Memory & Cognition
https://doi.org/10.3758/s13421-025-01774-1

Decoding the implausible: Mandarin sentence interpretation through the noisy channel model

Ruihua Mao[1] · **Sihan Chen**[2] · **Edward Gibson**[2]

Accepted: 22 July 2025

Abstract
The noisy channel language comprehension proposal posits that comprehenders detect and correct errors when interpreting sentences. This study replicates and extends Zhan et al. (2023), testing the model in Mandarin Chinese with three syntactic alternations: (1) Active–Passive–BA sentences, (2) Double Object (DO)–Initial position Prepositional Object (PO)–Final position PO sentences, and (3) Transitive–Initial position Adverbial Intransitive–Final position Adverbial Intransitive sentences. In each alternation, the first two structures were adopted from Zhan et al. (2023), while the third was introduced in this study. These alternations require different numbers and types of edits to transform implausible sentences into plausible ones. Participants read test items and answer corresponding comprehension questions, which indicate whether they interpret the item literally. The results aligned with Zhan et al. (2023)'s findings, indicating that Mandarin participants were most likely to make inferences for implausible sentences resulting from deleting or inserting a single morpheme, followed by those formed by a noun phrase exchange across a function word, and least likely to make inferences for implausible sentences obtained through a noun phrase exchange across a main verb. The inclusion of novel structures reinforces the robustness of the noisy-channel framework and highlights how language-specific properties influence language comprehension.

Keywords Rational inference · Noisy channel · Mandarin · Sentence processing

Introduction

Past theories of sentence processing typically assumed that there are no errors in the input (Frazier & Fodor, 1978; Gibson, 1991; Hale, 2001; Levy, 2008a). However, in actual language use, producers may misspeak or mistype, and comprehenders may mishear or misread. Additionally, environmental noise can lead to potential misunderstandings. Despite errors and potential misunderstandings, comprehenders might successfully recover the intended information. It is therefore necessary for a sentence processing theory to consider the existence of errors and misunderstandings in real-world communication. Figure 1 illustrates the framework for how communication occurs across a noisy channel, as proposed by Shannon (1949), Levy (2008b) and Gibson et al. (2013). Producers use an intended sentence (s_i) to express their intended meaning (m_i). Across the noisy channel, the intended sentences will be corrupted to the perceived sentences (s_p) by the producer errors, comprehender errors, or environmental noise. Comprehenders then need to decode the perceived sentence to uncover its intended meaning (m_p). If the meaning of the perceived sentence matches the meaning of the intended sentence, communication is considered successful.

Equation 1 was put forward in Gibson et al. (2013) as an approximation of the probabilistic inferences that take place in the noisy channel framework. Here, the left-hand $P(s_i|s_p)$ represents the probability for a rational comprehender to infer the intended sentence s_i given the perceived sentence s_p. $P(s_i)$ represents the prior probability of the intended sentence in the world (its meaning and its form), and the noise model $P(s_i \rightarrow s_p)$ represents the likelihood of the intended sentence being corrupted to the perceived sentence. According to this noise model, larger changes are less likely compared to smaller ones. Hence, the probability to infer the intended sentence based on the perceived sentence $P(s_i|s_p)$ is pro-

✉ Ruihua Mao
ruihua.mao@etu.u-paris.fr

[1] Laboratoire de Linguistique Formelle, Université Paris Cité, Bât. Olympe de Gouges, 5ème étage, 8, Rue Albert Einstein, 75013 Paris, France

[2] Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Published online: 03 October 2025

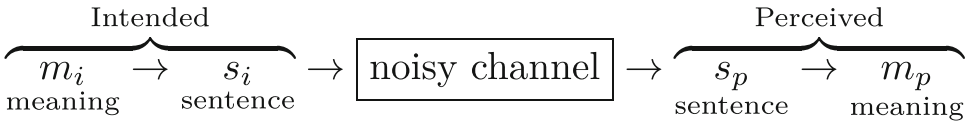

Fig. 1 Schematic representation of the noisy channel

portional to both the real-word probability of the perceived sentence $P(s_i)$ and the likelihood of the intended sentence being corrupted to the perceived sentence $P(s_i \rightarrow s_p)$.

$$P(s_i|s_p) \propto P(s_i)P(s_i \rightarrow s_p) \quad (1)$$

Previous studies

Several studies have been conducted to evaluate the noise model in the noisy-channel framework, using English materials. Gibson et al. (2013) investigated five English alternations as shown in Table 1. The critical items were materials that were all literally implausible, following their syntax, but could be interpreted plausibly with a variable number of edits across materials. Participants answered a corresponding Yes or No comprehension question, which allowed the experimenters to infer whether the item was taken literally (according to the syntax) or nonliterally (with an inference). For example, in the active/passive alternation example, the comprehension question was "Did the girl kick something/someone?" If a participant answered "Yes", they were probably following the non-literal interpretation; if they answered "No", they were probably following the literal interpretation.

Based on the results of their five English alternations, Gibson et al. (2013) proposed a noise model composed solely of function word insertions or deletions, treating exchanges as combinations of insertions and deletions. They found that comprehenders were more willing to take the non-literal interpretation when: (1) fewer changes were needed from the implausible perceived sentences to derive the semantically plausible interpretation; (2) a deletion, rather than an insertion, was necessary to achieve the plausible interpretation, explained by the Bayesian size principle (Xu & Tenenbaum, 2007): compared with the deletion of a word, the insertion of a word has smaller likelihood P($s_i \rightarrow s_p$), since a deletion only involves selecting a word from a sentence, but an insertion necessitates selecting a word from the produce's entire vocabulary; (3) there were errors in the filler items, leading comprehenders to perceive an increased error rate; and (4) the base rate of implausible sentences was increased.

Table 1 Example sentence stimuli used in Gibson et al. (2013). Each item varies in sentence construction (i.e., active or passive) and plausibility. Plausible sentences (left column) can be potentially corrupted into implausible sentences (right column) by different kinds of noise operations (middle column)

Plausible Conditions	Edits	Implausible Conditions
(1a)Plausible Active The girl kicked the ball.	Two insertions	**(1c)Implausible Passive** The girl was kicked by the ball.
(1b)Plausible Passive The ball was kicked by the girl.	Two deletions	**(1d)Implausible Active** The ball kicked the girl.
(2a)Plausible Subject-locative Onto the table jumped a cat.	One deletion + One insertion	**(2c)Implausible Object-locative** The table jumped onto a cat.
(2b)Plausible Object-locative The cat jumped onto a table.	One insertion + One deletion	**(2d)Implausible Subject-locative** Onto the cat jumped a table.
(3a)Plausible Transitive The tax law benefited the businessman.	One insertion	**(3c)Implausible Intransitive** The tax law benefited from the businessman.
(3b)Plausible Intransitive The businessman benefited from the tax law	One deletion	**(3d)Implausible Transitive** The businessman benefited the tax law.
(4a)Plausible DO goal The mother gave the daughter the candle	One insertion	**(4c)Implausible PO goal** The mother gave the daughter to the candle.
(4b)Plausible PO goal The mother gave the candle to the daughter	One deletion	**(4d)Implausible DO goal** The mother gave the candle the daughter.
(5a)Plausible DO benefactive The cook baked Lucy a cake.	One insertion	**(5c)Implausible PO-benefactive** The cook baked Lucy for a cake.
(5b)Plausible PO benefactive The cook baked a cake for Lucy.	One deletion	**(5d)Implausible DO-benefactive** The cook baked a cake Lucy.

Poppels and Levy (2016) introduced exchanges as its own noise operation into the noise model. They replicated three alternations of Gibson et al. (2013): Active/Passive, Transitive/Intransitive and DO/PO benefactives, and added a different syntactic alternation: Prepositional Phrase Exchanges (PP exchanges). Examples of PP exchanges are shown in (1):

(1) Plausible, canonical: The package fell from the table to the floor.
Plausible noncanonical: The package fell to the floor from the table.
Implausible canonical: The package fell from the floor to the table.
Implausible noncanonical: The package fell to the table from the floor.

According to the noise model proposed by Gibson et al. (2013), four edits (two deletions and two insertions) would be required for the plausible sentences in (1) to be recovered from the implausible ones. Given that such a significant number of edits is highly unlikely under the noise model of Gibson et al. (2013), which does not contain exchanges, we would expect few inferences when participants encounter these kinds of implausible sentences.

However, Poppels and Levy (2016) showed that participants still made many inferences when they encountered the implausible PP exchange sentences. Indeed, the inferences were more frequent than the implausible active/passive conditions, which require two edits according to Gibson et al. (2013). Consequently, Poppels and Levy (2016) argued that exchanges should be considered a separate noise operation. They also noted that the environment for the exchanges matters: exchanges may be easier when not across a main verb into subject position, as would be needed in the active-passive materials in Table 1. This difference might explain why the inference rate is lower for active-passive materials, where exchanges occur across a main verb, compared to the higher inference rate for PP exchange sentences, where such complex exchanges are not necessary.

