Cognitive Aspects of Relative Clause Production In Oral and Written Narratives

Ming-Ming Pu

UNIVERSITY OF MAINE, FARMINGTON

mingpu@maine.edu

Abstract

The present study aims to investigate relative clause production in Chinese and English from a cognitive-function approach that explores underlying cognitive, semantic, and discourse-pragmatic factors operative in discourse processing. With this approach we are able to account for both general and specific distributional patterns of RCs between the two languages on the one hand and between speech and writing on the other, using narrative data elicited from native speakers of both languages.

1 Introduction

The processing of relative clauses has been an important focus of research in various fields of linguistics and psychology over the past two decades. Such clauses, with their structural complexity and morpho-syntactical differences among diverse languages, provide rich and unique data for linguistic analyses from typological studies investigating language universals to neurolinguistic experiments examining memory mechanisms in language comprehension.

The present study investigates further the well-documented processing asymmetry between subject and object relative clauses and uncovers important distributional patterns other than the asymmetry. While most prior relative-clause research has focused on studies of sentence or reading comprehension of a particular language, the present study is based on spoken and written narrative data from native speakers of Mandarin Chinese and American English, examining and discussing the occurrence and distribution of relative clauses between these two morpho-

syntactically different languages on the one hand and between speech and writing on the other.

2 Prior research

In general, prior research on relative clauses (RCs) has found that subject relative clauses (SRCs) are easier to process than object relative clauses (ORCs), with the exception[1] of Hsiao & Gibson (2003) and Carreiras et al. (2009). A number of important psycholinguistic theories have been proposed to explain the effect of the processing asymmetry between SRCs and ORCs in English such as the working-memory accounts, integration cost accounts, and word order accounts. The memory-based theories (Frazier & Fodor 1978, Gibson 1998, 2000, Lewis 1996) consider working memory load as the determining factor for the processing difficulty of ORCs: There are a larger number of linguistic elements that have to be retained in an ORC than an SRC until the sentence can be grammatically structured for comprehension, thus generating a heavier memory load. The integration cost theories (Ford 1983, Gibson 1998, 2000, Hsiao & Gibson, 2003) claim that the subject-object asymmetry is due to the integration of, or linear distance between, a head NP and its trace in the RC: There are more intervening elements or referents between the head NP and its trace for an ORC than an SRC, hence a larger integration cost for the former. The word order hypotheses (Bever 1970, Prideaux & Baker 1986, Tabor et al. 1997) posit that utterances of a language that assume the canonical word order of the language should be easier to process. In English, ORCs have a non-canonical order of OSV, as compared to the normal SVO order of SRCs, and are hence harder to process.

The study of Chinese RCs has also attracted considerable attention. Both Li & Thompson (1981) and Chu (1998) have a chapter respectively on RCs in their discussion of the Chinese grammar: the former examines semantic and pragmatic factors involved in the formulation and interpretation of RCs, the latter focuses on the grounding status of such clauses in discourse. More recently, Pu (2007) and Tao (2002) both find that RC distribution in Chinese discourse is quite skewed and discuss the preferred and dispreferred RC patterns in terms of discourse functions. While Pu demonstrates that these patterns result from the interaction of cognitive constraints with semantic and discourse features of RC constructions, Tao shows that one type of RCs, i.e., temporal RCs, occur most frequently because they serve the discourse function of marking episode boundaries rather than expressing temporality, a device indispensable in oral narratives.

Moreover, distributional patterns of RCs are found different between oral and written discourse. In his quantitative study investigating differences between spoken and written language, Biber (1988) has shown that in English discourse RCs occur more frequently in writing than

---

[1] Hsiao & Gibson (2003) and Carreiras et al. (2009) reported an ORC preference in Chinese and Basque respectively.

speaking because out of the two potential positions for RCs, subject and object, only the latter tends to be utilized in speeches. Further, both *that*- and *which*-RCs modifying object heads occur more frequently than those modifying subject heads in general, and for those RCs on object positions, *that*-RCs occur more frequently than *which*-RCs in spoken English. Though typologically different from English, Chinese RCs exhibit a similar distributional pattern: RCs occur more frequently in written than oral discourse, and RC types used in the former is different from that in the latter (Chen 1997, Pu, 2007).

