

Anaphora resolution in causal coherence relations in Mandarin Chinese

Jun Lyu & Elsi Kaiser (University of Southern California)

Introduction: In a sentence involving causal relations like “Jane angered Peter because...”, the ensuing pronoun typically refers to the *cause* of the event *Jane* (e.g., Hartshorne, 2014). However, the processing effort depends, among other factors, on the saliency of the intended antecedent. As a structurally salient constituent, the subject tends to be privileged for coreference, known as the *subject preference* (Frederiksen, 1981; Crawley et al., 1990). A less explored factor is *thematic role*. Previous research shows that thematic roles higher on the Thematic Hierarchy (Jackendoff, 1972/1987; Grimshaw, 1990) are more prominent, as a higher ranked role (e.g., agent vs. patient) often takes the subject position given syntactic flexibility, which makes one wonder if thematic-role-related prominence affects anaphora resolution. This study probes the influence of both subject preference and thematic role on real-time pronoun processing in Chinese.

Methods: In a self-paced reading (SPR) study (163 natives), we crossed *grammatical position* (subject/non-subject) of the cause and its *thematic role* (agent/patient) in the context sentence, followed by a target sentence (20 target items and 20 fillers). The pronoun in target sentences refers to the gender-unambiguous antecedent, the cause of the event (see (1) for an example). To promote patients to subject position, passive BEI-construction in Chinese was used. Note that the “agent” and “patient” in this study are defined in line with Dowty (1991) who outlined several properties of proto-agents/proto-patients but argued that agents and patients should be determined based on their similarities to proto-agents/proto-patients. Thus, the agents in this study are more agent-like relative to patients and patients are more patient-like relative to agents.

Predictions: The *subject preference account* predicts a main effect of *grammatical position* while the *thematic hierarchy account* predicts a main effect of *thematic role*.

Results: **Figure 1** shows the RT patterns (stats: *Imer* in R) for all conditions in the target sentence. For better visualization that isolates the impact of grammatical positions and thematic roles, see **Figure 2**. RTs for comprehension question response are plotted in **Figure 3**. The critical pronoun region 3 showed that both main effects and the interaction were significant ($p < .05$), because the non-subject/agent condition in (1b) elicited longest RTs compared to other conditions ($p < .001$). Region 4 revealed a main effect of *grammatical position* ($p < .01$) and a significant interaction ($p < .001$), because while patient-hood lead to reading slowdowns in subject conditions ($p < .001$), agent-hood lead to reading slowdowns in non-subject positions ($p < .01$). Region 5 is characterized only by a main effect of *grammatical position* ($p < .005$), demonstrating a subject preference. Statistical model for region 6 was not significant ($ps > .05$). Interestingly, the final two regions showed that patient conditions lead to faster RTs ($ps < .05$). These results overall only support the *subject preference account* at best, but not the *thematic role account*. The divergent thematic role effect at Region 4 is highly intriguing, shown clearly by Figure 2 (top row). In fact, based on the effect size and the timing of the effect, the following processing ease hierarchy can be derived: Subject/Agent > Subject/Patient = Non-subject/Patient > Non-subject/Agent. Finally, comprehension question response latencies showed that patient conditions overall lead to faster responses ($p < .05$), similar to the SPR reading patterns at the last two regions.