Gibson et al. (2017) investigated whether sentence interpretation is influenced by the speaker's accent and found that comprehenders are more likely to interpret accented speech non-literally compared to non-accented speech. This tendency can be attributed to listeners assuming a higher error rate in foreign-accented speech, prompting them to make more inferences to arrive at a plausible interpretation. In other words, when listeners encounter an implausible sentence spoken with a non-native accent, they are more likely to infer that the speaker intended a plausible message but produced an error, compared to when the same sentence is spoken with a native accent. This finding is compatible with the noisy channel predictions and aligns with the earlier noise model proposed by Gibson et al. (2013) and Poppels and Levy (2016), supporting the robustness of the noisy channel model across different modalities.

Ryskin et al. (2018) provided compelling evidence that exchanges are indeed treated as a distinct category of error in sentence processing. They asked participants to correct sentences if they noticed errors, and they also manipulated context by adding exposure sentences that contained different error types (deletions, insertions, exchanges, mixed errors, and no errors). Their results revealed that participants were more likely to correct sentences through insertions and less likely to correct sentences through exchanges when the exposure conditions contained deletions or insertions. This also supports the idea that the noise model is sensitive not only to the base rate of noise but also to the nature of the errors presented.

In real-world communication, sentences are typically embedded within a discourse context, while all the aforementioned research utilized isolated sentences. Chen et al. (2023) investigated how different kinds of context would affect people's likelihood to interpret a sentence literally. They used three kinds of contexts (a supportive context, a non-supportive context, and a null context) to test how people interpret two syntactic alternations: active/passive, and DO/PO in context. They found that, when an implausible sentence was preceded by a supportive context, people were more likely to interpret it non-literally than when a non-supportive context or no context preceded it. Poliak et al. (2024) tested the noisy channel in a more naturalistic task: listening to lyrics, and the results suggest that noisy channel processing predictions are applicable to natural tasks and materials.

There have also been several previous studies that have investigated the noisy-channel approach conducted in non-English languages, but most of these investigated aspects of the prior syntactic probability within those languages (e.g., Liu et al., 2020; Keshev & Meltzer-Asscher, 2021; Poliak et al., 2023; Dodd et al., 2024, in Mandarin Chinese, Hebrew, Russian, and Arabic, respectively).

Of most relevance to the current work, Zhan et al. (2023) tested three pairs of Mandarin syntactic alternations, closely aligned in design to the English materials in Gibson et al. (2013), utilizing both the written and spoken modality. The corresponding noise operations were modified according to language differences, and exchanges were treated as a separate noise type according to Poppels and Levy (2016). The materials and edits were shown in Table 2.

The results showed that comprehenders were most likely to make inferences when they perceived sentences made implausible by deletions or insertions than those made implausible by the exchange across a function word, and they made the fewest inferences while perceiving implausible sentences made by the exchange across a main verb. Overall, the

Table 2 Example sentence stimuli used in Zhan et al. (2023). Each item varies in sentence construction (i.e., active or passive) and plausibility. Plausible sentences (left column) can be potentially corrupted into implausible sentences (right column) by different kinds of noise operations (middle column).

Plausible Conditions	Edits	Implausible Conditions
(1a)Plausible Active 奶奶打碎了这个碗。 Grandma break-ASP this-CL bowl "Grandma broke the bowl"	NP exchange across main verb	**(1c)Implausible Active** 这个碗打碎了奶奶。 This-CL bowl break-ASP grandma "The bowl broke Grandma"
(1b)Plausible Passive 这个碗被奶奶打碎了。 This-CL bowl bei grandma break-ASP "The bowl was broken by Grandma."	NP exchange across function word	**(1d)Implausible Passive** 奶奶被这个碗打碎了。 Grandma bei this-CL bowl break-ASP "Grandma was broken by the bowl."
(2a)Plausible DO 老林付了清洁工五十块钱。 Laolin pay-ASP cleaner fifty-CL money 'Laolin paid the cleaner fifty yuan."	Insertion	**(2d)Implausible serial verb** 老林付了清洁工给五十块钱。 Laolin pay-ASP cleaner gei fifty-CL money "Laolin paid the cleaner to fifty yuan."
(2b)Plausible serial verb 老林付了五十块钱给清洁工。 Laolin pay-ASP fifty-CL money gei cleaner "Laolin paid fifty yuan to the cleaner."	Deletion	**(2c)Implausible DO** 老林付了五十块钱清洁工。 Laolin pay-ASP fifty-CL money cleaner "Laolin paid fifty yuan the cleaner."
(3a)Plausible transitive: 清水溶解了食盐。 Clear-water dissolve-ASP salt "The clear water dissolved the salt."	NP exchange across main verb	**(3c)Implausible transitive:** 食盐溶解了清水。 Salt dissolve-ASP clear-water "The salt dissolved the clear water."
(3b)Plausible intransitive 食盐在清水里溶解了。 Salt zai clear-water li dissolve-ASP "The salt dissolved in the clear water."	NP exchange across function word	**(3d)Implausible intransitive** 清水在食盐里溶解了。 Clear-water zai salt li dissolve-ASP "The clear water dissolved in the salt."

Mandarin results are similar to the English results, demonstrating that Mandarin speakers also interpret the implausible sentences in a rational way, integrating both the likelihood of different error types and their world knowledge. In addition, Zhan et al. (2023) found that results from the spoken modality experiment were generally consistent with results from the written modality experiment, but participants made fewer inferences for implausible conditions in the listening modality. They attributed this difference to the absence of spaces between characters in Mandarin Chinese orthography, which may hinder participants' ability to recognize implausible sentences, leading them to interpret implausible sentences directly as plausible ones, thereby resulting in higher reference rates. This process might reflect a perceptual misrecognition, which can be viewed as a passive repair mechanism within the broader noisy-channel framework, rather than the overt inference mechanism that we focus on in this study in which comprehenders are aware of implausibilities and infer more plausible alternatives.

In the current study, we replicate and extend the work of Zhan et al. (2023) using the written modality. Zhan et al. (2023) chose certain syntactic structures in their materials, in their effort to understand rational inference in Mandarin, but there exist other frequently used sentence structures in Mandarin, apart from those in Zhan et al. (2023), that could also serve as controls in each syntactic alternation. Hence, we considered one additional control structure in each of the alternations that Zhan et al. (2023) explored, aiming to better capture the variability and complexity of Mandarin syntax.

In the active/passive alternation, we added the *Ba* construction which uses subject-object-verb (SOV) word order. The *Ba* construction is commonly used and often preferred by Mandarin speakers over the canonical subject-verb-object (SVO) order in many disposal and causative contexts (Tsao, 1987). In the DO/PO alternation, the coverb *gei*("to") of the PO construction can appear either in sentence-final position, as tested in Zhan et al. (2023), or immediately after the subject and before the main verb. We included the latter variant in our design. In the transitive/intransitive alternation, the locative phrase of the intransitive constructions can appear between the subject and the main verb, as tested in Zhan et al. (2023), or at the beginning of the sentence, so we included the sentence-initial variant in our design. These sentence structures are frequently used by native Mandarin speakers and require different noise operations to repair implausible conditions to the corresponding plausible ones. By incorporating these additional constructions, we aim to test the applicability of noisy channel theory across a broader range of syntactic contexts, thereby providing a more comprehensive understanding of rational inference in Mandarin Chinese and offering new insights into how language-specific properties shape inference likelihood.

Methodology

Zhan et al. (2023) investigated three syntactic alternations in Mandarin Chinese, and each experiment manipulated two variables: syntactic alternations and plausibility, making each a 2x2 Latin square design. We replicated the three syntactic alternations in Mandarin Chinese from Zhan et al. (2023), and added one construction into each alternation, transforming each experiment into a 3x2 Latin-square design. Experiment 1 examined the Active/Passive/*Ba* alternation, as illustrated in (2); Experiment 2 examined the DO/Initial-Position PO/Final-Position PO alternation, as shown in (3); Experiment 3 examined the Transitive/Final-position Adverbial Intransitive/Initial-position Adverbial Intransitive alternations, as presented in (6).

The three experiments were conducted together, with each experiment containing 24 items varying in construction and plausibility ("target items"). There were also 48 shared fillers. Participants accessed the study via MIT Surveyor (https://mit-surveyor.com/), where they viewed each item and then answered a corresponding yes/no comprehension question, which can be used to determine if they interpreted the item literally or non-literally. Half of the comprehension questions were implausible, while the remaining ones were plausible, with the coding for interpretation adjusted accordingly. Following previous studies (e.g., Gibson et al., 2013), plausible target items and fillers serve as attention checks. Participants who failed to answer 75% of these plausible materials correctly were excluded from our data analysis. The presentation order was randomized for each participant, and the experiment took approximately 15 min.

Experiment 1: The Active/Passive/Ba alternation

Materials

Example materials for Experiment 1 are shown in (2). The Active conditions and Passive conditions were the same as Zhan et al. (2023), with the *Ba*-construction added. Although Mandarin typically follows a SVO word order, it also uses the *Ba*-construction, which follows a SOV word order. This construction follows the pattern [(NP1)-ba-NP2-VP], where NP1 is the AGENT or CAUSER of the event, and NP2 is the PATIENT or AFFECTEE (Li & Thompson, 1989).

