

Noisy-channel Bayesian Inference in Mandarin Listening Tasks

Abstract

Previous research has found that English speakers may infer that an implausible utterance is the result of noise processes and adopt a more plausible but nonliteral interpretation of the sentence they believe the speaker intended to say. Several studies have found evidence in support of English sentence comprehension as rational Bayesian inference across a noisy channel. However, this phenomenon is underexplored in non-English languages such as Mandarin Chinese. Additional data may serve to illuminate whether noisy-channel processing is a universal language processing faculty and if there are differences in its function across languages. Here, we explore the case of word substitution and exchange errors using a sentence recall paradigm and provide experimental evidence for Bayesian priors toward plausible decoding for sentences taking on the active/passive verb forms and past/future sentence forms, but relatively noise-free decoding on time/location sentence structures. Findings from this work suggest the need for follow-up studies to replicate these results on a larger scale.

Keywords: rational analysis; noisy-channel comprehension; non-literal interpretation; Mandarin Chinese

Introduction

In the past, multiple experiments have shown that comprehenders (readers or listeners) may adapt non-literal interpretations of sentences whose literal reading is unlikely due to inserted, deleted, substituted, or exchanged words (Gibson, Bergen, & Piantadosi, 2013; Poppels & Levy, 2016). For example, the sentence “The mother gave the candle the daughter” is more likely to be interpreted as “The mother gave the candle *to* the daughter” or “The mother gave the *daughter* the *candle*” than the literal meaning of the candle receiving the daughter (Poppels & Levy, 2016). This phenomenon is known as noisy-channel communication. Intuitively, a human listener has priors of how plausible or implausible the literal interpretation of a sentence is going to be. Depending on how strong the priors are, this affects the likelihood of them interpreting the meaning of a sentence literally or decoding it to a more plausible interpretation. Studying noisy channel communication provides insight into how humans parse language and has interesting implications for improvement of natural language models, which struggle with human-level error detection and correction.

Previous work in this domain has centered around English-language listening comprehension tasks. In this work, we hope to extend these findings to other languages, namely Mandarin Chinese. Different languages have syntactic differences that affect what sorts of noise operations can convert

plausible sentences to syntactically sound but semantically implausible sentences (Cai, Pickering, & Branigan, 2012), so it is worth examining whether comprehenders’ noise inferences reflect these differences and whether the noisy channel account is a valid explanation of comprehender behavior in non-English languages in the first place.

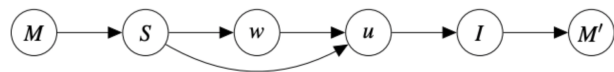


Figure 1: Schematic of noisy channel communication (Poppels & Levy, 2016). M is the intended meaning, which is encoded in structure S with surface form w . As Poppels & Levy demonstrate, S and w jointly determine the actual utterance u , which is processed to produce the input to the comprehender I .

It has been theorized that the computation that human comprehenders perform is rational in terms of Bayesian inference, meaning that we aim to model human judgements using Bayes’ rule. Specifically, from Figure 1, comprehenders infer $P(M'|I)$ from $P(I|M)P(M)$, where M is the literal meaning of the interpreted phrase, I is the input to the interpretation process, and M' is the inferred meaning. $P(I|M)$ is the noisy channel, and $P(M)$ is the prior.

Meanings M and inputs I may be *plausible* or *implausible*. To simplify the task, for a certain *plausible* meaning M , we only allow the corresponding input I to be the same plausible interpretation (the original message) or the corresponding implausible interpretation (the ‘noisy’ or ‘repair’ message), and vice versa for an *implausible* meaning M .

We hypothesize that the prior $P(M)$ places a higher probability on the plausible meanings of any given message. If this is true, implausible inputs will be more likely to be decoded as the plausible interpretation than the other way around. However, the strength of this prior would decrease if there are many implausible prompts presented in close succession, especially in an experimental setting where out-of-the-ordinary conditions are expected (Gibson et al., 2017; Ryskin, Futrell, Kiran, & Gibson, 2018; Arehalli & Wittenberg, 2021).