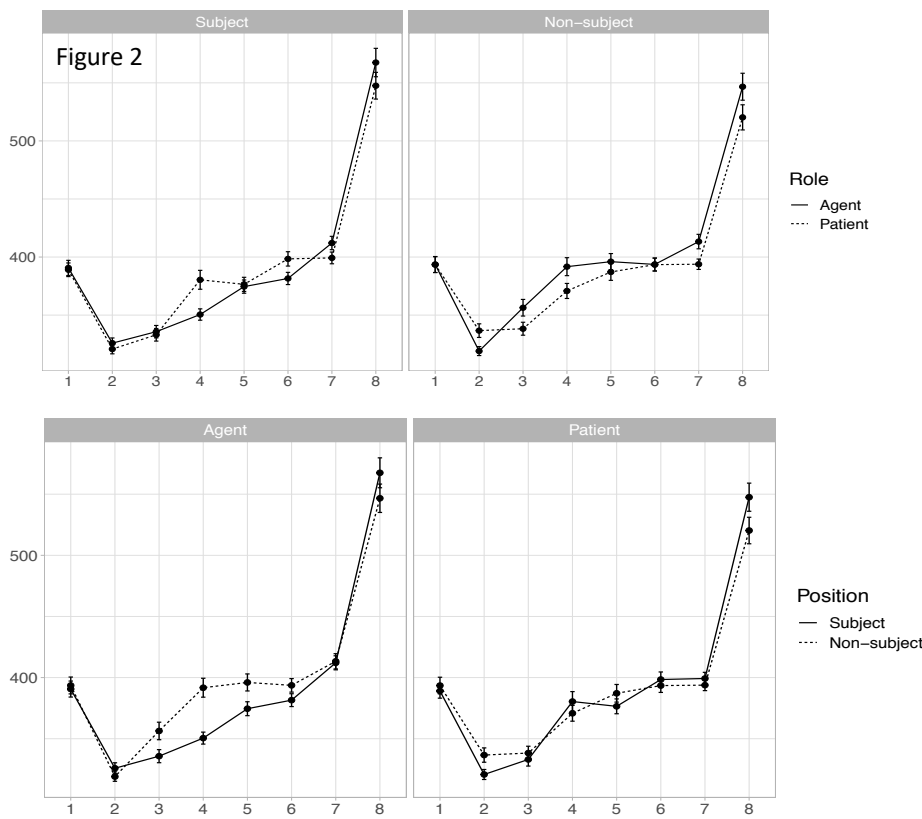
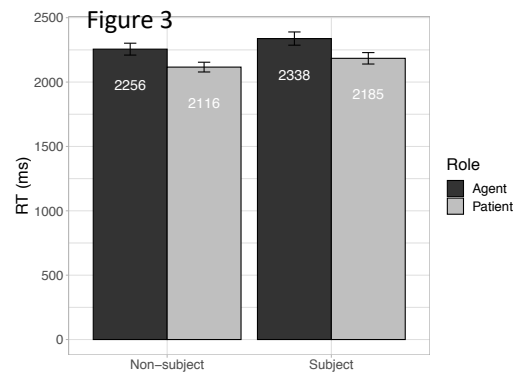
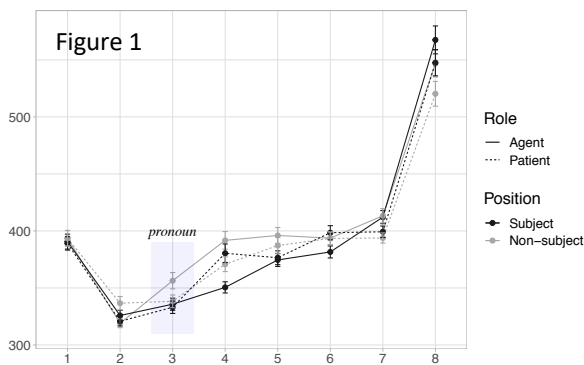
Discussion: To account for the unexpected contrast in Figure 2 (top row), we propose a Mapping Principle, similar to Ferreira (1994): an agent in the sentence must be mapped to the subject position and a non-agent must be mapped to the non-subject position. Thus, when the pronoun refers to the cause of the event that violates the Mapping Principle, parsing difficulty occurs. Crucially, the Mapping Principle alone cannot fully account for the data as the bottom-right panel in Figure 2 suggests an absence of the Mapping Principle penalty, as the subject/patient condition did not lead to longer RTs compared to the mapped non-subject/patient condition. However, as shown by Table 1, when Subject Preference and Mapping Principle are acknowledged to play independent roles, the processing ease pattern we observed earlier can be explained. As to the response latencies, we hypothesize that the semantic representation of the event has a “agent-verb-patient” configuration, which helps explain why retrieval of patients is easier.

- (1) Example target stimuli (translated from Chinese; 20 sets in total; left: context; right: target sentence)
- a. Subject/Agent: **Jane** upset Peter. } This is/ because/ **she**/ in/ the game/ not know/
 - b. Non-subject/Agent: Peter BEI **Jane** upset. } team/ work.
 - c. Subject/Patient: **Peter** BEI Jane blamed. } This is/ because/ **he**/ in/ the game/ not know/
 - d. Non-subject/Patient: Jane blamed **Peter**. } team/ work.

Table 1. Constraints active in pronoun resolution in a causal discourse.

Condition	Subject Preference	Mapping Principle	Penalty
a. Subject/Agent	-	-	0
b. Non-subject/Agent	1	1	2
c. Subject/Patient	-	1	1
d. Non-subject/Patient	1	-	1

Processing ease:
 a. Agent/Subject > c. Subject/Patient, d. Non-subject/Patient > b. Non-subject/Agent



Selected references

[1] Crawley, Rosalind A., Rosemary J. Stevenson & David Kleinman. 1990. The use of heuristic strategies in the interpretation of pronoun. *Journal of Psycholinguistic Research*.
 [2] Dowty, D. 1991. Thematic proto-roles and argument selection. *Language*.
 [3] Ferreira, F. 1994. Choice of passive voice is affected by verb type and animacy. *JML*.
 [4] Hartshorne, J. K. 2014. What is implicit causality? *LCN*.
 [5] Jackendoff, R. S. 1972. *Semantic interpretation in Generative Grammar*. MIT press.