(2) Mandarin Active – Passive – Ba

a. Plausible Active
妹妹 关上了 窗户。
younger.sister close-ASP window
'The younger sister closed the window.'

b. Plausible Passive
窗户 被 妹妹 关上了。
window Bei younger.sister close-ASP
'The window was closed by the younger sister.'

c. Plausible Ba
妹妹 把 窗户 关上了。
younger.sister BA window close-ASP
'The younger sister has the window closed.'

d. Implausible Active **(NP exchange across main verb)**
窗户 关上了 妹妹。
window close-ASP younger.sister
'The window closed the younger.sister.'

e. Implausible Passive **(NP exchange across function marker *Bei*)**
妹妹 被 窗户 关上了。
younger.sister BEI window close-ASP
'The younger sister was closed by the window.'

f. Implausible Ba **(NP exchange across function marker *Ba*)**
窗户 把 妹妹 关上了。
window BA younger.sister close-ASP
'The window has the younger sister closed.'

A comprehension question is provided with each item. For the example item, the plausible comprehension question is "妹妹关上了什么东西/什么人吗?"(Did the younger sister close something/somebody?). "Yes" is the literal response for plausible conditions, and "No" is the literal response for the implausible conditions. The implausible comprehension question is "什么东西/什么人关上了妹妹吗?"(Did something/somebody close the younger sister?). In this case, "No" is the literal response for the plausible items, and "yes" is the literal response for the implausible items.

An exchange of noun phrases across the main verb is necessary to get the implausible active sentence (2d) from the corresponding plausible one (2a). Regarding the implausible Passives (2e), they can be corrected by exchanging noun phrases across the passive function marker "Bei" in the plausible sentence (2b), and the implausible Ba sentence (2f) can also be rectified by exchanging noun phrases across the function word "Ba" into (2c). Since an NP exchange across the main verb is regarded as a low-probability noise operation compared with an NP exchange across a function word according to the noise model that has been proposed thus far (Garrett, 1975; Zhan et al., 2023), the noisy-channel approach predicts that people should interpret the implausible active sentence (2d) more literally compared with the other two implausible conditions.

Summing up, the noisy-channel approach predicts the literal interpretation rate sequence: **Plausible conditions** > **Implausible Active** > (**Implausible passive = Implausible Ba**).

Participants

Eighty-one native Mandarin speakers were recruited through advertisements posted on social media. All participants self-reported as native speakers of Mandarin Chinese. They also took part in both Experiment 2 and Experiment 3 during the same session and were compensated with 25 Chinese Yuan (approximately 3.45 USD) for their participation.

Data analysis

All data analyses in this study were conducted using the R statistical software (R Core Team, 2021). We employed a Bayesian generalized linear mixed-effect model, implemented with the brms package (Bürkner, 2017, 2018, 2021).

We performed two main analyses in Experiment 1. The first analysis compared literal interpretation rates between plausible and implausible sentences. In this analysis, plausibility was treated as a fixed effect, while items and participants were treated as random effects, including random by-item and by-participant slopes for plausibility. The predictor variable was coded using dummy coding, with the reference level (implausible) coded as 0 and the other level (implausible) coded as 1. The model specification is presented in Equation (2). To model the binary outcome, we employed a Bernoulli distribution with a logit link function. We specified weakly informative priors, setting normal(0, 2) for the fixed effects coefficients to balance prior assumptions with data-driven inference. To ensure model convergence and minimize divergent transitions, we set adapt_delta to 0.95. The model was fitted using four Markov chains, each running for 4000 iterations, and utilized eight CPU cores to optimize computation speed.

$$Response \sim Plausibility + (1 + Plausibility \mid Participant) + (1 + Plausibility \mid Item) \quad (2)$$

The second analysis examined literal interpretation rates specifically within implausible conditions to determine whether participants made more inferences when implausibility resulted from exchanges across a function word compared to exchanges across a main verb.

Ideally, one would analyze the literal interpretation rate of different sentence constructions for both plausible and implausible sentences, and the noisy-channel framework would predict an interaction between construction and plausibility, in that participants were likely to interpret all plausible sentences literally, but they had different literal interpretation rate for implausible sentences, depending on the noise operations necessary to repair the sentence to a plausible alternative.

However, a problem with analyzing the plausible conditions (and interactions with them) is that the responses to plausible versions are close to 100% (ceiling), which makes logistic regression unreliable due to inflated standard errors for interaction terms. This issue is reflected in the derivation of Wald-type confidence intervals, where coefficient estimates may become unreliable when predictors result in near-perfect classification of the outcome (i.e., values close to 0 or 1).[1]. Therefore, we only examined implausible sentences in the following analyses of Experiment 1.

In this analysis, Construction was treated as a fixed effect, while items and participants were treated as random effects. We also applied dummy coding here, with the reference level coded as 0 and the other levels coded as 1. The model specification is shown in Equation (3). To thoroughly assess differences between constructions, the model was run twice: once with the implausible passive condition as the reference level and once with the implausible active condition as the reference level. The model settings, including priors and chain specifications, were identical to those described above, as all Bayesian model settings across this study were kept consistent to ensure comparability.

$$Response \sim Construction + (1 + Construction \mid Participant) + (1 + Construction \mid Item) \quad (3)$$

In all Bayesian models reported in the paper, we present the 95% credible interval to characterize the posterior distribution, along with the posterior probability of direction. This refers to the proportion of posterior samples that fall on the same side of zero as the posterior mean of the beta coefficient (Makowski et al., 2019). We interpret posterior probabilities of 0.95 or higher as strong evidence for an effect, values between 0.90 and 0.95 as moderate evidence, and values below 0.90 as indicating little or no evidence for an effect.

Results and discussion

Each participant's accuracy on the fillers was above 87%, so we retained all of their data. Proportions of literal interpretation are illustrated in Fig. 2, and the results are summarized in Table 3.

In Experiment 1, the literal interpretation rates of the plausible conditions were all near ceiling (Plausible Active: 0.978; Plausible Passive: 0.941; Plausible Ba: 0.966). Bayesian analysis provided very strong evidence in favor of literal interpretation for plausible conditions compared to implausible conditions. Regarding the implausible conditions, there was strong evidence that implausible active conditions were

[1] While this issue is not always explicitly emphasized in the published literature, it is widely acknowledged in applied statistical practice. See, for example, Levy (2025, p. 150): https://github.com/rlevy/quantitative-inference-spring-2025/blob/main/assets/slides/2025-03-05-categorical-predictors-interactions-logistic-regression.pdf

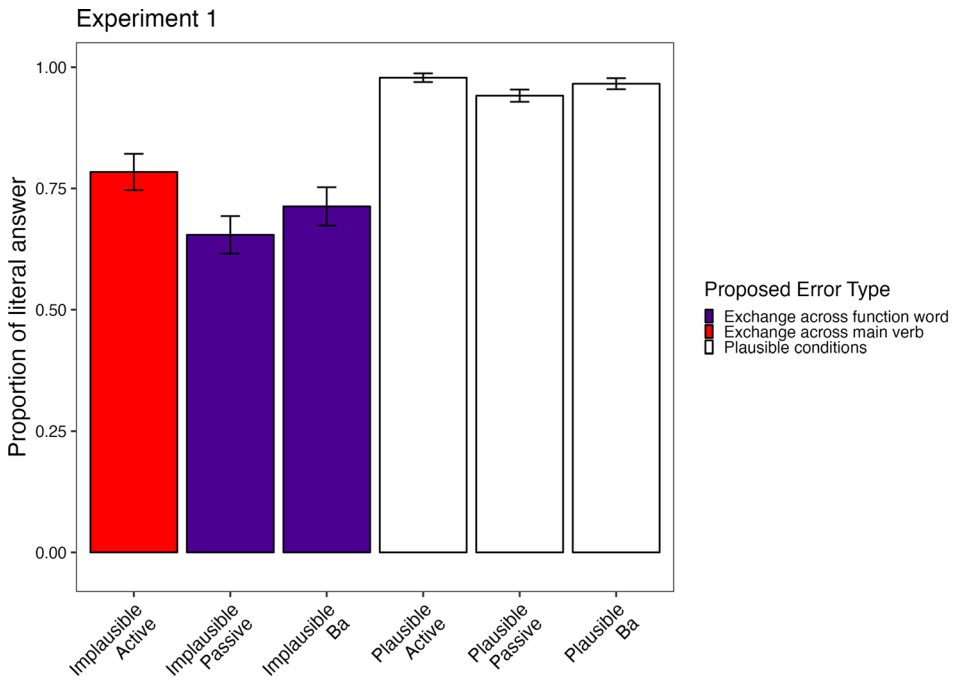

Fig. 2 Experiment 1: Percentage of trials where participants relied on the literal syntax for interpretation of the sentences. *Error bars* represent 95% confidence intervals

interpreted more literally than either the implausible Passive condition or the implausible *Ba* condition. This finding aligns with the noisy-channel prediction, as the former involves an exchange across the main verb, whereas the latter two involve an exchange across function words.