3 Factors underlying RC processing

The present study explores (a) how cognitive operations conspire with semantic and discourse-pragmatic factors to generate the often skewed distributions of RC patterns in discourse, as observed in many languages, and (b) how general cognitive strategies such as easing memory burden and avoiding ambiguity interact with language-specific features that result in different distributional patterns between oral and written discourse. The study is usage-based and data-driven, which compares and discusses the occurrence and distribution of RC types in Chinese and English discourse, both oral and written, to show why and how the distributional patterns arise and function the way they do in both languages.

3.1   Chinese and English RCs

Chinese and English have the same basic word order of SVO (Huang & Chui, 1997, Li & Thompson, 1981) and share certain features of relativization. The RC construction in either language consists of a head NP and a modifying subordinate clause in which the relativized NP is zero in form. The differences, however, are more than the similarities, as illustrated by the following Chinese example[2] with its English translation.

1) nage [Ø$_i$ bei   xiaohai *de*] lüke$_i$ mai.le wan fang.bian.mian
   that   bring child   traveler bought bowl ramen.noodle
   The traveler$_i$ [(who$_i$) carried a child] bought a bowl of ramen noodles.

The RC formation in (1), e.g., *nage [Ø$_i$ bei xiao.hai de] lüke$_i$* (Lit.:'that [Ø$_i$ carried a child *de*] traveler$_i$'), exhibits some remarkable differences from that of English *the traveler$_i$ [who$_i$ carried a child]*. The first is the linear position of RC: Unlike an English RC that follows its head NP, a Chinese RC precedes its NP. The second is the use of relative pronoun such as 'who' or 'that/which': Whereas an English RC may have an obligatory or optional relative pronoun, a Chinese RC lacks such a relative pronoun and is marked by the particle *de* at the end of the RC. The third is the placement

---

[2] All Chinese examples are rendered in *pinyin* in the first line and followed by a word-for-word gloss in English in the second line, with the English translation of the utterance at the end.

of a head NP determiner such as articles, demonstratives, and classifiers. In English such a determiner always immediately precedes its head NP (i.e., Det+NP+RC); in Chinese it may be separated from the head by a RC, as in (1), where the determiner *nage* (roughly equivalent to 'that' in English) is separated from its head *lüke* ('traveler'), rendering a split head NP (i.e., Det+RC+NP), although a unified head NP (i.e., RC+Det+NP) such as the one in (2) also occurs in Chinese discourse.

> 2) ta renshi [$\emptyset_i$ qi che *de*] nage ren$_i$
> he know ride bike that person
> He knew that person [(who) was riding a bike].

In (2), the determiner *nage* immediately precedes its head *ren* ('person'), just as it would in English.

It is clear that the different positions of RC with respect to its head in Chinese and English would result in distinct embedding structures. English RCs follow their head NPs, and hence those modifying a grammatical subject are often the imbedding kind, interrupting the main clause processing (i.e, S[RC]V(O)), whereas those modifying other grammatical roles (e.g., object, object of preposition, subject complement or predicate nominal, and subject in an existential structure) are the non-embedding kind. Chinese RCs, on the other hand, precede their head NPs, and thus most RCs would interrupt the main clause (e.g., those modifying object, object of preposition, or subject complement). While the only possible non-embedding RCs are those that modify a grammatical subject at sentence-initial position, they become interruptive when their head NPs are the split kind. Examples (1-2) both have interrupting RCs although they modify a subject and an object head respectively. The English renditions of both examples, on the other hand, show that (1) has an interrupting RC and (2) does not.

3.2   Cognitive strategies

We argue that clause-level syntactic analysis alone is not sufficient nor adequate to explain the observed preference for one type of RC structure over another because RCs are not processed as isolated dependent clauses but used mainly as a grounding and reference-tracking device in discourse (Chen 1997; Fox & Thompson 1990, Givón 1993, Pu 2007) and thus have to be studied in the rich discourse context of the entire NP+RC or RC+NP construction. As the speaker uses a relative clause structure at a certain juncture of discourse, it is important for her/him to consider the relative accessibility of the head NP and decide what type of RC to modify it so as to help the interlocutor, albeit mostly subconsciously, recognize her/his communicative intent with a minimum justifiable mental effort. In this decision-making process, the accessibility, the semantic features, and the discourse-pragmatic status of the head NPs as well as the processing ease of the RC all play an important role.