Our contributions include: (1) evidence that inferences made in human language comprehension are Bayesian ratio-

nal in languages other than English, (2) evidence for word substitution but not content word exchanges as noise operations admitted by speakers of Mandarin Chinese (3) further evidence for noisy channel communication as a model for the universal capability of language comprehension.

Related Work

Theories of sentence processing are an important subfield within the study of linguistics and human cognition. Earlier theories made the simplifying assumption that the input to language processing is an error-free sequence of words. Gibson et al. (2013) expanded the theory of human sentence comprehension to offer an explanation for how humans can understand language given imperfect input and how human language processing might be well-designed to recover intended meaning from noisy input by introducing the idea of language comprehension via a rational Bayesian noisy-channel language comprehender.

Poppels and Levy (2016) expanded the theory of noisy-channel language comprehension to provide evidence that the structure of the noisy channel allows noise in an utterance to be dependent on not only the surface string of words, but the underlying grammatical structure which produced them. They showed that English-speaking comprehenders will hypothesize positional exchanges as a plausible noise operation for function words but not content words, which is impossible without structural dependencies. Zhang, Ryskin, and Gibson (2023) explore noisy-channel comprehension using the case of depth-charge illusions.

Several efforts have been made to understand what factors can influence people’s Bayesian priors on noise operations. Gibson et al. (2017) showed that comprehenders adopt a higher noise prior and make noise inferences more often when a speaker has a foreign accent. Ryskin et al. (2018) showed that comprehenders can be primed to adopt a higher prior on *specific* noise operations, such as deletions, in addition to adjusting their overall error rate. Arehalli and Wittenberg (2021) investigated the effect of the ratio and structure of filler items on people’s judgments using white noise audio distortions rather than word-based speech errors.

Like most areas of linguistics, the majority of noisy-channel language comprehension research has been conducted in English. However, there have been some efforts to explore this phenomenon in other languages. Liu (2022) examined the case of subject-verb-object order in Mandarin Chinese. We hope to expand upon previous work and provide further evidence for noisy-channel language comprehension in Mandarin Chinese using different grammatical structures.

Methods

Participants were presented with a series of audio phrases that could be plausible or implausible on a literal interpretation. Some phrases, which we will call prompt phrases, had both a plausible and an implausible form, and were randomly selected to be either plausible or implausible in the

experiment. We used these phrases to track the priors of plausible/implausible interpretations. Additionally, we had filler phrases which were only present in one form, to prevent the proportion of implausible phrases from being too high.

After hearing each group of audio phrases, participants would answer a series of yes/no questions about whether they heard a list of selection phrases in the preceding audio. Selection phrases were a randomized list of distractor items which did not appear in the audio, filler items from the audio, and a random selection of the original prompt phrases and plausibility inversions of prompt phrases.

Differences in Mandarin and English

Mandarin Chinese differs from English in that many of the syntactic alternations used by Poppels and Levy (2016) can not be achieved by inserting/deleting/exchanging function words, but also require changing the order of content words. The direct-object/prepositional-object benefactive forms, for example, manifest as the following:

[SUBJECT] [VERB] 给了 [BENEFACTEE] [OBJECT]

in the DO form and

[SUBJECT] [VERB] 了 [OBJECT] 给 [BENEFACTEE]

in the PO form. (Example from (Cai et al., 2012))

Note that the indirect object/benefactee must follow the verb in the DO form, while the direct object must follow the verb in the PO form. The preposition 给 also ‘sticks’ to the benefactee.

Because content word exchanges were likely to also be an unallowed noise operation in Mandarin Chinese, we used some different grammatical forms for our experiment which could achieve the alternation by word substitution. We chose to include the time/location alternation, which is a single content word exchange, to test this hypothesis.

Prompt phrases

We considered three forms of plausible/implausible prompt phrases in this experiment.

Active/Passive The first form we considered was active/passive, which takes on the form

[SUBJECT] + [PREPOSITION] + [OBJECT] + [VERB]

in the active form and

[OBJECT] + [PREPOSITION] + [SUBJECT] + [VERB]

in the passive form.