An unexpected pattern emerged when comparing the Implausible *Ba* condition and the Passive condition. This comparison showed a moderate positive effect size, with a 90.5% probability of being greater than zero, despite both conditions requiring the same correction operation: NP-exchange across function words. One possible explanation for this effect could be the higher frequency of *Ba* structures compared to passive *Bei* structures in Mandarin. According to structural frequency bias theory, a structure with a lower prior probability may make comprehenders more likely to interpret it as a more frequent alternative (Liu et al., 2020). In order to evaluate this possibility, we searched for sequences consisting of "*Bei*+Noun+Verb" and "*Ba*+Noun+Verb" in two Mandarin corpora: Universal Dependencies (https://universaldependencies.org/) and the BLCU Chinese Corpus (BCC, https://bcc.blcu.edu.cn/) (Xun et al., 2016). In Universal Dependencies, the *Ba* structure was observed 100 times, compared to 60 occurrences of the *Bei* structure. In the BCC Corpus, the *Ba* structure occurred 317,938 times, while the *Bei* structure appeared 179,672 times. Both corpora thus exhibited a similar frequency bias in favor of the *Ba* structures over the passive *Bei* structures in Mandarin, consistent with a structural frequency bias for the *Ba* structures.

Table 3 Mean results, range of the 95% credible interval, and probability of beta being different from zero for the Bayesian models in Experiment 1

Comparison in Experiment 1	**Estimate**	**95% Credible Interval**	$p(\beta \neq 0)$
Comparison 1			
Plausibility (plausible vs. implausible)	2.75	[1.65, 3.98]	1.000
Comparison 2: Passive as reference			
construction(Active vs. Passive)	2.16	[0.82, 3.79]	1.000
construction (Ba vs. Passive)	0.60	[-0.33, 1.54]	0.905
Comparison 2: Active as reference			
construction (Ba vs. Active)	-0.75	[-1.59, 0.10]	0.958
construction (Passive vs. Active)	-1.39	[-2.57, -0.18]	0.985

Experiment 2: The DO/Initial-Position PO/Final-Position PO alternation

Materials

Example test materials for Experiment 2 are shown in (3). The Mandarin DO construction, as in (3a) and Final-Position PO constructions using the coverb gei "to", as in (3b), is analogous to its English counterpart. In this construction, the direct object typically denotes what is being transferred by the main verb, while the indirect object represents what is being affected by the action (Li & Thompson, 1989).

However, the coverb "gei" (to) can appear in two positions: either in the final position (3b), which was tested in Zhan et al. (2023), or directly after the subject and before the main verb. Hence, we also included conditions (3c) and (3f) to test this variation. Because of the different positions of "gei erzi" (to the son), we label them as Final-Position PO and Initial-Position PO, respectively.

(3) Plausible DO – Final-Position PO – Initial-Position PO

a. Plausible DO
爸爸 送了 儿子一部 手机。
Father send-ASP son one-CL mobile.phone
'The father sent the son a mobile phone.'

b. Plausible Final-Position PO
爸爸 送了 一部 手机 给 儿子。
Father send-ASP one-CL mobile.phone GEI son
'The father sent a mobile phone to the son.'

c. Plausible Initial-Position PO
爸爸 给 儿子 送了 一部 手机。
Father GEI son send-ASP one-CL mobile.phone
'The father sent the son a mobile phone.'

d. Implausible DO **(Deletion)**
爸爸 送了 一部 手机 儿子。
Father send-ASP one-CL mobile.phone son
'The father sent a mobile phone the son.'

e. Implausible Final-Position PO **(Insertion)**
爸爸 送了 儿子 给 一部 手机。
Father send-ASP son GEI one-CL mobile.phone
'The father sent the son to a mobile phone.'

f. Implausible Initial-Position PO **(NP exchange across the verb)**
爸爸 给 一部 手机 送了 儿子。
Father GEI one-CL mobile.phone send-ASP son
'The father sent the son to a mobile phone.'

For this example item, the comprehension question is: "儿子收到了某个东西/某人了吗?" (Did the son receive something/someone?). The Literal response would be "No" for the implausible conditions, and "Yes" for the plausible conditions.

The implausible DO (3d) could result from exchanging the two post-verbal NPs in (3a) or deleting the function word "gei" (to) in (3b). We discuss this alternation in terms of deletion, but both are possible (Ryskin et al., 2018). Concerning the implausible Final-position PO (3e) both inserting "gei" (to) into the Plausible DO condition and exchanging the two post-verb NPs are viable operations. We discuss this in terms of insertion (Ryskin et al., 2018). As for the implausible Initial-position PO (3f), it can arise from an NP exchange across the main verb. Because the edit probability varies inversely with the literal interpretation rate, the noisy-channel theory predicts the following literal interpretation proportion sequence: **Plausible conditions >Implausible Initial-Position PO>Implausible Final-Position PO>Implausible DO.**

Participants

As mentioned above, the participants for Experiment 2 were the same as the participants for Experiment 1. The three experiments were all distractors for each other in one larger experiment.

Data analysis

As in Experiment 1, we employed a Bayesian generalized linear mixed-effects model in R (R Core Team, 2021) to analyze the data, using the brms package. First, we compared the literal interpretation rates between plausible and implausible sentences. In this model, plausibility was included as a fixed effect, while items and participants were treated as random effects, with random by-item and by-participant slopes for plausibility. We applied dummy coding to the predictor variables throughout the analysis, where the reference level (implausible) was coded as 0 and the other levels as 1. The model specification is shown in Eq. (4). A normal prior distribution N(0,2) was set for the fixed effect coefficients, reflecting a prior expectation of effect sizes centered around zero with a standard deviation of 2. To facilitate convergence, all parameter estimates were initialized at 0. The model computation was expedited by using 8 cores in parallel. We ran four Markov chains with 4000 iterations each to ensure a sufficient number of posterior samples and reliable convergence diagnostics.

$$Response \sim Plausibility + (1 + Plausibility \mid Participant) + (1 + Plausibility \mid Item) \quad (4)$$

Next, we analyzed the literal interpretation rates for implausible conditions to determine whether participants made the most inferences when the implausibility could be resolved by deletions, more inferences when repaired

by insertions, and the fewest inferences when resolved by exchanges across the main verb. The model specification is shown in Eq. (5). The model settings were identical to those described above. As in Experiment 1, responses to all plausible conditions were near ceiling, making logistic regression unreliable, so we also limited our analysis to the three implausible conditions in Experiment 2. To thoroughly assess differences between constructions, we ran the model twice: once with Initial-position PO as the reference level and once with Final-position PO as the reference level. We applied dummy coding to the independent variable, coding the reference level as 0 and all other levels as 1.

$$Response \sim Construction + (1 + Construction \mid Participant) + (1 + Construction \mid Item) \quad (5)$$

Results and discussions

The proportions of literal interpretation for Experiment 2 are illustrated in Fig. 3, and the results are summarized in Table 4. The literal interpretation rates for plausible conditions were near ceiling (DO: 0.948; Final-Position PO: 0.954; Initial-Position PO: 0.926), providing very strong evidence in favor of literal interpretation for plausible conditions compared to implausible conditions, as indicated by the Bayesian analysis.

Bayesian analysis provided strong evidence that the literal interpretation rate for the implausible Initial-Position PO condition was higher than that for both the implausible DO condition and the implausible Final-Position PO condition. This finding aligns with the predictions of the noisy-channel theory.

Furthermore, Bayesian analysis provided no difference between the implausible DO condition and the implausible Final-Position PO condition, despite the former being potentially reparable by deletion and the latter by insertion. Zhan et al. (2023) also observed this result and suggested that, in colloquial Mandarin, implausible Final-Position PO sentences are highly acceptable and are interpreted similarly to plausible conditions. They proposed that comprehenders interpret "gei" as the main verb "give" rather than the preposition "to", as illustrated in (4).

(4) 爸爸 送了 儿子, 给了 一部 手机。
Father send-ASP son (something) GEI-ASP one-CL mobile.phone
'The father sent the son (something). (Father) gave (son) a mobile phone.'