First of all, minimum justifiable mental efforts have to do with processing easy of utterances. Generally, language units that do not tax our cognitive resources should be easier and faster to process. In RC processing two cognitive strategies would interact with grammatical properties of RCs, yielding certain types of constructions easier to process than others. One strategy has to do with our working memory (and storage) limitations, and the other depends on our experience with various structures of utterances, both of which reflect how our mind, with limited resources, processes information recruitment and usage. The two cognitive strategies are CLOSURE and NORMAL FORM (see also Andersen 1989, Frazier 1979, Gibson 1998, 2000, King & Kutas 1995, Prideaux 2000).

> CLOSURE: A constituent (e.g., clause) that is not internally interrupted by another constituent or discourse entity will require less processing resources (i.e., will be less difficult to process) than that same constituent that is internally interrupted.

CLOSURE is a generalized Gestalt principle that hinges on working-memory limitations: Because we have only limited cognitive resources to allocate to a processing task at hand, CLOSURE would facilitate our assembly of one processing unit as quickly as possible so we can clear our active (verbatim or syntactic) buffer to deal with the next unit. In contrast, a unit containing another embedded unit does not allow for early closure; we would have to suspend processing one unit to attend to the other, hence increasing demands on working memory.

There is ample evidence in the psycholinguistic literature that supports the general processing strategy of CLOSURE. Prideaux and Baker (1986), for example, have found in their experimental study that English speakers and writers tend to use RCs to modify sentence objects than subjects because object-modifying RCs do not interrupt the matrix clause while subject-modifying RCs are center-embedded, which impose heavier memory loads for the sentence completion. They have also demonstrated that embedded units that interrupt the early closure of a sentence, be they RCs, subject clauses or prepositional phrases, are harder to process in both production and comprehension tasks. The preference for non-embedded RCs, especially in oral discourse production, has also been reported in Biber (1988), who explains that the tendency for RCs to occur with object heads in speaking is due to the processing ease of such RCs because "the speaker must contend with real-time production constraints" (p. 156), and the late closure of a clause is cognitively more demanding. Further, studies in other languages such as Japanese, Korean, Hungarian, Ukranian (Lynkowsky 1980, MacWhinney & Pléh 1988, Prideaux & Baker 1986) have also found the tendency for speakers to use non-embedded RCs, regardless of the word order and whether the RC precedes or follows its head NP in a language.

The other strategy, NORMAL FORM or markedness, addresses the ease with which our mind deals with canonical, familiar patterns rather than novelty and deviation (Fenk-Oczlon, 2001; Givón, 1993; Prideaux, 2000, Tabor et al. 1997). Givón (1993) suggests three criteria for the theoretical construct of markedness: structural complexity, discourse distribution, and cognitive complexity. He argues that the unmarked or canonical case, the general norm, should be structurally simpler or neutral, more frequent in discourse, and easier to process, whereas the marked case, the counter norm, should be structurally more complex, less frequent in discourse, and harder to process. Hence we have

> NORMAL FORM: The unmarked member of a set of forms is the general norm. It tends to be structurally simpler, more frequent in discourse, and should be easier to process than the corresponding marked form.

Take English for example. An SRC with the normal word order of [RPs (subject) VO], as in *the cat [that chased the dog],* is the unmarked form, and an ORC with the non-normal order of [RPo (object) SV], as in *the cat [that the dog chased]* is the marked form. NORMAL FORM is similar to the linguistics experience account, which argues that infrequent structures in a given configuration would be difficult to activate because of competition with more available frequent structures (Gennari & MacDonald 2008, McRae et al. 1998, Tabor, et al. 1997, Trueswell et al. 1994). In a series of experiments, Prideaux and Baker (1986) has found that in English SRCs are judged more natural, read and responded to faster, and used more frequently than ORCs by participants in various production and comprehension tasks. In Hungarian, too, SRCs are found easier to process than ORCs in the canonical word order of SVO/SOV (MacWhinney & Pléh 1988).