Specifically, we consider the active/passive prepositions 把 and 被.

For a given sentence of active form

[SUBJECT] 把 [OBJECT] [VERB] 了

(“the [SUBJECT] [VERB]-ed the [OBJECT]”), the following passive form is equivalent:

[OBJECT] 被 [SUBJECT] [VERB] 了

(“the [OBJECT] was [VERB]-ed by the [SUBJECT]”).

If the above sentences are of the plausible form, the implausible form can be obtained by switching the active preposition to a passive preposition or vice versa, thus exchanging the subject and object roles, in each of the above cases. An example is included in Table 1.

Time/Location We also considered the time/location form, where the object of the sentence is specifically a location or time. Concretely, these sentences take on the format

[SUBJECT] + [VERB] + [LOCATION]

or

[SUBJECT] + [VERB] + [TIME].

In this case, the above forms are the plausible forms, and the implausible forms are when the object and subject are switched. In this case, every plausible sentence only takes on one form, and has one implausible analogue. For example, the sentence “小白生于北京” (Xiaobai was born in Beijing) is the plausible form, while “北京生于小白” (Beijing was born in Xiaobai) is the implausible form. Analogous examples using time are found in Table 2.

Past/Future The final form of prompts we considered was the past/future form. In English, these sentences take on the form

Yesterday [SUBJECT] [VERB]-ed [OBJECT]

or

Tomorrow [SUBJECT] will [VERB] [OBJECT]

in the plausible condition, and take on the form

Yesterday [SUBJECT] will [VERB] [OBJECT]

or

Tomorrow [SUBJECT] [VERB]-ed [OBJECT]

in the implausible condition. The sentences in Chinese are almost an exact translation of the English form (with yesterday=昨天, tomorrow=明天), an example of which is shown in Table 3. A plausible sentence is converted to an implausible sentence by deleting one tense particle and inserting the opposite tense particle.

In addition to the above prompt phrases that could take one of several forms, we included generic filler phrases that only had one form. A summary of the frequencies of all sentences used in the investigation is found in Figure 2.

Experimental Procedures

All participants of the study (n=11) were self-identified Mandarin speakers. Each participant was presented with a survey consisting of six rounds of experimentation. In each round, the participant was prompted with 6 audio phrases and asked to fill a series of 10 yes/no selection questions indicating whether or not they heard a given sentence in the audio round. We chose to use selection questions in lieu of comprehension questions as in Poppels and Levy (2016) due to the limited dataset size and the lack of variety in the pool of possible

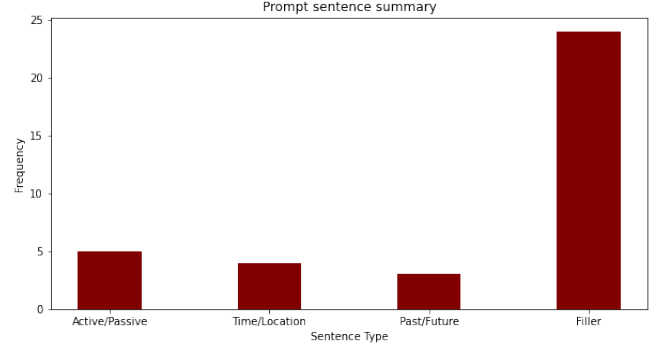


Figure 2: Summary of phrases used in experiment. Each active/passive phrase had four possible forms, each time/location phrase had two possible forms, and each past/future phrase had four possible forms.

comprehension questions, as well as the comprehension questions being overly contrived directly from the prompts themselves. This design also prevents respondents from looking back at the prompts when they make judgments about what they heard. Surveys were administered via Google Forms.

Each round of experimentation consisted of two prompt phrases and four filler phrases, making for a total of 12 prompt phrases per participant. When generating one round of an experiment, prompt phrases were selected at random and set to one of the possible forms with uniform probability. All audio phrases were repeated in the selection questions. Prompt phrases were flipped with 50% probability, meaning that there was a 50% chance that implausible phrases would become plausible phrases, and vice versa. If inverting the plausibility led to multiple equivalent phrases, one of the phrases was selected uniformly at random.