Wang (2011) suggests that "gei" primarily functions in three ways: as a verb, as a preposition, and as part of a verb phrase ("gei"-VP). Lv (1999) categorized the verb function of "gei" into two types: Dative "gei" (to give) and Causer "gei" (to cause). Dative "gei" can take either double objects or just one of them. This understanding led participants to consider example (3e) as plausible, thereby reducing the rate of literal

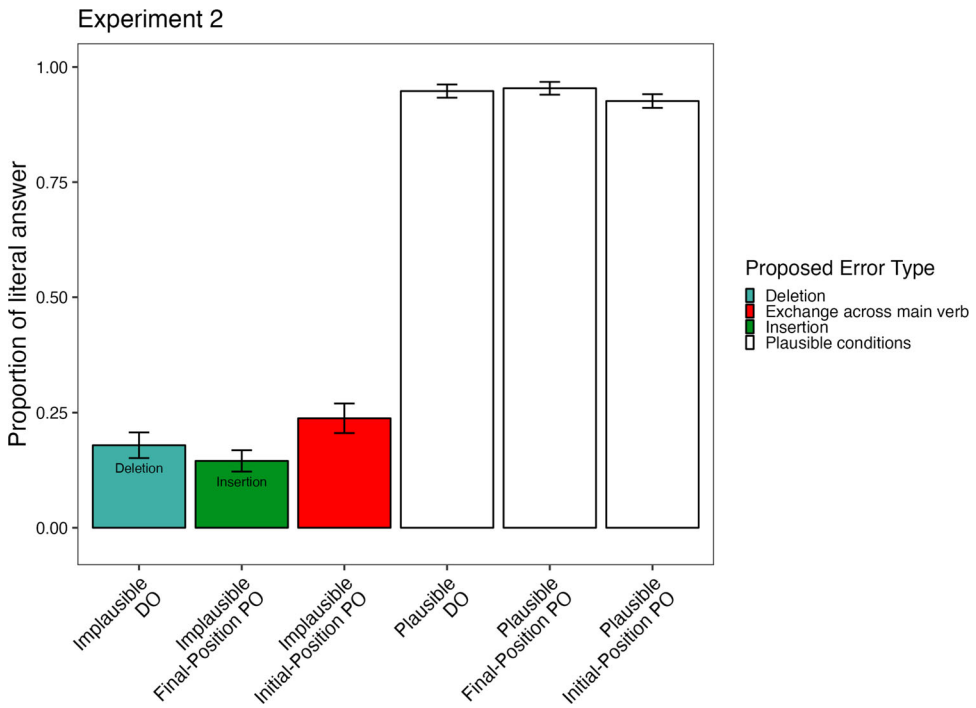

Fig. 3 Experiment 2: Percentage of trials where participants relied on the literal syntax for interpretation of the sentences. *Error bars* represent 95% confidence intervals

Table 4 Mean results, range of the 95% credible interval, and probability of beta being different from zero for the Bayesian models in Experiment 2

Comparison in Experiment 2	Estimate	95% Credible Interval	$p(\beta \neq 0)$
Comparison 1			
Plausibility (Plausible vs. Implausible)	5.91	[4.80, 7.10]	1.000
Comparison 2: Final-position PO as reference			
Construction(Initial-position PO vs. Final-position PO)	0.86	[0.15, 1.54]	0.991
Construction (DO vs. Final-position PO)	0.18	[-0.78, 1.10]	0.673
Comparison 2: Initial-position PO as reference			
Construction (DO vs. Initial-position PO)	-0.675	[-1.53, 0.07]	0.964
Construction (Final-position PO vs. Initial-position PO)	-1.23	[-2.19, -0.44]	0.998

interpretation in the results and narrowing the differences between implausible DO and Implausible Final Position PO.

Although "gei" may be misinterpreted as a main verb rather than a preposition in both implausible Final-position PO and Initial-position PO conditions (see example 5), participants were more likely to interpret the implausible Initial-position PO sentences literally. In addition to the predicted error type for the Initial-position PO condition (NP exchange across the main verb), this pattern may also be because the "gei" in Final-position PO structures is more likely to be misinterpreted as a main verb, according to several native speaker intuitions.

(5) 爸爸 给了 一部 手机, 送了 儿子。
Father GEI-ASP one-CL mobile.phone, send-ASP son
'The father gave (son) a mobile phone, sent the son (something).

Experiment 3: The Transitive/Final-position Adverbial Intransitive/Initial-position Adverbial Intransitive alternations

Materials

Example materials for Experiment 3 are shown in (6). As in (6a), the Mandarin transitive structure resembles English, but Mandarin features two intransitive forms. In Mandarin, the locative phrase has the structure: [Zai-NP-Locative.particle], with common locative particles including *"li"* (in), *"shang"* (above), *"xia"* (under), etc. Li and Thompson (1989). The locative phrase can appear between the subject and main verb, as shown in (6b), or at the beginning of a sentence, as illustrated in (6c). The former was tested by Zhan et al. (2023), labeled as the "Final-Position Intransitive" condition; we introduced the latter, which we label the "Initial-Position Intransitive" condition.

(6) Transitive – Final-Position Intransitive – Initial-Position Intransitive

a. Plausible Transitive
热水 溶解了 咖啡粉。
hot.water dissolve-ASP coffee.powder
'The hot water dissolved the coffee powder.'

b. Plausible Final-Position Intransitive
咖啡粉 在 热水 里溶解了。
coffee.powder **zai** hot.water **li** dissolve-ASP
'The coffee powder dissolved in the hot water.'

c. Plausible Initial-Position Intransitive
在 热水 里 咖啡粉 溶解了。
zai hot.water **li** coffee.powder dissolve-ASP
'In the hot water, the coffee powder dissolved.'

d. Implausible Transitive **(NP Exchange across the main verb)**
咖啡粉 溶解了 热水。
coffee.powder dissolve-ASP hot.water
'The coffee powder dissolved the hot water.'

e. Implausible Final-Position Intransitive **(Exchange across preposition *Zai*)**
热水 在 咖啡粉 里 溶解了。
hot.water **zai** coffee.powder **li** dissolve-ASP
'The hot water dissolved in the coffee powder.'

f. Implausible Initial-Position Intransitive **(Exchange across locative particle *Li*)**
在 咖啡粉 里 热水 溶解了。
zai coffee.powder **li** hot.water dissolve-ASP
'In the coffee powder, the hot water dissolved.'

Participants saw the following comprehension question: "咖啡粉是否被什么溶解了?" (Was the coffee powder dissolved by something?). "Yes" is the literal response for implausible conditions, while "No" is the literal response for plausible conditions.

An NP exchange across a main verb is required to repair the implausible transitive (6d), and an NP exchange across a preposition can repair the implausible Final-Position Intran-

sitive (6e). As for the implausible Initial-Position Intransitive condition(6f), an NP exchange across the locative particle is required. The NP exchange across a preposition and across a locative particle are both types of NP exchanges across function words, so we expect that the literal interpretation rate to follow this pattern: **Plausible conditions > Implausible Transitive > Implausible Final-Position Intransitive = Implausible Initial-Position Intransitive**.

Participants

As mentioned above, the participants for Experiment 3 were the same as the previous experiments.

Data analysis

Like in Experiment 1 and Experiment 2, we employed a Bayesian generalized linear mixed-effects model in R (R Core Team, 2021) to analyze the data, using the brms package. First, we compared the literal interpretation rates between plausible and implausible sentences in Experiment 3. In this model, plausibility was included as a fixed effect, while items and participants were treated as random effects, with random by-item and by-participant slopes for plausibility. We applied dummy coding to the independent variable, with the reference level coded as 0 and the other levels coded as 1. The model specification is shown in Eq. (6). A normal prior distribution N(0,2) was set for the fixed effect coefficients, reflecting a prior expectation of effect sizes centered around zero with a standard deviation of 2. All parameter estimates were initialized at 0 to facilitate convergence. To expedite computation, the model utilized 8 cores in parallel. Four Markov chains were run, each with 4000 iterations, to ensure a sufficient number of posterior samples and reliable convergence diagnostics.

$$Response \sim Plausibility + (1 + Plausibility \mid Participant) + (1 + Plausibility \mid Item) \quad (6)$$

Next, as in Experiments 1 and 2, we first analyzed the literal interpretation rates for implausible conditions to determine whether participants made more inferences for those implausible conditions caused by exchanging across the function word than those caused by exchanging across the main verb. The model specification is shown in Eq. (7). The model settings, including priors, chain settings, and dummy coding scheme, were identical to those described above. We ran this model twice: once with the Initial-position Intransitive condition as the reference level and once with the Transitive condition as the reference level.

In Experiment 3, items and participants were entered as random intercepts, with random by-participant slopes and by-item slopes for constructions, as shown in Eq. (7).

$$Response \sim Construction + (1 + Construction \mid Participant) + (1 + Construction \mid Item) \quad (7)$$

However, unlike in Experiments 1 and 2, in Experiment 3, the literal interpretation rate for the Final-Position Intransitive condition was not quite at the ceiling (0.861), and therefore we conducted an additional analysis to test for interaction effects between construction and plausibility. The model specification is shown in Equation (8). We ran the model twice. The first time, for the variable Construction, we coded 1 for the Initial-position Intransitive and Final-position Intransitive conditions, and 0 for the Transitive condition, followed by subsequent scaling. For the variable Plausibility, we coded 1 for implausible and 0 for plausible, also followed by subsequent scaling. The second time, for the variable Construction, we coded 1 for the Initial-position Intransitive and Transitive conditions, and 0 for the Final-position Intransitive condition, followed by subsequent scaling.

$$\begin{aligned} Response \sim\ & Construction * Plausibility \\ & + (1 + Construction * Plausibility \mid Participant) \\ & + (1 + Construction * Plausibility \mid Item) \end{aligned} \quad (8)$$

Results and discussions

Figure 4 presents the proportions of literal interpretations, while the results of Model (6) and Model (7) are summarized in Table 5. Consistent with the findings from Experiment 1 and Experiment 2, the literal interpretation rates for the plausible conditions in Experiment 3 were high (Transitive: 0.935; Final-Position Intransitive: 0.861; Initial-Position Intransitive: 0.895). These results provide very strong evidence in favor of literal interpretation for plausible conditions compared to implausible conditions, as indicated by the Bayesian analysis.

