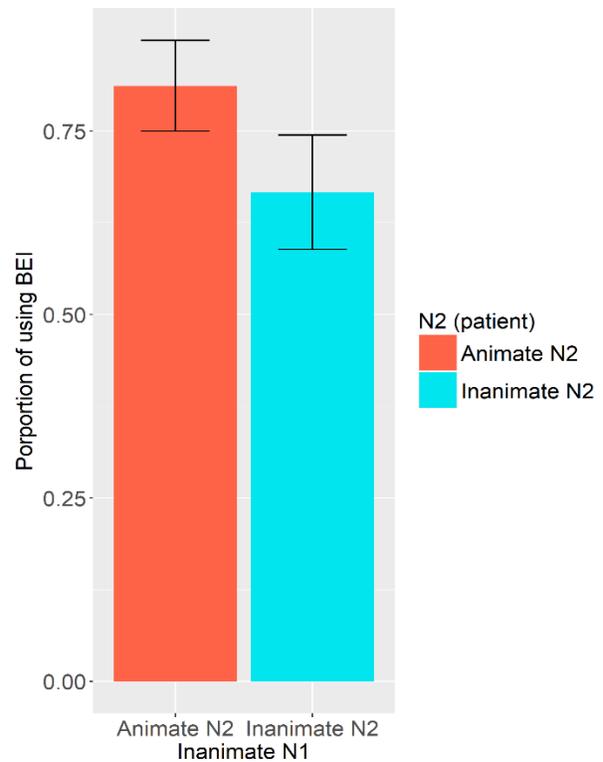


Fig. 2. Proportion of using BEI in Exp2 by patient/N2 animacy (animate vs. inanimate) The error bars show 95% confidence interval.

References: [1] Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *PNAS*, 201216438. [2] Hsiao, Y., & MacDonald, M. C. (2016). Production predicts comprehension: Animacy effects in Mandarin relative clause processing. *JML*, 89, 87-109. [3] Roger Levy. (2008). A noisy-channel model of rational human sentence comprehension under uncertain input. *Proceedings of the 13th Conference on EMNLP*, pp. 234–243.