Moreover, the present study contends that both discourse-pragmatic and semantic factors conspire with the two cognitive strategies in RC production, which motivates speakers to use a particular RC construction to modify a given head. The former has to do with the accessibility of the head NP as reflected in the information status, topicality and discourse function of the head and its relative clause, and the latter includes semantic properties of the head NP such as humanness, agentivity, saliency, and referentiality. The degree of the mental accessibility and semantic information of a head NP would influence not only the encoding of its syntactic role, definiteness, refentiality, and positioning of the head NP but also what kind of modifying RC would be selected for the head and how often this construction would occur in discourse (Fox & Thompson 1990, Pu 2007).

We thus proposed a cognitive-functional principle to account for the occurrence and distribution of various RC constructions in discourse:

> The cognitive constraints of memory demand conspire with the independently motivated semantic and discourse-pragmatic factors

to determine not only the observed prevalent asymmetry between subject and object RCs but also the specific distributional patterns of RCs in a particular language with respect to speech and writing.

Upon encoding a complex NP (i.e., an NP with its modifying RC), the speaker would choose a RC that grounds its head NP in a certain way, based on her/ his assessment of the head NP accessibility (in the hearer's mind) and discourse function of the RC. The resulting structure would best serve its communicative purpose (specific discourse function) while satisfying a minimum justifiable mental effort on the part of both speaker and hearer (relative processing ease).

4 A narrative Study

The Chinese and English RC data used in this study came from an experimental (narrative) study in which native speakers of both languages were asked to watch a 5-minute video clip, *The New Doorbell,* and then describe it in either spoken or written form. The video clip is a cartoon movie about a man who installs in his apartment a new doorbell that plays a melody and then waits anxiously for visitors to ring it. However, his mood turns from joyfulness to disappointment and finally to anger because no one rings his doorbell. The clip is a silent color video with background music; no written language ever appears on the screen except the title, *The New Doorbell*, which is shown at the beginning of the movie in both Chinese and English.

Thirty native speakers of each language participated in the narrative study, all university undergraduates of non-language majors, about half of which were male students. The participants were randomly assigned to two groups of equal number: one group recounted the video clip in spoken language and the other in written form. The oral data were tape-recorded and later transcribed, and the written data were collected immediately after participants finished writing.

4.1 Data

All RCs found in our data modify either a grammatical subject (including subject complement and existential subject) or object (including object of preposition). Table 1 summarizes the RC types in both languages, where S stands for subject, SC for subject complement or existential subject, O for direct object and OP for object of preposition.

Table 1. Relative Clause Distribution Data

Language	Subject head			Object head			Total	
	S	SC	(%)	O	OP	(%)	N	(%)
Chinese	33	7	(57.14)	30	0	(42.86)	70	100.00
English	21	8	(44.62)	27	9	(55.38)	65	100.00
Total	54	15	(51.11)	57	9	(48.89)	135	100.00

There are four RC types in terms of grammatical structure—SS, SO, OS and OO: the first letter indicates the grammatical role of the head NP, and the second the type of relative clause that modifies the head. The following examples illustrate each type of RCs (placed within square brackets).

A. SS (subject RC modifying subject head):

   3) The first person [that came into the hall] was a little girl.
   4) There was this lady [who lived upstairs].

B. SO (object RC modifying subject head):

   5) The video [that we watched] was about a man and his new doorbell.

C. OS (subject RC modifying object head):

   6) A man installs a new doorbell [that plays music].
   7) She then proceeds up to her apartment [which is directly above his].

D. OO (object RC modifying object head):

   8) He signed the letter [the postman gave him].

4.2 Overall results

For the Chinese narrative, relative clauses are used much less frequently in the oral than the written task, and the pattern is reversed for the English narrative. The difference in frequency distribution between Chinese oral and written narratives is statistically significant (x^2=32.9*, p<0.001), while that between the English narratives is not (x^2=1.86, p>0.1). Nonetheless, the distributional patterns among different types of RCs are quite similar across modalities and languages. Table 2 tallies the raw data in RC types and their percentages for each of the experimental conditions.

