Information structure constraints and complex NP islands in Chinese

Dawei Jin
University at Buffalo, SUNY

Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar

Freie Universität Berlin
Stefan Müller (Editor)
2013
pages 110–120

Abstract

This paper presents an analysis of the complex NP island effects in Chinese. I follow Ginzburg & Sag (2000)’s analysis of in situ wh-interrogative construction and propose that feature percolation from the non-head clause daughter to the head daughter is required for a proper treatment of in situ wh-relative. A semantic analysis of the idiosyncrasy of weishenme ‘why’ reveals that a definite reading is forced for a wh-relative when weishenme stays in situ. This requirement causes feature percolation into relative head to fail. In this way I show that the island effects in Chinese can be independently ruled out in the grammar as a case of contradiction.

1 Introduction

This paper proposes that the complex NP island constraints (henceforth: CNPC) in Chinese wh-interrogatives receive an information-structural explanation. I argue that mainstream treatments of CNPC in terms of movement constraints fail to predict the interpretational distinctiveness associated with different wh-phrases. On the other hand, island facts follow naturally from an independently motivated constraint on relative clause’s propositional content, motivated by this distinctiveness. I adopt an HPSG implementation used in representing in situ wh-interrogatives (Ginzburg & Sag, 2000). I show this framework allows us to impose fine-grained interactional constraints that capture the relation between wh-phrases and semantic interpretation.

The rest of this paper will be structured as follows: Section 2 reviews the core data surrounding CNPC effects in Chinese; Section 3 summarizes previous theories on Chinese strong islands, couched in the transformational framework, and discusses their shortcomings; Section 4 examines the behavioral differences between reason adverbial weishenme ‘why’ and other wh-phrases in some detail, and derives the CNPC effects by an information-structural constraint based on this distinction; Section 5 presents an HPSG implementation of the above mentioned analysis; Section 6 concludes the paper.

2 Data

It has long been noted (Huang 1982; Aoun & Li, 1993; Tsai, 1994; Huang et al, 2009) that in situ wh-phrases in Chinese can circumvent the canonical CNPC effects, where proper interpretation cannot be established when a wh-phrase is associated with a relative-clause internal position. As (1) shows, when a wh-phrase is overtly fronted, both Chinese and English induce island effects; however, such effects disappear when Chinese wh-phrases stay in situ.

(1)  a. ?? What do you like the person [who wrote _ ]?  
     b. ?? Shenme, ni xihuan [ xie _ ] de ren?  
     What, you like write REL person  
     c. Ni xihuan [shei xie _ ]  de    shu?  
     You like who write REL book

1
d. Ni xihuan [ _ xie shenme] de shu?
   ‘What topic, do you like book(s) that describes _?’

On the other hand, it has been claimed since Huang (1982) that CNPC obtains for in situ reason-adverbial, weishenme ‘why’, illustrated below.

(2) #Ni xihuan [ta weishenme xie _] de shu?
   ‘Why, do you like the books that he wrote _?’

Crucially, this contrast has been argued to be a matter of argument-adjunct distinction (Huang, 1982), given examples like the following, where island effects once again disappear when weishenme is replaced by an argumental reason wh-phrase, yinweishenme ‘because of what (reason)’.

(3) Ni xihuan [ta yinwei shennme xie _] de shu?
   ‘What reason do you like the book(s) that he wrote for that reason?’

3 Previous Analyses

The mainstream explanations of weishenme-induced CNPC (Huang et al, 2009; Cheng & Rooyrck, 2000; Cheng, 2009) have been to take the unruly behavior of weishenme as crucial evidence for the existence of covert movement. Specifically, these theories argue that for a wh-interrogative to receive proper interpretation in Chinese, the interrogative feature at the matrix scope position needs to be checked off at LF, the purported level of representation that provides the feed for semantic interpretation. One way to achieve feature checking is to move the wh-phrase to the matrix position at LF. However, the complex NP domain, which subsumes relative clause, constitutes a barrier against movement, inducing island effects. This explains why weishenme induces CNPC. To explain away the island-free behaviors of other wh-phrases, a separate, movement-free licensing mechanism for wh-interpretation, unselective binding (Pesetsky, 1987; Reinhart, 1998; Aoun & Li, 1993; Tsai, 1994), is proposed, which selectively targets wh-arguments.

