

報告番号	※	第	号
------	---	---	---

主 論 文 の 要 旨

論文題目 Ambiguity in the Processing of Japanese, Korean and
Mandarin Chinese Relative Clauses
氏 名 MANSBRIDGE Michael Patrick

論 文 内 容 の 要 旨

Chapter 1

Numerous models of sentence processing have been constructed through the lens of the processing of relative clauses (RC). In RC processing, there is a near universal phenomenon demonstrating that subject-extracted relative clauses (SRCs) are easier to process and comprehend compared to their object-extracted relative clause (ORC) counterparts. Following this, many researchers have crafted hypotheses of processing in order to explain the processing asymmetry between SRCs and ORCs. Considering the amount of models to explain this effect, it is difficult to reconcile which models are valid accounts of processing as well as formulating a unified sentence processing theory.

A key method to empirically test models is to investigate them within a large amount of languages with varying typological features to determine if a model can accurately predicate processing in each language. In terms of RC processing, studies had classically investigated processing within European languages which are all post-nominal (i.e., the RC follows the head noun it modifies). Accordingly, there has recently been a surge in studies investigating processing within prenominal languages (i.e., the RC precedes the head noun it modifies). In the current study, RC processing was investigated in three East Asian languages, all of which have prenominal RC structures: Japanese, Korean, and Mandarin Chinese. However, all three of these languages have a confounding factor inherent to each language and their respective structure. Specifically, this factor is an initial clause-type ambiguity. In other words, since these languages lack relative pronouns or markers at the left edge (i.e., the start or boundary of the clause), a simple sentence (i.e., matrix clause) interpretation will likely be taken instead of an RC one. This misparse will continue until a viable cue to disambiguate the ambiguity of clause-type: The head noun in Japanese, the adnominal marker in Korean, and the relativizer in Mandarin. While recent studies have begun to use unambiguous RCs in these languages, there are a relatively few addressing it as an actual processing factor. Consequently, it is of empirical importance to

investigate the manner in which ambiguity influences the processing of RCs in each language to better explain how processing unfolds. As such, the main aim of this study was to investigate RC processing in each language as a factor of ambiguity to explain how processing mechanisms interact with one another.

In terms of processing factors, this study limited the discussion to three prominent factors which have already been integrated into cue-based retrieval models. The purpose of this was to detail the interrelatedness of these three factors: integration of filler-gap dependencies, expectation-based processing (i.e., frequency effects) and similarity-interference. Also, a fourth factor, outside the scope of cue-based retrieval, was introduced since it is a more recent theory for Japanese and Korean RCs: The object-before-subject-bias (OBSB).

The current study utilized eye-tracking methods on native speakers of each of these languages. Chapters 2-4 detail the processing of RCs in each language. This study asks three basic questions: i. Are ORCs more difficult to process than SRCs for each of the three languages? ii. Does (i) change as a factor of ambiguity? iii. Which processing models account for the data? Are they influenced by ambiguity and how are they related?

Chapter 2

This chapter detailed the processing of relative clauses in Mandarin Chinese. Mandarin Chinese is a relatively unique language with an exceptionally rare language typology in regards to relative clauses. Mandarin is an SVO language with prenominal relative clauses (i.e., the RC precedes the head noun). Because of this pattern, ORCs in Mandarin follow the canonical SVO word order of Mandarin while SRCs instead follow a non-canonical VOS. In terms of other languages, SRCs typically follow the canonical word order while ORCs do not. Even though this typological feature is divergent, Mandarin SRCs are still more frequent than ORCs in corpora similar to that of most other languages.

While Mandarin has a relativizer particle (i.e., not a relative pronoun), it appears frequently at the right edge or boundary of the clause. Consequently, since there are no cues to encode an RC interpretation at the left edge or boundary, a parser will instead likely interpret the clause as a simple matrix clause. In other words, RCs in Mandarin are initially ambiguous. This would cause a strong garden path effect that would likely facilitate the processing of ORCs, despite their increased structural complexity, processing difficulties surrounding them, and being the less frequent structure.