Table 2. Overall RC Production

	SS		SO		OS		OO		Total	
	N	%	N	%	N	%	N	%	N	%
Chinese										
Oral	7	63.64	0	0.00	2	18.18	2	18.18	11	100.00
Written	31	52.54	2	3.40	14	23.73	12	20.34	59	100.00
Subtotal	38	54.28	2	2.86	16	22.86	14	20.00	70	100.00
English										
Oral	16	42.11	4	10.53	12	31.58	6	15.78	38	100.00
Written	9	33.33	0	0.00	9	33.33	9	33.33	27	100.00
Subtotal	25	38.46	4	6.15	21	32.31	15	23.08	65	100.00
Total	63	46.67	6	4.44	37	27.41	29	21.48	135	100.00

Even at a glance, the asymmetry between SRC and ORC in both languages is obvious: SRCs (SS and OS types) occur much more frequently in the narrative, a staggering 77% (54 out of 70 tokens) in Chinese and 71% in English (46 out of 65 tokens), indicating that normal form may be an important strategy used by speakers in general in RC production. The frequency difference between SRC and ORC is statistically significant in both languages (Chinese: $x^2=20.63^*$, p<0.001; English: $x^2=11.22^*$, p<0.001). Further, the general ranking order of RC distribution (oral+written) is the same for both languages: SS > OS > OO >> SO.

Nonetheless, the ranking order also reveals that the distributional rates within SRCs (between SS and OS) and ORCs (between OS and OO) are not the same, showing not only an SRC-ORC asymmetry but also an overall preference for SRCs that modify subject heads (SS) and a strong dispreference for ORCs that modify subject heads (SO). On the other hand, while the RC ranking order is almost the same between spoken and written narratives in Chinese, it differs between the two modalities in English. These specific distributional patterns are more complex than what can be explained solely by any of the theories discussed in §2, but can be accounted for by the proposed cognitive-functional principle.

4.3 English narrative

In addition to NORMAL FORM, the cognitive strategy of CLOSURE seems to operate robustly in RC production in English: Participants used more clause-final RCs than clause-medial RCs in both oral (63%) and written modalities (78%), as shown in Table 3.

Table 3. Clause-medial vs clause-final RCs in English

	Spoken			Written		
	Medial	Final	Subtotal	Medial	Final	Subtotal
SS	10	6	16	6	3	9
SO	3	1	4	0	0	0
OS	1	11	12	0	9	9
OO	0	6	6	0	9	9
Total N	14	24	38	6	21	27
%	36.84	63.16	100.00	22.22	77.78	100.00

It is interesting to note that the percentage of medial or interruptive RCs in the oral task is greater (37%) than that in the written task (22%), which appears to run counter to our assumption that the online oral production imposes a heavier cognitive demand than writing, hence fewer medial RCs should be found in the oral data. However, a further analysis of the data reveals three major differences in RC production between the two modalities, lending support to the proposed cognitive-functional principle in its account for RC distributional patterns. First, medial RCs, mostly the

SS type, function differently between the oral and written narratives: those in the former are used typically to modify GIVEN referents and the latter NEW. The following two examples are taken from the oral narrative.

9)   But the man [who installed the doorbell] goes and grabs
      the postman.
10) And the guy inside [that had put the doorbell on] got really angry.

In both utterances, the head NP of the RC encodes a known referent, i.e., the protagonist of the story, and the RC also provides information that has been established earlier in discourse. Since the online task is cognitively more strenuous, the speaker would try to help the hearer quickly locate and track referents by providing additional, though perhaps redundant, information about the identity of the referents under concern. This GIVEN+GIVEN information structure is typical of the SS utterances in the oral task, which could be easily comprehended in the unfolding discourse, thus easing the memory burden caused by center embedding. Of the 10 medial RCs in the SS structure in the oral data, 8 are the GIVEN+GIVEN type. In the written data, on the other hand, the overwhelming majority of SS utterances (8 out of 9) are of an entirely different kind, i.e., NEW+ NEW. For example,

11) A young man [who wore a pair of sunglasses] walked upstairs.
12) Next comes a short man [who knocks at the door].

Both utterances introduce a subject head that refers to a new referent in the story, and the RC in each sentence characterizes the new referent by describing what *he* looks like or does at the moment, which is also new information to the reader. The information density is hence much higher in the written narrative than that in the spoken one because writers/readers are not pressed for online processing, and hence are relatively freed from cognitive constraints of short-memory effect (Chafe 1992). Writers usually have time to plan, edit and revise their clause, sentence and discourse; readers can always stop when they come across comprehension problems, and reread and reprocess at a more leisure pace.
    The second major difference (related to the first) between the English oral and written narratives is the use of existential structure such as the following:

13) There was a little girl [who had a ball].
14) There was a deliveryman [who came to the door].