However, this line of reasoning faces several difficulties.

Theoretically, a movement-based explanation should predict that the island effects disappear in overt pied piping, since it involves extraction of the entire complex NP domain, and therefore the in-situ wh-phrase should not cross any barrier. Overt pied-piping of the whole NP chunk to topicalized position is commonly attested in Chinese filler-gap constructions, as (4) illustrates.

(4) a. [Shei xie _] de shu, ni xihuan?
   ‘Books written by who, do you like?’

b. [ _ xie shenme] de shu, ni xihuan?
Write what REL book, you like
‘Books which describe what, do you like?’

However, overt pied-piping fails to rescue *weishenme*-islands, as the following shows

(5) #[Ta weishenme xie _] de shu, ni xihuan?
    He why write REL book, you like
    ‘Books which he wrote why, do you like?’

Under a movement-based theory, this fact seems mysterious because there seems to be no non-stipulatory reasons why pied-piping should be ruled out as an option in (5).

Empirically, other adjuncts or adverbials are also island-free, as (6) shows.

(6) a. Ni hui mai [ _ mai duoshaoqian] de shu?
    You will buy sell how.much REL book
    ‘How much will you buy the book(s) that were sold for that amount of money?’

b. Ni xuyao [ na’er neng maidao _] de shu?
    You need where can buy=RES REL book
    (REL: resultative)
    ‘Where do you need the books that can be bought at that place?’

Therefore, the purported argument-adjunct asymmetry, motivated by the contrast between (2) and (3), is only apparent. The actual contrast w.r.t. CNPC effects involves *weishenme* versus all other *wh*-phrases. It seems hardly desirable that a structural mechanism is formulated upon one data point alone and is forced to rule out all the remainder.

Furthermore, structural theories fail to take account of the fact that the acceptability for *wh*-phrases in a relative is interpretation-dependent. Crucially, I argue that only generic readings are available for the aforementioned island-free examples. Because there is no definite determiner in Chinese, whether a relative head receives generic or definite readings is normally resolved by contexts. However, as (7) exemplifies, when a definite reading is forced via the presence of the demonstrative *nei* ‘that’, CNPC effects arise even for *in situ wh*-arguments.

(7) a.#Ni xihuan [shei xie _] de nei-ben shu?
    You like who write REL DEM-CL book
    #‘Who, do you like that book that _ write?’

b #Ni xihuan [ta yinwei shenme xie _] de nei-ben shu?
    You like he because.of what write REL DEM-CL book
    #‘What, do you like that book that he wrote because of _?’

Conceivably, a structural theory may argue that definiteness markers can be barriers of movement. Indeed, Huang (1982) proposes exactly this kind of explanation for the following English example.
However, the same explanation cannot be extended to the anti-definiteness effect in Chinese, because in such cases no movement occurs: *wh*-arguments undergo unselective binding, and binding, according to standard structural assumptions, is not sensitive to movement barriers (Cheng & Rooyrk, 2000; Cheng, 2009).

These suggest that we should look for the explanation for weishenme-induced CNPC within the interpretational component of grammar. Below I propose such an analysis.

### 4 My Analysis

The semantics of *why* has long been noticed to be peculiar crosslinguistically (Bromberger, 1992; Szabolcsi & Zwarts, 1997). Recent literature has presented various treatments for *why*’s idiosyncrasy, e.g. *why* favors high attachment or late insertion (Ko, 2005). I will adopt Tsai (2008)’s proposal that Chinese *weishenme* takes the underlying event as its internal argument and functions as a sentential operator; On the other hand, the argumental reason *wh*-phrase *yinweishenme* ‘because of what’ modifies the underlying predicate and functions as a derived predicate in the manner of VP-adverbials. I argue this formulation can readily account for the differing interpretations elicited by the two *wh*-phrases.

For example, although the semantic distinction of *weishenme/yinwei shenme* does not yield logically distinct interpretations when a single event is under discussion, different interpretations arise when a multiple event reading is elicited through the introduction of a quantifier.