Results and Discussion

Figure 3 shows the accuracy, precision, and recall of participant responses to the selection questions. We hypothesize that this is because it is more phonetically distinct than the other 2 prepositions used in active/passive forms and was likely to be more prominently perceived and would skew the data when it was present.

We see that the filler phrases had the highest accuracy and precision, setting a baseline value. For the active/passive forms, plausible phrases had overall higher precision, accuracy, and recall than implausible phrases. This implies that for the active/passive forms, there is a strong prior on the plausible interpretation. Moreover, there is a phonetic aspect to this as well: The words 被 and 把 are similar-sounding, meaning that listeners are likely to assume that the preposition was misheard.

For the time/location forms, implausible sentences had a higher recall and accuracy than plausible sentences. This means that most of the implausible sentences were heard and identified in the selection process. The results from time/location forms differ from active/passive forms likely

Sentence	Translation	Form	Plausibility
男孩把热狗吃了。	The boy ate the hot dog.	Active	Plausible
热狗被男孩吃了。	The hot dog was eaten by the boy.	Passive	Plausible
热狗把男孩吃了。	The hot dog ate the boy.	Active	Implausible
男孩被热狗吃了。	The boy was eaten by the hot dog.	Passive	Implausible

Table 1: All active/passive forms for a given example sentence.

Sentence	Translation	Content	Plausibility
小白生于北京。	Xiaobai was born in Beijing.	Location	Plausible
北京生于小白。	Beijing was born in Xiaobai.	Location	Implausible
王陵生于2002年。	Wang Ling was born in 2002.	Time	Plausible
2002年生于王陵。	2002 was born in Wang Ling.	Time	Implausible

Table 2: Time/location form. Both time and location example sentences are shown.

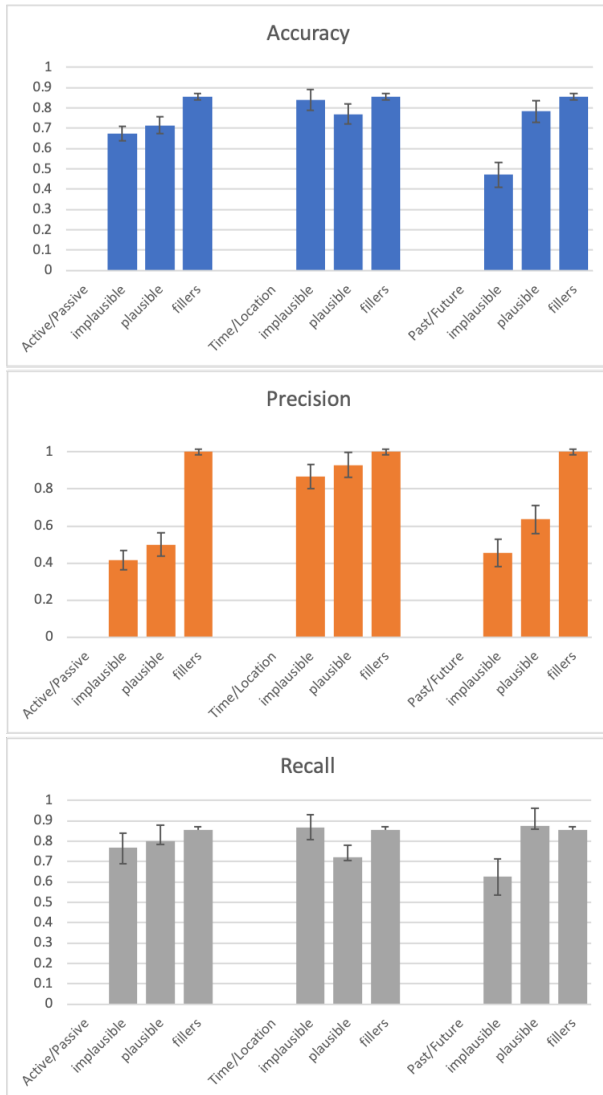


Figure 3: Summary of participant responses for each group of experiment items.

because the implausible form is more obvious, meaning that they were more likely to stick in the participant’s short term memory. Therefore, the priors for implausible sentences is probably stronger once one implausible sentence of the time/location form was heard. However, the differences between plausible and implausible are not statistically significant and all values were relatively high, meaning that time/location sentences were overall “easier” to identify for listeners.