While writers often introduced a new referent in the story by coding it as a subject NP (e.g., examples 11 and 12), speakers did so by coding it as the subject of an existential clause, as in (13-14) above. The processing of an existential subject head taking a RC has two cognitive advantages over that of a regular subject NP+RC construction. On the one hand, it has the

subject placed after 'there be ...", a typical GIVEN BEFORE NEW structure, which signals to the listener the incoming of certain new information, and on the other hand it has the RC placed at the end of the clause, which does not interrupt the main clause processing. Six out of the 16 SS utterances in the oral data are of the existential type, which occurs only once in the written data. In online comprehension when cognitive demand is high, the use of such a structure to introduce new referents helps reduce processing difficulties.

The third difference between the two modalities is the use of relative pronoun in ORCs. Though optional in ORCs, the relative pronoun frequently occurs in the oral narrative, serving as a signal to inform hearers of the advent of a RC and prepare them for its processing. Writers, on the other hand, typically opted for the omission of the relative pronoun, presumably because of the fact that cognitive constraints in written discourse are relaxed and readers can afford ambiguity and reprocessing to a greater extent than hearers. Examples (15) and (16) are typical of the ORC utterances in the oral and written tasks respectively.

15) He liked the sound [that the doorbell made].   ORAL
16) He enjoyed the tune [the doorbell plays].   WRITTEN

Overall, 8 out of all 10 ORCs (the SO and OO types) in the oral data are led by a relative pronoun, while only 3 out of 9 ORCs in the written data are so preceded.

4.4 Chinese narrative

The most striking difference between Chinese oral and written narratives is the frequency distribution: RCs used in the latter are almost five times more than those in the former regardless of the fact that the oral narrative is in general longer than the written one (an average of 87 versus 76 propositions per narrative respectively). Although NORMAL FORM may account for much greater occurrence of SRCs than ORCs in general, we argue that the skewed frequency distribution between the two modalities manifests the cognitive strategy of CLOSURE at work in Chinese as well. Whereas the strategy operates at a local level of clauses in English, it does so at a more global level in Chinese discourse due to the fact that the majority of Chinese RCs are of the embedding kind. On the one hand, sentences with embedded RCs impose greater demand on memory than the one-clause sentence (Just & Carpenter 1992, Martin 1987, Prideaux & Baker 1986), and on the other hand, Chinese RCs placed before their head also deny a quick access to the head. In oral discourse when speakers and hearers are more severely constrained by cognitive activities of memory and attention than their counterparts in writing, the processing of embedded RCs would only add to cognitive demand. Consequently, Chinese speakers would in general try to avoid RC production in oral tasks so as not to tax cognitive resources.

The processing constraints underlying spoken language affect writing and reading to a much lesser extent due to the slower, editable, and retraceable language medium that allows writers to construe more complex sentences at ease and use relative clauses more frequently. The interaction of cognitive activities and syntactic properties of RCs results in the overall different frequency distribution of RCs between oral and written narratives.

Another major difference is the occurrence of split head NP versus unified ones, the former having the structure of Det+RC+NP (as in example 1) and the latter that of RC+Det+NP (as in example 2). Theoretically, only a clause-initial subject head may take a non-interruptive RC and thus a unified NP would be expected in the SS and SO constructions for the benefit of early CLOSURE, especially in the oral narrative task when cognitive demand is greater. However, split head NPs occur more frequently in the oral than the written narrative: 100% of all SS and SO utterances (7 out of 7) in the oral data feature a split head, and the percentage reduces to 73% (24 out of 33) in the written data. Why is this so? The answer lies again in how cognitive strategies interact with specific linguistic structures of a language. Since Chinese RCs lack a leading relative pronoun or any other clause-initial marker, such a clause with a unified head, e.g., [RC ØsVO *de*] SV(O), sounds/looks just like a common main clause with a zero subject until the particle *de* is reached at the end of the RC. The structure of RC+Det+NP would very likely lead the hearer down a 'garden-path,' thus risking being reprocessed at the end of the RC. A determiner placed before a RC, on the other hand, would signal to hearers the forthcoming head NP and have them prepared for continued processing at the end of the RC. The signaling function of the determiner would be of greater importance in cognitively more constrained oral discourse because hearers are less able to afford ambiguity and reprocessing than readers. Hence, as is the given+given information structure preferred in the English oral narrative, the split head structure is preferred in the Chinese oral narrative, both allowing for processing ease in the cognitively more demanding oral task.