(9) a. Lisi yinwei shenme cizhi?
   *Lisi because.of what resign*
   ‘What reason does Lisi have for resigning?’

b. Lisi weishenme cizhi?
   *Lisi why       resign*
   ‘Why did Lisi resign?’

c. Weishenme duoshu ren cizhi?
   *Why       most person resign*
   ‘Why most people, not few people, resign? (What is the singular reason that causes most people in the salient discourse to resign?)’

d. Duoshu ren  yinwei shenme cizhi?
   *Most person because.of what resign*
   ‘What reason did most people have for resigning? (What reasons can account for the majority cases of resignations?)’

e. Weishenme meiyouren/henshaoren cizhi?
   *Why       nobody/few person resign*
   ‘Why nobody/few people resigned?’

f. Meiyouren/henshaoren yinwei shenme cizhi?
   *Nobody/few people because.of what resign*
   ‘What reasons did nobody/few people have for resigning?’
Generalized quantifiers like *most/few* need to quantify over properties/predicates. Therefore they can only take *yinweishenme* as argument, whereas *weishenme* ranges above the entire quantified event as its argument. Thus when we adopt a strictly compositional semantic derivation by interacting reason *wh*-phrases with other constituents, the resulting logical interpretations will differ. This, I argue, underlies the purported island effects in Chinese.

To begin with, interpretation of a clause containing a *wh*-phrase requires the propagation of the interrogative feature (Fiengo et al, 1988; von Stechow, 1996). For example, when an interrogative NP is contained within another NP, as in *pictures of who*, the head NP will also be construed as interrogative. Just as *who* is a quantifier ranging over individuals, *pictures of who* may be construed as a quantifier ranging over pictures sets defined by their owners. This feature percolation idea has been implemented using different semantic frameworks, but the basic intuition remains the same.

In a *wh*-relative, *wh*-feature percolation requires the head noun to denote a set of entities defined in terms of the properties specified in the *wh*-phrase. For example, in (10)

(10) Ni xihuan [shei xie _] de shu?
    *You like  who write  REL book*
    ‘Who is the person s.t. you like book(s) that (s)he wrote?’

The embedded *wh*-argument *shei* ‘*who*’ denotes a salient set of individuals who have written books, and the question ranges over any books that bear the property of being written by this set of individuals.

A definite reading, where the set of books are already salient from context, and we are inquiring after its author, is not available. That is to say, the identity of the books cannot be known *a priori*, but has to crucially rely on anchoring the identity of the individual who writes them.

Similarly, in (11), a set of alternative sets of books are characterized in terms of a set of discourse-salient reasons as follows

(11) Ni xihuan [ta yinwei shennme xie _] de shu?
    *You like  he because.of what write REL book*
    ‘What reason do you like the book(s) that he wrote for that reason?’

For example, imagine we have a context where a book A was written for reason R₁, a book B was written for reason R₂, etc. The *wh*-relative in (11) would pick out the set of books {A, B,…}, which are defined in terms of the set of reasons {R₁, R₂,…}. Crucially, the *wh*-feature must percolate from the clause-internal *wh*-phrase to the head, so that the identity of the head noun is determined by the property specified within the *wh*-phrase.

On the other hand, *weishenme* cannot lend itself to such an interpretation, because a *weishenme*-question necessarily solicits the cause of a particular event which is denoted by the propositional argument that *weishenme* takes. Therefore if in (11) *yinweishenme* is replaced by *weishenme*, the relative clause will derive a class of propositions as follows:
Given that \textit{weishenme} is anchored to a particular event, it cannot provide a classification base that derives multiple events. In fact, it is only logically coherent with a specific reading, i.e. there exists a reason that causes his writing a particular book that the addressee likes, and we are wondering what this reason is. Therefore, the discourse referent of the head noun is not anchored by the relative clause, it must be known \textit{a priori}, and by locating this referent, we are retrieving the reason for this particular event.

This results in a paradox, because the propositional content within the relative clause plays no role in identifying the head noun, therefore feature propagation is impossible.

The anti-definiteness effect follows from the same reason: a demonstrative indicates that the discourse referent of the head noun is an \textit{a priori} known entity salient from prior discourse. However, since the relative clause's propositional content is interrogative, it cannot serve to anchor this referent, therefore similar contradiction occurs.

5 Formalization

Below I present an HPSG formalization based on Ginzburg & Sag's (2000) analysis of \textit{wh}-interrogatives. I show HPSG mechanisms neatly account for this island theory in terms of feature constraints, without incurring the difficulties encountered by a movement-based theory.