Regarding the implausible conditions, Bayesian analysis provided strong evidence that participants interpreted implausible transitive sentences more literally compared to both implausible Initial-position intransitive and implausible Final-position intransitive sentences. This finding aligns with predictions, as the transitive condition involves an exchange across the main verb, while the intransitive conditions involve exchanges across function words.

Furthermore, there was no evidence of differences between the Implausible Final-Position Intransitive and Implausible Initial-Position Intransitive conditions. This result also aligns with predictions, as both conditions are derived from the same operation: an NP exchange across a function word.

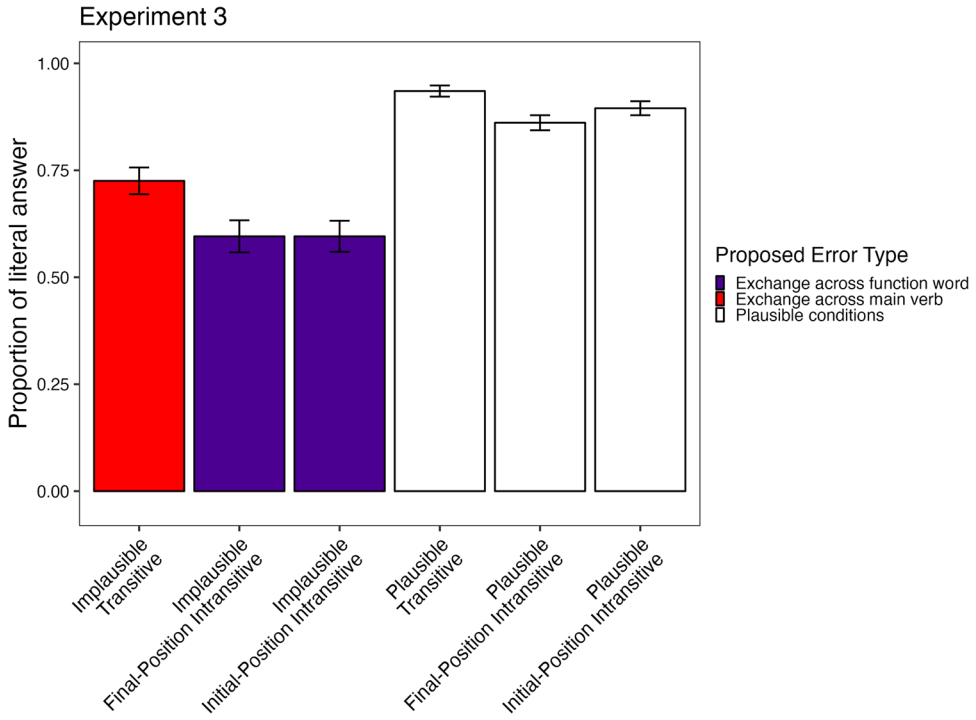

Fig. 4 Experiment 3: Percentage of trials where participants relied on the literal syntax for interpretation of the sentences. *Error bars* represent 95% confidence intervals

The only difference between these two implausible intransitive conditions lies in the position of the adverbial phrase, which may serve different discourse functions. For example, (5e) may place greater emphasis on the subject, while (5f) may highlight the location where the subject was dissolved. This pattern might suggest that function-word-based exchanges may take precedence over information structure effects within noisy-channel processing.

The results of the interaction analysis are summarized in Table 6. Bayesian analysis provided strong evidence that implausible conditions elicited less literal interpretations than plausible conditions. There was also strong evidence that the Initial-position Intransitive condition showed lower literal interpretation rates than the Transitive condition, and that the Final-position Intransitive condition also showed lower literal interpretation rates than the Transitive condition. In addition, there was modest evidence that the Initial-position Intransitive condition elicited higher literal interpretation rates than the Final-position Intransitive condition. For the interaction between Construction and Plausibility, there was no evidence supporting an interaction effect. The absence of an interaction effect, together with the presence of a main effect, suggests that the observed pattern is at least partially driven by structural priors (Poliak et al., 2023; Poppels & Levy, 2016). Within the noisy channel framework, comprehenders integrate their prior expectations about syntactic constructions with the incoming input. Less frequent or more marked constructions might carry lower prior probability and

Table 5 Mean results, range of the 95% credible interval, and probability of beta being different from zero for the Bayesian models in Experiment 3

Comparison in Experiment 3	**Estimate**	**95% Credible Interval**	$p(\beta \neq 0)$
Comparison 1			
Plausibility (Plausible vs. Implausible)	2.48	[1.42, 3.62]	1.000
Comparison 2: Initial-position Intransitive as reference			
Construction(Final-position Intransitive vs. Initial-position Intransitive)	0.01	[-0.44, 0.46]	0.518
Construction (Transitive vs. Initial-position Intransitive)	0.93	[0.41, 1.47]	1.000
Comparison 2: Transitive as reference			
Construction (Final-position Intransitive vs. Transitive)	-0.86	[-1.33, -0.40]	1.000
Construction(Initial-position Intransitive vs. Transitive)	-0.85	[-1.42, -0.29]	0.998

Table 6 Mean results, range of the 95% credible interval, and probability of beta being different from zero for the Bayesian models in Experiment 3

Interaction in Experiment 3	Estimate	95% Credible Interval	$p(\beta \neq 0)$
Transitives as reference			
Plausibility(Implausible vs. Plausible)	-2.73	[-3.84, -1.69]	1.000
construction(Initial-position Intransitive vs. Transitive)	-0.89	[-1.48, -0.31]	0.998
construction (Final-position Intransitive vs. Transitive)	-1.16	[-1.73, -0.63]	1.000
construction(Initial-position Intransitive vs. Transitive) x Plausibility	-0.05	[-1.24, 1.17]	0.540
construction (Final-position Intransitive vs. Transitive) x Plausibility	0.60	[-0.42, 1.66]	0.874
Final-position intransitives as reference			
Plausibility(Implausible vs. Plausible)	-2.41	[-3.39 , -1.51]	1.000
construction (Transitive vs. Final-position Intransitive)	1.38	[0.55, 2.33]	0.999
construction(Initial-position Intransitive vs. Final-position Intransitive)	0.50	[-0.18, 1.24]	0.924
construction (Transitive vs. Final-position Intransitive)x Plausibility	-0.43	[-1.46, 0.52]	0.803
construction(Initial-position Intransitive vs. Final-position Intransitive)x Plausibility	-0.52	[-1.38, 0.33]	0.888

be more prone to reinterpretation, even when the sentence is plausible.

Across-experiment data analyses

In the previous section, we conducted two main analyses: (1) a comparison between implausible conditions and plausible conditions within each experiment, and (2) a comparison between implausible conditions within each experiment. All the models are summarized in Comparison 1 and Comparison 2 of Table 7. The results were generally compatible with the predictions of the noisy-channel theory.

In this section, we compare the literal interpretation rates between conditions rendered implausible by different operations across experiments. Specifically, we examine the literal interpretation rates for three types of implausible conditions: those created by exchanges across the main verb (Active, Initial-position PO, and Final-position PO), those created by exchanges across a preposition (Passive, Ba, Initial-position Intransitive, and Final-position Intransitive), and those created by deletions or insertions (DO and Final-position PO). Insertions and deletions were grouped together because there is no evidence to suggest a difference between these two operations in both Zhan et al. (2023) and our study. These three types of constructions were collectively referred to as "alternation", which was entered as the fixed effect in the model. Items and participants were included as random intercepts with random by-participant and by-item slopes for alternation. The data were analyzed using a Bayesian generalized linear mixed-effects model in R, implemented with the brms package. The model specification is presented in Comparison

Table 7 Comparisons and Corresponding Analyses

Comparison1: Implausible vs. Plausible within each experiment	**Analyses**
Exp1	
Exp2	Response∼Plausibility + (1+Plausibility | Participant)+(1+Plausibility litem)
Exp3	
Comparison2: Comparison between implausible conditions within each experiment	**Analyses**
Exp1	
Exp2	Response∼Construction + (1+Construction|Participant)+(1+Construction|item)
Exp3	
Comparison3: Comparison between implausible conditions made by different operations across experiments	**Analyses**
Exchange across main verb(Active+Initial PO+Transitive) vs. Exchange across function word(Passive+Ba+ Initial intrans+Final intrans) vs	Response∼Alternation + (1+Altenation|Participant)+(1+Altenation|Item)
Deletion/Insertion (DO+ Final PO)	

3 of Table 7. Dummy coding was applied to the independent variable, with the reference level coded as 0 and the other levels coded as 1. This allowed for straightforward interpretation of the fixed effects. A normal prior distribution N(0,2) was specified for the fixed effect coefficients, representing a prior belief that effect sizes are centered around zero with a standard deviation of 2. All parameter estimates were initialized to 0 to enhance convergence stability. The model made use of parallel processing with 8 cores to accelerate computation. Four Markov chains were executed, each running for 4000 iterations, to generate a sufficient number of posterior samples and ensure robust convergence diagnostics. The model was run twice: first with NP exchange across the function word as the reference level, and the second time with NP exchange across the verb as the reference level.