4.5 General dispreference for SO

Of all four RC types, the SO construction occurs only marginally in both languages: Two are found in the Chinese written data and four in the English oral data. For example,

49) And the first person [that he hears] comes upstairs.
50) [ta   xingxingkuku zhuangshang *de*] menling   meiyou   renhe   yongchu.
    he take.pains   install   door.bell have.no any   use
    The doorbell (that) he took pains to install was useless.

The strong dispreference for SO utterances results from the disadvantage that the particular structure suffers with regard to cognitive constraints on the one hand and discourse-pragmatic and semantic properties on the other. The same factors that motivate the occurrence and distribution of SS, OS and OO constructions account for the dispreference of the SO structure as well, simply because they all operate against the formation of SO. First, the cognitive strategies of CLOSURE and NORMAL FORM disfavor the production of SO because the ORC does not have the normal word order and interrupts the main clause processing. Second, the discrepancy between the syntactic coding and the semantic features of the head NP discourages the use of SO, where a subject head usually encodes a human referent but an ORC tends to modify a nonhuman head coded as object (in OO), a notable feature observed in several studies (Chen 1997, Fox & Thompson 1990, Pu 2007). Third, the particular discourse-pragmatic function of object relative clause also contributes to the low frequency of SO utterances because the subject of a main clause, usually human and topical, is more inclined to take a subject relative clause that describes, characterize or reinstate the head rather than an object relative clause that relates it or attributes it to another referent or entity. All factors seem to conspire against the production of SO in discourse, which yield the lowest frequency distribution of all four RC constructions.

6 Conclusion

The present study has shown a general SRC-ORC asymmetry in both Chinese and English narrative discourse, as observed in prior RC research, and discussed further important differences in RC production between oral and written narratives in each language. Both general and specific RC patterns can be accounted for by the proposed cognitive-functional principle that explores underlying cognitive, semantic, and discourse-pragmatic factors operative in RC processing. The overwhelming preference for SRC over ORC shows the effect of NORMAL FORM at work in RC formation in both languages, and the effect of CLOSURE IS also robust in English. Furthermore, the strong preference for SRC to modify a subject head and dispreference for ORC to modify a subject head are due to the harmony (or disharmony) between the head NP and its RC in terms of semantic properties of the head and discourse function of the RC. On the other hand, the differences in RC production between oral and written narratives can readily be explained by different cognitive demand in speech and writing. These remarkably similar patterns across the two modalities and languages have underscored certain universal characteristics of RC formation in discourse.

Although we have yet to explore and discuss relative clause production in other kinds of discourse (e.g., spontaneous conversation, explanatory or expository texts), which may be more complex than what have been found in the present study, we believe that a cognitive-functional approach, such

as the one proposed in our study, to the analysis of syntactic structure and its function is adequate, explanatory and profitable because grammar is not autonomous, but rather shaped by various cognitive, semantic, pragmatic and discourse forces as well as strategies speakers develop to better communicate their meanings and intentions.

References

Andersen, A. H. 1989. Markedness theory-The first 150 years. In: O.M Tomic (Ed.) *Markedness in synchrony and diachrony* pp. 11-46 . New York: Mouton de Gruyter.

Bever, T.G. 1970. The cognitive basis for linguistic structures. In: J. R. Hayes (Ed.) *Cognition and the development of language* pp. 279-362 . New York: John Wiley and Sons.

Biber, D. 1988. Variations *across speech and writing*. Cambridge: Cambridge University Press.

Carreiras, M., Duñabeitia, J.A., Vergara, M., de la Cruz-Pavía, I., & Laka, I. 2009. Subject relative clauses are not universally easier to process: Evidence from Basque. Cognition, 115 (1), 79-92.

Chafe, W. 1992. The flow of ideas in a sample of written language. In: Mann, W.C. and S. A. Thompson (Eds.) *Discourse description: Diverse linguistic analysis of a fundraising text*. Amsterdam: John Benjamims. pp. 268-94.

