



## Introduction

- East Asian sentence-final particles (SFPs) express a range of subtle, speaker-oriented meanings that pertain to the way the speaker conveys her belief states to the listener (e.g. Chu 1998; Simpson 2014; Constant 2014). Specifically, we present a formal characterization of the Mandarin Chinese sentence-final particle *de* based on cross-entropy as apart of the Rational Speech Act model (Shannon 1948, Jäger 2007, Goodman & Stuhlmüller 2012)
- By motivate employing information-theoretic pragmatic methods as a way to obtain a fine-grained understanding of the non-literal meaning contributions expressed by speaker-oriented particles, the current work is in line with recent developments towards a complex interface approach to meaning, encompassing the speech act theory (e.g. Krifka 2020) and the realm of probabilistic semantics/pragmatics

## Data

*De* is attached to a proposition. The proposition can be used as:

- a complete answer (no subpart)
- a partial (non-complete) answer that addresses a subquestion in the immediate context

First, *de* optionally attaches to a complete resolving answer that fully addresses the current Question Under Discussion (QuD), with the speaker providing complete knowledge.<sup>1</sup>

Context: Who should I find if I want to get a refund this week?  
(1) zhezhou shi Cai laoshi fuze baoxiao *de*.  
this.week COP Cai teacher handle refund DE  
‘For this week, teacher Cai handles refunds.’

Second, the speaker with complete knowledge may also address the current QuD by supplying multiple contrastive partial answers. In such case the placement of *de* favors the partial answer that covers a larger number of situations, and disfavors the answer partitioning fewer situations.

(2) ??Zhouer yizhi dao zhouwu/tongchang, shi Wang laoshi fuze baoxiao (DE)  
Tuesday all.the.way until Friday/usually COP Wang teacher handle refund DE  
zhouyi/ou’er, shi cai laoshi fuze baoxiao (??DE).  
Monday/occasionally, COP Cai teacher handle refund DE  
‘From Tuesday until Friday teacher Wang handles refunds; On Monday, teacher Cai handles refunds.’

Third, when the speaker has only incomplete information that addresses the current QuD, she may attach *de* to a partial answer as long as she conveys with it all she knows.

(3) zhouyi shi Cai laoshi fuze baoxiao *de*, wo bu zhidao shengxia jitian de anpai.  
Monday COP Cai teacher handle refund DE I NEG know rest several.day REL arrangement  
‘On Monday, teacher Cai handles refunds *de*. I don’t know about the rest of the week.’

## Formal Implementation

Components needed to quantify informativity:

- a set  $T$  of all possible worlds (or equivalence classes of worlds, where two worlds are equivalent if they answer the QUD in the same way);
- a speaker, who holds a belief state  $S(t)$  over the worlds  $t \in T$  that she would like to convey;
- a listener, who forms a belief state  $L(t|m)$  over worlds  $t \in T$ , dependent on some message  $m$ .

Informativity could be measured as the amount of information provided about the speaker's belief state, representable as the negated cross entropy (Shannon 1948) between the speaker's and listener's belief states after the speaker communicates message  $m$ ,  $-H(P_s(\cdot), L(\cdot|m)) = \sum_{t \in T} P_s(t) \log L(t|m)$

- Utterances bringing the listener's belief state closer to the speaker's have a higher informativity
- Utterances which contradict worlds deemed possible by the speaker have a negative infinite informativity

- Let  $T$  be the set of all who's-in-charge assignments, each such assignment mapping one day of the workweek to one of two people who handle refunds
- Let  $S(t) = \delta_{t,s}$ , where  $s \in T$  is the unique world described in the example
- let  $L(t|m)$  be the **uniform** distribution over all worlds consistent with the literal meaning of  $m$  (a literal listener with a uniform prior over worlds)

The informativity of a message  $m$  is  $\log L(s|m)$ . If  $m$  is compatible with  $s$ , this is the negative logarithm of the number of worlds compatible with  $m$ .