Previous experimental studies on Mandarin RCs have revealed mixed results. Sometimes ORCs were found to be easier to process and comprehend and other times SRCs were easier. However, these past studies had the ambiguity as a confounding factor. Those who found ORC advantages claimed that due to a linear integration/retrieval metric for filler-gap dependencies, ORCs are easier since the gap is more local to the filler (i.e., the head noun) in terms of linear distance, compared to SRCs. On the other hand, those who found SRC advantages instead claimed that structural accessibility of the subject position and increased frequency of the RC structure should make SRCs easier than ORCs. Furthermore,

some researchers contend that those ORC advantages found were mere reflections of the garden path effect. More recently, however, researchers began exploring RCs in Mandarin using less ambiguous sentence designs. While the relativizer cannot be fronted, other syntactic cues that can create clause boundaries and increase the RC interpretation were used. These studies found that when Mandarin RCs were unambiguous, ORCs were more difficult to process and comprehend in comparison to SRCs as a result of their lower frequency. However, there were still marginal indications that ORCs were initially easier to process.

The current study on Mandarin RCs, explored processing in terms of ambiguity. Two separate eye-tracking studies were conducted using native speakers of Mandarin from Mainland China, opposed to Taiwan or other countries. Experiment 1 used a strictly ambiguous design while Experiment 2 introduced ambiguity as an experimental factor. This study showed that both ORCs and SRCs have different processing requirements depending on the locus and time course during reading. The results revealed that ORC reading was possibly facilitated by linear/temporal integration and canonicity. On the other hand, similarity-based interference made ORCs more difficult, and expectation-based processing were more prominent for unambiguous ORCs. Overall, RC processing in Mandarin should not be broken down to a single ORC (dis)advantage, but understood as multiple interdependent factors influencing whether ORCs are either more difficult or easier to parse depending on the task and context at hand.

Chapter 3

This chapter detailed the processing of Korean relative clauses. Korean is an SOV language with prenominal RCs. In Korean, SRCs are more frequent than ORCs. While Korean has an adnominal marker suffixed to the embedded RC verb (i.e., the right edge), there are no RC cues at the left edge of the clause. Therefore, there is initial clause type ambiguity for Korean.

Previous studies on Korean RCs show that, with respect to processing, ORCs are more difficult to process at the head noun than SRC within temporarily ambiguous contexts. ORCs, however, are predicted by expectation-based processing models for surprisal to incur a greater processing cost during early processing stages at the RC verb, since it is a likely locus of disambiguation for RCs in Korean, and because ORCs are a less frequent structure compared to SRCs. Consequently, the current study investigates whether processing difficulty for ORCs manifests itself at the RC verb using eye-tracking methods, both ambiguous and unambiguous RCs, and both a sentential-decision task and comprehension task. The results revealed significantly increased go-past reading times for ORCs at the RC verb and head noun for both ambiguous and unambiguous RCs. We believe this is a result of a less frequent structure being more difficult to parse during disambiguation. Additionally, the findings at the head noun support an integration metric based on structural-phrase integration rather than linear/temporal metrics. Ambiguity, however, had little effect on the overall processing pattern for Korean RCs. The lack of

influence from ambiguity may be attributed to Korean having distinct loci for disambiguation and integration. Consequently, expectation-based models of processing accurately predicted difficulty for ORCs at the locus of disambiguation in Korean.

Chapter 4

This chapter detailed the processing of Japanese relative clauses. Japanese is an SOV language with prenominal RCs. In Japanese, SRCs are more frequent than ORCs. Japanese RCs have no RC markers or cues within typical externally-headed relative clauses. Therefore, the head noun itself acts as the cue to provide the RC interpretation. Because of this clause-type ambiguity, a garden path effect would occur up to the head noun. Since embedded noun-modifying structures lack overt marking at the left boundary or edge of the clause, it can be initially difficult to distinguish matrix clauses from RCs and complement clauses (e.g., *-koto*, *-to iu*, *-jijitsu*). Accordingly, the parser only becomes aware of the RC structure at a locus of disambiguation which happens to be the head noun.