Chen, S. 1997. *A discourse analysis of relative clauses in spoken Chinese: A study based on grammatical reflexes and information structure*. Unpublished MA thesis. National Taiwan University.

Chu, C. 1998. *A discourse grammar of Mandarin Chinese*. New York: Peter Lang Publishing.

DuBois, J.W. 1980. Beyond definiteness: The trace of identity in discourse. In: W. Chafe (Ed.) *The pear stories*. Norwood, N.J.: Ablex. pp. 302-74.

DuBois, J.W. 1987. The discourse basis of ergativity. *Language, 63*, 805-55.

Fenk-Ozslon, G. 2001. Familiarity, information flow and linguistic form. In: J. Bybee & P. Hopper (Eds.) *Frequency and the emergence of linguistic structure* Amsterdam: John Benjamins. pp. 431-448.

Ford, M. 1983. A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior*, 22, 203–218.

Fox, B. A. & Thompson, S. A. 1990. A discourse explanation of the grammar of relative clauses in English conversation. *Language, 66(2),* 297-316.

Frazier, L. 1979. *On comprehending sentences: Syntactic parsing strategies*. Bloomington: Indiana University Press.

Frazier, L. & Fodor, J. D. 1978. The sausage machine: A new two-stage parsing model. *Cognition*, 6, 291–325.

Gennari, S. P. & MacDonald, M. C. 2008. Semantic indeterminacy in object relative clauses. *Journal of Memory and Language*, 58, 161–187.

Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1–76.

Gibson, E. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In: Y. Miyashita, A. Marantz, & W. O'Neil (Eds.) *Image, language, brain*. Cambridge, MA: MIT Press. pp. 95–126.

Givón, T. 1993. *English Grammar: A Function-based Introduction, Vol. I & II.* Amsterdam: John Benjamins.

Hsiao, F. & Gibson, E. 2003. Processing relative clauses in Chinese. *Cognition*, 90(1), 3-28.

Huang, S. & Chui, K. 1997. Is Chinese a pragmatic order language? In: C. Cheng (Ed.) *Chinese languages and linguistics IV: Typological studies of languages in China.* Taipei, Taiwan. pp. 51-80.

Just, M.A. and Carpenter, P.A. 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99(1),* 122-149.

King, J., & Kutas, M. 1995. Who did what and when? Using word- and cause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7, 376–395.

Lewis, R. 1996. A theory of grammatical but unacceptable embeddings. *Journal of Psycholinguistic Research*, 25, 93–116.

Li, C. N, and Thompson, S. 1981. *Mandarin Chinese: A functional reference grammar.* Berkeley and Los Angeles: University of California Press.

Lynkowsky, P.E. 1980. *Development of relative clauses: Comprehension strategies in English and Ukranian.* Unpublished Ph.D. Thesis, University of Alberta.

MacWhinney, B. & Pleh, C. 1988. The processing of restrictive relative clauses in Hungarian. *Cognition*, 29, 95–141.

Martin, R.C. 1987. Articulatory and phonological deficits in short-term memory and their relation to syntactic processing. *Brain and Language, 32,* 159-192.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. 1998. Modeling the influence of thematic fit and other constraints in on-line sentence comprehension. *Journal of Memory and Language*, 38, 283–312.

Prideaux, G. D. 2000. Discourse analysis and psycholinguistics: Breaking down barriers. Presidential Address, Annual meeting, Linguistic Association of Canada and the United States. In: A. Lommel & A. Melby (Eds.) *LACUS Forum, 26,* 1-16.

Prideaux, G.D. & Baker, W. 1986. *Strategies and structures: The processing of relative clauses. Current Issues in Linguistic Theory 46.* Amsterdam: John Benjamins.

Pu, M-M. 2007. The distribution of relative clauses in Chinese discourse. *Discourse Processes*, 43(1), 25-53.

Tabor, W., Juliano, C., & Tanenhaus, M. K. 1997. Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language & Cognitive Processes*, 12, 211–272.

Tao, H. 2002. Semantic and discourse properties of relative clauses in Chinese oral narratives . *Contemporary Research in Modern Chinese, 4,* 47-57.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. 1994. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33, 285–318.