As logarithms are monotonic, informativity can be measured by counting worlds: utterances with more compatible worlds are less informative.

**Explaining example (2):**

From Tuesday to Friday is compatible with 2 worlds; on Monday compatible with 16 worlds  
Thus the first partial answer is less uncertain and more informative  
*De* is predicted to be attached there!

With the quantifier *usually*, assuming it represents more than half of the weekdays, the proposition is compatible with 16 worlds.  
*Occasionally*, assuming a meaning of at least one weekday, the second proposition is compatible with 31 worlds.  
The first proposition is thus more informative. *DE* is predicted to be attached there!

## Compared with Entailment-based Approaches

Based on an entailment-based approach, it could be postulated that *de* simply requires its preajacent occupy the highest level of summation among the elements being considered. However, it faces the following difficulties:

- The events that each adverb quantifies over do not directly stand in a proper-part relationship, and hence are not comparable.
- It has nothing to say about cases where the presentational order of information affects the epistemic knowledge about the speaker belief state (i.e. a listener expects a conventional temporal/numeral order, but a speaker disobeys it).

## Extension

First, we extend to the cases where the speaker and the listener have an arbitrary common-knowledge prior.

For the example (2) with additional simplified assumptions where

- The speaker and listener both hold non-uniform priors, with higher probabilities assigned to worlds in which teacher Wang is in charge except the days when she is on vacation. • the speaker knows that the listener doesn't have the information when teacher Cai is on vacation.
- the probability of teacher Cai being on vacation per day is 1 percent, and each day is independent from each other,
- The result thus yielded an approximation -0.0174 after hearing the proposition ‘On Monday, teacher Cai is in charge of the refund’, larger than the one for ‘From Tuesday to Wednesday, teacher Wang handles refund’ (which equals to -2).

We predict that the *de* particle prefers to attaching to the utterance *On Monday teacher Cai handles refund*, since it is more informative in this situation (which coincides with native speakers intuition).

**Second**, a recent work from van Rooy (2004) argues that whether the **mention-some reading** suffices to resolve the question depends on the expected utility of the answers. Our proposal could be connected to the mention-some/all discussion, in a way the *de* particle could be potentially used as a diagnostic to see if the informativity actually influence the blocking of mention-some in plural-marked questions (Xiang 2017). Further scrutiny is necessary, since the maximal informativity we evaluated is exclusively from the speaker's perspective without recursion.

**Third**, The cross-entropy approach for *de* can be expanded to capture **sentence-final particles cross-linguistically** that are shown to compare the interlocutors' belief states (e.g. Cantonese *ge*, Japanese *no*, Malay *mia*). For example, Speaker A saw a butter on the table and complained that she doesn't like the salted butter, then B uttered in Japanese: “This is the butter without salt inside **no**!” The drastic difference between prior and posterior belief states (salted vs. unsalted) should be reflected by a high Kullback-Leibler (KL) divergence (Baldi 2002):  $H(L(\cdot|m), L(\cdot)) - H(L(\cdot|m))$  More explicitly, we could quantify how "far" the listener's belief state has shifted upon hearing this utterance, based on  $S(m) = -\log(L(\textit{unsalted}))$ , where  $L$  is the listener's prior.

## References

**Constant, Noah. 2014.** *Contrastive topic: Meanings and realizations*. University of Massachusetts dissertation. **Frank, Michael, Noah Goodman. 2012.** Predicting pragmatic reasoning in language games. *Science* 336.6084: 998-998. **Franke, Michael, Jäger, Gerhard, 2016.** Probabilistic pragmatics, or why Bayes’ rule is probably important for pragmatics. *Zeitschrift für Sprachwissenschaft* 35 (1), 3–44. **Noah Goodman, Andreas Stuhlmüller. 2013.** Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, 5(1):173-184.

<sup>1</sup>A pilot acceptability results are on OSF platform as an Appendix for the current work.