Predictions

For the between-experiment predictions, we hypothesized that, across the alternations, all conditions requiring an NP exchange across a main verb would exhibit a higher literal interpretation rate than conditions requiring an NP exchange across a function word. Additionally, conditions involving deletions or insertions were expected to show the lowest literal interpretation rates. Specifically, conditions (2d), (3f), and (6d) were predicted to have higher literal interpretation rates than conditions (2e), (2f), (6e), and (6f), while conditions (3d) and (3e) were expected to demonstrate the lowest rates.

Results and discussion

The overall result pattern graph is shown in Fig. 5, and the results of the comparison between implausible conditions made by different operations across experiments are summarized in Table 8.

Our results provide very strong evidence that both implausible conditions resulting from NP exchanges across a main verb (Mean=0.582) and those resulting from NP exchanges across a function word (Mean=0.640) were interpreted more literally than conditions generated by deletions or insertions (Mean=0.162), which is consistent with the predictions of noisy-channel theory.

There is also evidence that implausible conditions involving NP exchanges across a function word tend to be interpreted less literally than those involving exchanges across a main verb, as predicted. However, we can see the remarkably low literal interpretation rate observed in the implausible Initial-position PO condition, even though it is formed by NP exchange across the main verb–just like the Implausible Active and Implausible Transitive conditions in the other two experiments. There are at least two possible explanations for this discrepancy. First, it might reflect a case of substitution rather than exchange. For instance, in the implausible Initial-position PO sentence (3f) "爸爸给一部手机送了儿子" (Father to a mobile phone sent the son), participants might substitute "gei" with another function word such as "*ba*". Additionally, as previously noted, "*gei*" can function as both a preposition meaning "to" and as a main verb meaning "*give*". Consequently, condition (6f) may also be interpreted as (5). Thus, the proposed implausible Initial-position PO condition(6f) was perceived as plausible by Mandarin participants, rendering it ineffective for testing the noisy-channel hypothesis. It is therefore necessary for future studies to explore this condition using another word that solely functions as the preposition "to", such as the Mandarin preposition "xiang".

General discussion

Several papers have now examined the impact of noise operations on English sentence comprehension through the

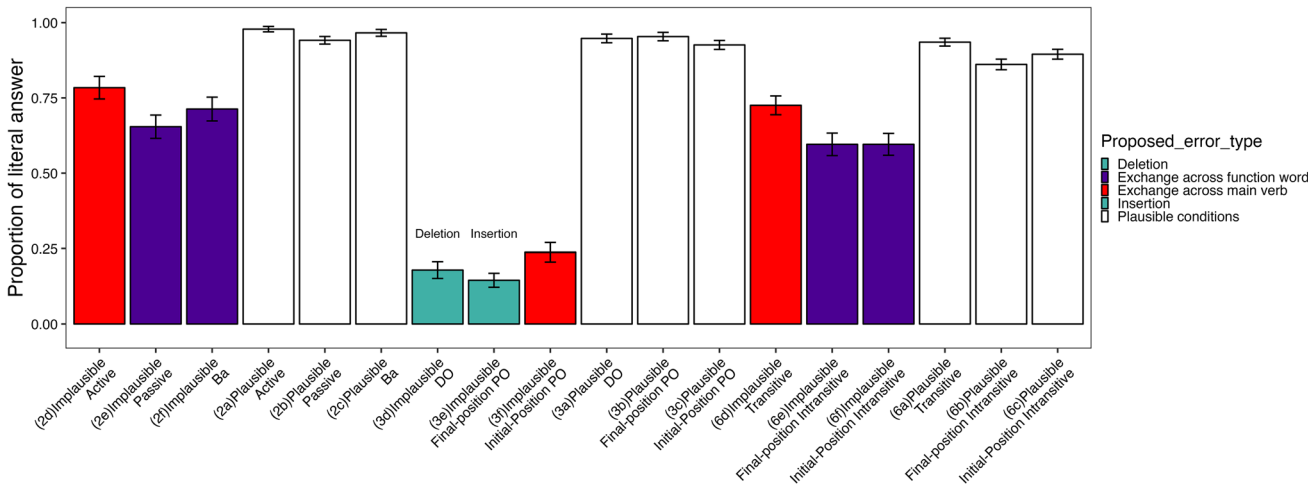

Fig. 5 Overall graph of three experiments

Table 8 Mean results, range of the 95% credible interval, and probability of beta being different from zero for the Bayesian models in across-experiment comparison

Comparison	Estimate	95% CI	Probability
Run 1: NP Exchange across Function Word as Reference			
Deletion/Insertion vs. Exchange across Function Word	-1.37	[-2.22, -0.65]	0.999
Exchange across Verb vs. Exchange across Function Word	0.34	[-0.13, 0.77]	0.931
Run 2: NP Exchange across Verb as Reference			
Deletion/Insertion vs. Exchange across Verb	-1.28	[-2.09, -0.57]	1.000
NP Exchange Function Word vs. Exchange across Verb	-0.36	[-0.79, 0.12]	0.935

lens of the noisy-channel theory, but the number of studies testing other languages is limited. This study aimed to extend the noisy-channel model of sentence comprehension to Mandarin Chinese by examining three sets of syntactic alternations: Active–Passive–Ba sentences, DO–Initial position PO–Final position PO, and Transitive–Initial position Adverbial Intransitive–Final position Adverbial Intransitive sentences.

Our results replicated the patterns observed in English (Gibson et al., 2013) and the previous Mandarin study (Zhan et al., 2023), with Mandarin participants showing a similar tendency to make inferences for implausible sentences that are more easily converted into plausible ones through fewer and simpler edits. Specifically, sentences that required a single deletion or insertion elicited the most inferences, while more intricate modifications, such as noun phrase exchanges across a function word, led to fewer inferences. The most complex edit, involving noun phrase exchanges across the main verb, resulted in the fewest inferences. These parallels across languages underscore the robustness of the noisy channel theory in explaining sentence processing.

However, we also observed three unexpected patterns. The first unexpected pattern is the difference between the implausible Passive condition and the implausible Ba condition, despite both error types involving NP exchanges across function words. In the discussion part of Experiment 1, we attributed this difference to the frequency disparity between the two constructions in Mandarin, where Ba constructions are generally more frequent than Passive constructions. However, if we consider substitution as a potential error type, the observed difference between the Ba and Passive conditions might also be explained within this framework. While Poliak et al. (2023) explored potential misinterpretations in Russian and argued that substitution errors are unlikely, the characteristics of Mandarin make substitution more plausible. As a non-alphabetical language with no spaces between characters, Mandarin lacks clear word boundaries, increasing the likelihood of single-character substitutions. Phonologically, both "Bei" and "Ba" start with the same initial consonant sound (/b/), which enhances their perceptual similarity. Moreover, both constructions are highly frequent in Mandarin and are often discussed as complementary structures, making them cognitively accessible and familiar to native speakers. Therefore, if both conditions involve substitution rather than exchange, the higher frequency of Ba structures compared to Bei structures could still account for the difference in literal interpretation rates.

The second unexpected pattern concerns the lack of difference between implausible DO sentences and implausible Final-position PO sentences, even though repairing the former typically involves insertions and the latter deletions. Surprisingly, the numeric values even showed a tendency opposite to our predictions. We attribute this result to the possible misinterpretation of "给" (gei) as a main verb rather than a preposition. Participants might interpret "给" (gei) as "give" rather than "to", leading to an interpretation where the structure is perceived as plausible rather than implausible.

The third unexpected pattern is the notably low literal interpretation rate for the implausible Initial-position PO condition, despite it being formed by NP exchange across a main word, just like other conditions in Experiment 1 and Experiment 3. One possible explanation is that "给" (gei) was also misinterpreted as the main verb "give" rather than a preposition. Additionally, this pattern might also be explained by considering substitution rather than exchange, where instead of performing an NP exchange, participants might have substituted "给" (gei, give) with verbs like "用" (yong, use) or "拿" (na, take). This substitution would result in a different sentence structure that appears more plausible to the comprehender, leading them to directly interpret the sentence as plausible. These possibilities highlight the complexity of interpreting substitution errors and underscore the need for further investigation into whether substitution should be treated as an independent error type in Mandarin.

In addition, we partially replicated the results in Zhan et al. (2023) regarding the transitive/intransitive sentences. Both our study and Zhan et al. (2023) found that implausible transitive sentences were more likely to be interpreted literally, compared with initial-position implausible intransitive sentences. However, there were discrepancies in how likely participants interpreted each type of sentence: in Zhan et al.

(2023), participants were more likely to interpret plausible intransitive sentences literally (> 90% literal interpretation rate) than in this study. In addition, participants in this study were more likely to interpret implausible sentences literally (> 60% in Fig. 4), compared to Zhan et al. (2023) (40% and 60% for implausible initial-position intransitive and implausible transitive, respectively). This could possibly be due to the idiosyncrasy of participants, since the two studies were conducted on different platforms and data were collected years apart.