For the past/future forms, we notice a similar pattern to the active/passive forms where the plausible items have a higher precision, recall, and accuracy rate than the implausible items. Again, this could be because the implausible sentences of the past/future forms are “harder” to identify as implausible from first listen, as prior logical knowledge would be needed to identify the sentences themselves as implausible. The logical (plausible) sentences were identified more easily than the illogical sentences.

Figure 4 shows a finer breakdown of accuracy rates for all the combinations of plausible/implausible audio and plausible/implausible selection item. Take note in particular of the plausible→implausible and implausible→plausible repairs.

For the active/passive forms, we can see that the lower accuracy rate for implausible audio items is indeed driven by a lower accuracy for the implausible→plausible case. This is evidence that people are indeed making noise inferences and hypothesizing that the intended sentence might be the plausible sentence, which has a higher prior probability, when presented with the option. Interestingly, this effect was not strong enough for people to claim they had not heard the implausible sentence when presented with the original implausible sentence as a selection option.

Similarly to the aggregated data, the effect, if it is there, is much weaker for the time/location forms. The accuracy rate for the implausible→plausible repair is slightly lower, but it may be noise in the data. These results may provide evidence that content word swaps are not in the inventory of noise operations considered by Mandarin speakers, or could be explained based on the string-edit prominence of the change re-

Sentence	Translation	Time	Plausibility
昨天我去了学校。	I went to school yesterday.	Past	Plausible
昨天我会去学校。	I will go to school yesterday.	Past	Implausible
明天我去了学校。	I went to school tomorrow.	Future	Implausible
明天我会去学校。	I will go to school tomorrow.	Future	Plausible

Table 3: All past/future forms for a given example sentence.

quired. Further experimentation using content words of varying length may be needed to confirm.

For the past/future forms, we see an even more dramatic effect and a much lower accuracy for the implausible→plausible repair, indicating that people were much more likely to hypothesize they had heard the plausible sentence after hearing the impossible audio than the other way around. Similarly to the active/passive forms, people did not go so far as to claim they had not heard the implausible sentence when presented with the original implausible sentence at the same frequency, though the accuracy rate for the implausible→implausible case is lower than the plausible→implausible repair.

Conclusion and Future Work

In this investigation, we explored Bayesian rational inference over a noisy channel as an explanation for noise inferences made by Mandarin speakers. We observed that Mandarin-speaking comprehenders behave in a Bayesian rational manner with a high prior on plausible sentences for active/passive and past/future forms, as plausible sentence items achieved higher precision, recall, and accuracy than implausible items; and people were more likely to say they had heard a plausible sentence after listening to a corresponding implausible sentences. For time/location, this effect was less noticeable since the participants were more easily able to identify both plausible and implausible forms.

The grammatical phenomena we investigated in this work are not sufficient to confirm whether or not noisy-channel processing in Mandarin Chinese is structure-sensitive. The word substitution and content word exchange sentences we used as experimental items can be explained entirely in terms of string-edit noise models, and future work will be needed to confirm whether structural sensitivity is present.

Future work would benefit from spending more time to search through papers on Mandarin Chinese linguistics and consult with language experts to select appropriate grammatical forms that can illuminate different noise operations and the presence of structural sensitivity, if it does exist.

With more time and a larger participant pool, it would also be interesting to perform different manipulations on the statistics of the filler items to examine the effect on people’s priors.

A long-term goal to strive for is to perform similar experiments on a large number of different languages using similar methods and use the data collected to draw conclusions about the role of noisy-channel comprehension in the universal human mechanism of language comprehension.