While there are previous studies which have eliminated this ambiguity resulting in an attenuated difficulty at the head noun, these studies, however, have not effectively investigated the classical processing asymmetry between subject- and object-relatives during reading comprehension. In the current study, eye-tracking was employed on native Japanese speakers to determine how RC processing is influenced by ambiguity. The results revealed different processing patterns for each ambiguity context. Specifically, for the ambiguous RC context, ORC difficulties were observed during overall measures of processing at the head noun and RC verb which can be attributed to a variety of processing factors such as structural-integration, similarity-interference, expectation and the object-before-subject-bias. In contrast, when the clause structure was unambiguous, ORC difficulty was mainly observed during early processing measures at the head noun. This shift in processing behaviour is indicative of effects of expectation-based processing, i.e., since ORCs are less frequent than SRCs, they are more difficult to process. Thus, when the clause becomes less ambiguous, expectations become a more salient processing feature. Overall, RC processing in Japanese is influenced by the level of ambiguity.

Chapter 5

This chapter detailed the relationship of the findings from Chapters 2-4. In this study, several language independent and dependent findings were observed. Notably, for all three languages, the chief processing difficulties were centred at the locus of disambiguation as predicted by expectation-based processing. For Mandarin and Japanese, but not Korean, ambiguity was observed to be a factor influencing the processing of RCs. When RCs were unambiguous, the effects from expectation-based processing became more pronounced. Integration effects were also observed for all three languages at the predicted locus of integration; however, if the clause became less ambiguous, these effects became

less observable for Japanese and Mandarin. For both Japanese and Korean, the results supported a structural-integration metric while a linear metric was supported in Mandarin. This may suggest that these metrics are parametric between languages or that the head noun retrieves the *wh*-operator instead of the gap in these two languages. For all three languages, some level of similarity-interference was observed at the head noun making ORCs more difficult. For Mandarin in particular, ORC advantages were observed regardless of ambiguity which suggests that the regularities of the language may support processing despite parsing a less frequent structure. While, the above findings fit within the framework of cue-based retrieval, the model needs to be updated to account for regularities facilitating processing.

OBSB only had little support for ambiguous RCs in Japanese and unambiguous RCs in Korean, which should not be subject to OBSB effects. As such, OBSB may likely only contribute little to the overall ORC processing difficulties for Japanese and Korean.

Chapter 6

In conclusion, the combined findings from all three studies revealed that multiple factors contribute to the overall processing of RCs in these languages. Specifically, memory-constraints (i.e., integration/retrieval decay, similarity-based interference) and expectation-based processing were shown to be interrelated factors. These factors caused fluctuation in the activation of the RC structures and lexical dependencies that either facilitated processing or increased processing difficulties. Furthermore, ambiguity was largely seen as a factor for both Japanese and Mandarin, but not Korean. The reasoning for this is that in the former two, the locus of disambiguation also serves as the locus of integration, but in Korean, in contrast, these two are at separate loci. These findings fit nicely into the framework of cue-based retrieval models. Also, cue-based parsing was also supported since unambiguous RCs were not shown to reveal expectation effects until the locus of disambiguation (i.e., the principal cue for the RC interpretation). However, there was one additional finding that was counter to normal cue-based parsing effects. That is canonical order facilitation in Mandarin for unambiguous RCs. Thus, canonical order facilitation should be included in the framework of cue-based parsing in contexts which are (i) unambiguous, but (ii) have principal cues responsible for extracting frequency expectation at the right edge of the clause. This factor may be unique for languages like Mandarin however. Overall, RC processing should be viewed as an interrelated model of various processing mechanisms.