A reviewer raised the concern that some misinterpretations might occur in the comprehension questions, particularly the implausible ones. We examined this possibility and concluded that such cases are unlikely. In Experiments 1 and 2, misinterpreting the implausible questions would require an NP exchange across the main verb–an error not predicted by the noisy channel model. In Experiment 3, such questions are even less likely to trigger noisy-channel inferences, as they lack close, plausible alternatives that could result from a likely edit.[2]

We present our findings through the lens of Noisy-Channel Theory, but they are also broadly consistent with the Good-Enough Processing Theory (Ferreira, 2003; Ferreira & Lowder, 2016). The Good-Enough Theory posits that language comprehension often relies on quick, approximate interpretations rather than detailed syntactic analysis. Instead of constructing fully accurate representations for every sentence, people frequently use heuristic processing to capture the general meaning, especially when precise parsing is not essential. This approach highlights a practical balance between efficiency and adequate understanding, prioritizing communicative success over grammatical perfection. For example, in garden-path sentences like "While Ana dressed the baby played in the crib", readers often misinterpret "dressed the baby" as a complete unit before realizing the mistake. Similarly, in passive sentences such as "The dog was bitten by the man", people frequently mistake the dog as the agent, reflecting a preference for surface-level processing when syntax becomes complex or conflicts with common sense. Such errors illustrate how people tend to favor fast, efficient processing rather than exhaustive analysis. Noisy-Channel Theory, while distinct in its formalization, aligns with this good-enough perspective. This theory suggests that when linguistic signals are noisy or ambiguous, the language processor reconstructs the most plausible interpretation based on prior knowledge and contextual cues. In this way, Noisy-Channel Theory can be seen as a more fine-grained and quantitative extension of the good-enough idea, modeling how initial interpretations may dominate when syntactic structures are prone to disruption. Together, these theories underscore how human language processing is shaped by practical communication needs rather than exhaustive accuracy, allowing for rapid and adaptive responses in everyday interactions.

In conclusion, we presented experimental findings from three types of syntactic alternations in Mandarin, which are broadly consistent with previous English studies and with the Mandarin findings of Zhan et al. (2023). These cross-linguistic parallels underscore the robustness of the noisy-channel model, while also highlighting the importance of language-specific properties in shaping comprehension and inference patterns.

Author Contributions Ruihua Mao: Conceptualization, Methodology, Formal analysis, Writing-original draft. Sihan Chen: Conceptualization, Methodology, Validation, Writing-review & editing. Edward Gibson: Conceptualization, Methodology, Writing-review & editing, Supervision, Project administration.

Funding The work was supported by a grant from the National Science Foundation (No. 2121074) "CompCog: Noisy-channel processing in human language understanding" to E. Gibson.

Data Availability All the raw data and materials are available at: https://osf.io/7usq6/?view_only=1bf8a922f8974a06a5bc6daf4fb57d0b.

Code Availability The code used for data analysis in this study is available at: https://osf.io/7usq6/?view_only=1bf8a922f8974a06a5bc6daf4fb57d0b.

Declarations

Conflict of Interest The authors have no conflicts of interest to disclose.

Ethical Approval This research received approval from the Committee on the Use of Humans as Experimental Subjects at the Massachusetts Institute of Technology.

Consent to Participate Participants completed the study through a web interface that presented an informed consent form before the study began. By proceeding, they provided their informed consent to participate.

Consent for Publication Not applicable.

[2] **Experiment 1, Item 13**
这个 钱包 捡到了 某个东西／某人 吗?
this wallet pick.up-ASP some-CL thing/person Q
'Did this wallet pick up something/someone?'
(it can be made plausible by an NP exchange across the main verb)

Experiment 2, Item 40
丝绸 收到了 某个东西／ 某人 吗?
silk receive-ASP some-CL thing/person Q
'Did the silk receive something/someone?'
(it can be made plausible by an NP exchange across the main verb)

Experiment 3, Item 61
太阳 是否 由于 什么 融化了?
sun whether due.to what melt-ASP
'Did the sun melt due to something?'
(no plausible close alternatives)

References

Bürkner, P.-C. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1):1–28. https://doi.org/10.18637/jss.v080.i01.

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395-411. https://doi.org/10.32614/RJ-2018-017.

Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, *100*(5):1-54. https://doi.org/10.18637/jss.v100.i05.

Chen, S., Nathaniel, S., Ryskin, R., & Gibson, E. (2023). The effect of context on noisy-channel sentence comprehension. *Cognition, 238*, Article 105503. https://doi.org/10.1016/j.cognition.2023.105503

Dodd, N., Boush, F., Leung, T., Ferreira, F., & Morgan, E. (2024). Noisy-channel processing in standard arabic relative clauses. In: *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 46. https://escholarship.org/uc/item/6vj2f252.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive psychology, 47*(2), 164–203. https://doi.org/10.1016/s0010-0285(03)00005-7

Ferreira, F., & Lowder, M. W. (2016). Prediction, information structure, and good-enough language processing. In: *Psychology of learning and motivation*, volume 65, pages 217-247. Elsevier. https://doi.org/10.1016/bs.plm.2016.04.002.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition, 6*(4), 291–325. https://doi.org/10.1016/0010-0277(78)90002-1

Garrett, M. F. (1975). The analysis of sentence production. In: Bower, G. H., editor, *Psychology of Learning and Motivation*, volume 9, pages 133-177. Academic Press. https://doi.org/10.1016/S0079-7421(08)60270-4.

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences, 110*(20), 8051–8056. https://doi.org/10.1073/pnas.1216438110

Gibson, E., Tan, C., Futrell, R., Mahowald, K., Konieczny, L., Hemforth, B., & Fedorenko, E. (2017). Don't underestimate the benefits of being misunderstood. *Psychological science, 28*(6), 703–712. https://doi.org/10.1177/0956797617690277

Gibson, E. A. F. (1991). A computational theory of human linguistic processing: Memory limitations and processing breakdown. Phd thesis, Carnegie Mellon University.

Hale, J. (2001). A probabilistic earley parser as a psycholinguistic model. In: Second meeting of the north american chapter of the association for computational linguistics. https://aclanthology.org/N01-1021/.

Keshev, M., & Meltzer-Asscher, A. (2021). Noisy is better than rare: Comprehenders compromise subject-verb agreement to form more probable linguistic structures. *Cognitive Psychology, 124*, Article 101359. https://doi.org/10.1016/j.cogpsych.2020.101359

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006

Levy, R. (2008b). A noisy-channel model of rational human sentence comprehension under uncertain input. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, pages 234-243. Association for Computational Linguistics. https://aclanthology.org/D08-1025/.

Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.

Liu, Y., Ryskin, R., Futrell, R., & Gibson, E. (2020). Structural frequency effects in comprehenders' noisy-channel inferences. In *Proceedings of the 26th Architectures and Mechanisms for Language Processing Conference*. Poster presentation.

Lv, S. (1999). *Eight hundred words in Modern Chinese*. Beijing: Commercial Press.

Makowski, D., Ben-Shachar, M. S., and Lüdecke, D. (2019). bayestestr: Describing effects and their uncertainty, existence and significance within the bayesian framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541.

Poliak, M., Kimura, H., & Gibson, E. (2024). Mis-heard lyrics: an ecologically-valid test of noisy channel processing. In: *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 46. https://escholarship.org/uc/item/3mf978x8.

Poliak, M., Ryskin, R., Braginsky, M., & Gibson, E. (2023). It is not what you say but how you say it: Evidence from russian shows robust effects of the structural prior on noisy channel inferences. Journal of experimental psychology. *Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0001244.

Poppels, T., & Levy, R. (2016). Structure-sensitive noise inference: Comprehenders expect exchange errors. In: *CogSci*. https://escholarship.org/uc/item/7vx580zr.

R Core Team. (2021). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ryskin, R., Futrell, R., Kiran, S., & Gibson, E. (2018). Comprehenders model the nature of noise in the environment. *Cognition, 181*, 141–150. https://doi.org/10.1016/j.cognition.2018.08.018

Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE, 37*(1), 10–21. https://doi.org/10.1109/JRPROC.1949.232969

Tsao, F.-F. (1987). A topic-comment approach to the ba construction. *Journal of Chinese Linguistics*, pages 1-54.

Wang, Y. (2011). The acquisition of the mandarin "gei" by cantonese-speaking learners. Master's thesis, The Chinese University of Hong Kong.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as bayesian inference. *Psychological review, 114*(2), 245. https://doi.org/10.1037/0033-295X.114.2.245

Xun, E.-D., Rao, G.-Q., Xiao, X.-Y., & Zang, J.-J. (2016). Da shuju beijing xia bcc yuliaoku de yanzhi [development of the bcc corpus in the context of big data]. *Yuliaoku Yuyanxue [Corpus Linguistics], 3*(1), 93–109.

Zhan, M., Chen, S., Levy, R., Lu, J., & Gibson, E. (2023). Rational sentence interpretation in mandarin Chinese. *Cognitive Science, 47*(12), Article e13383. https://doi.org/10.1111/cogs.13383