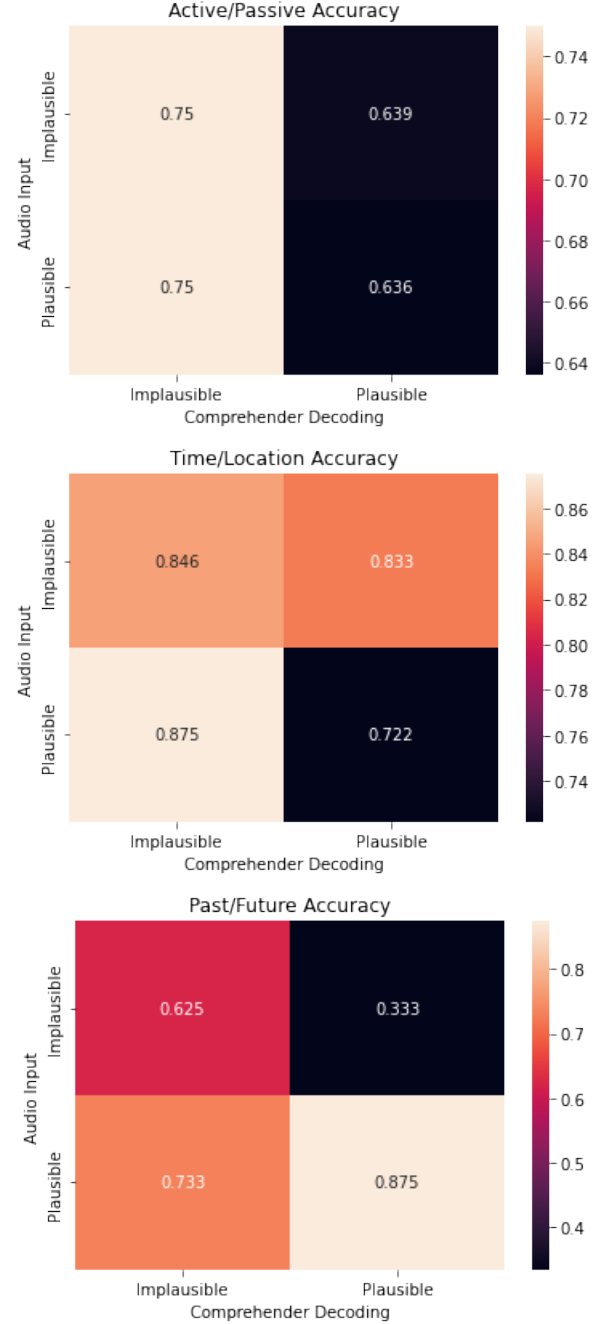


Figure 4: Breakdown of accuracy rates for plausible/implausible items that were decoded as plausible/implausible in the sentence participants had to identify as heard or not heard.

References

- Arehalli, S., & Wittenberg, E. (2021). Experimental filler design influences error correction rates in a word restoration paradigm. *Linguistics Vanguard*, 7(1).
- Cai, Z. G., Pickering, M. J., & Branigan, H. P. (2012). Mapping concepts to syntax: Evidence from structural priming in mandarin chinese. *Journal of Memory and Language*, 66(4), 833-849. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0749596X12000319> doi: <https://doi.org/10.1016/j.jml.2012.03.009>
- 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.
- 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.
- Liu, Y. (2022, Nov). *Structural frequency effects in comprehenders' noisy-channel inferences*. OSF. Retrieved from osf.io/h9tdf doi: 10.17605/OSF.IO/H9TDF
- Poppels, T., & Levy, R. (2016). Structure-sensitive noise inference: Comprehenders expect exchange errors. In *Cogsci*.
- Ryskin, R., Futrell, R., Kiran, S., & Gibson, E. (2018). Comprehenders model the nature of noise in the environment. *Cognition*, 181, 141–150.
- Zhang, Y., Ryskin, R., & Gibson, E. (2023). A noisy-channel approach to depth-charge illusions. *Cognition*, 232, 105346.