The Chinese \textit{in situ} \textit{wh}-relative construction can be treated as being subject to the constraints associated with matrix \textit{in situ} interrogative clauses (\textit{is-int-cl}) and embedded relative clauses (\textit{rel-cl}). Importantly, \textit{is-int-cl} possesses several peculiar properties compared to fronted \textit{int-cl}. First, it must allow the non-initial \textit{wh}-word in its \textit{in situ} position to bear a specified WH-value (Ginzburg & Sag, 2000). Second, association of this WH-value at the root clause level needs to be guaranteed, in order for the matrix content type to be \textit{question}. This requires the WH-value to percolate up via head, given there is no initial filler.

The following constraint by Ginzburg & Sag (2000) already allows WH-percolation to occur via the head-argument path:

\begin{equation}
\text{(13) WH-Amalgamation Constraint}
\end{equation}

\[
\text{word} 
\Rightarrow \left[ \begin{array}{l}
\text{SS[WH ARG ST } \langle \text{WH [S1]}, \ldots, \text{WH [Sn]} \rangle
\end{array} \right]
\]

As is discussed in my previous analysis, a crucial step to guarantee a question reading is to allow the WH-value to be shared with the relative head, so that it can percolate up all the way through the matrix clause head. This is formulated in a separate WH-amalgamation constraint:

\begin{equation}
\text{(14) Rel WH-Amalgamation Constraint}
\end{equation}
Finally, NPs with a definite reading will have a $[\text{DEF}+]$ feature, whereas generic NPs receive a $[\text{DEF}–]$ feature. This feature is specified in the CONTENT since it is a semantic/pragmatic feature. Thus a demonstrative head NP is marked $[\text{DEF}+]$, so is the head NP of a *weishenme*-relative clause.

On the other hand, whenever a WH-value is specified for a word, the word must be indefinite. This is because a WH-word inherently ranges over a set of properties, and thus cannot ground a particular discourse referent. I incorporate this requirement by stipulating it as the following constraint:

\[(15) \text{A noun with a non-empty WH value must be } [\text{DEF}–]\text{ in its CONTENT.}\]

|
\[
\text{noun } \Rightarrow \begin{cases} 
\text{SS}[\text{WH} & \begin{cases} \text{ne} \\
\text{CONT} & [\text{DEF}–] \end{cases} \\
\text{WH} & \text{STORE} \begin{cases} () \end{cases} \\
\text{VP} & \begin{cases} \text{WH} & \text{STORE} \begin{cases} (1) \end{cases} \\
\text{WH} & \text{STORE} \begin{cases} (2) \end{cases} \\
\text{rel} - \text{cl} & \begin{cases} \text{SS}[\text{WH} & \begin{cases} (2) \end{cases} \\
\text{MOD} & (1) \\
\text{SLASH} & \text{CONT} \begin{cases} (\text{SPECIFIC} & \begin{cases} – \end{cases} \end{cases} \end{cases} \\
\text{PARAMS} & \begin{cases} (\Omega) \end{cases} \\
\text{question} & \begin{cases} () \end{cases} \\
\text{CONT} & \begin{cases} () \end{cases} \\
\end{cases} \end{cases}
\end{cases}
\]

For an *wh*-argumental *is-int-cl* like (11), the WH value can percolate all the way up until getting associated at the root clause. First, through (13) it percolates to the embedded clause level, and then to the relative head via (14). Then (13) applies again, until it reaches the matrix root clause. Importantly, the relative head NP is marked $[\text{DEF}–]$ when it inherits the WH-value. Also, the semantics of the relative clause results in a generic reading for the head NP, which also requires a $[\text{DEF}–]$ value. These two requirements give compatible results, and the whole derivation can be implemented as follows:

\[(16)\]

For a *weishenme* *is-int-cl* as in (2) (repeated below)

\[(17)\]  
\[\begin{align*}
\text{#Ni xihuan } [\text{ta weishenme xie} ] \text{ de } & \text{ shu?} \\
You & \text{ like } \text{ he why write REL book} \\
\text{# 'Why, do you like the books that he wrote __?'}
\end{align*}\]

The WH-value of the relative clause percolates up to the head NP and requires it to be specified as $[\text{DEF}–]$. However, the semantics of the relative clause imposes
a definite reading for the head NP, marking it [DEF+]. These two competing feature valuations create contradiction, and as a result the type weishenme is-int-cl is ruled out by the interacting constraints and doesn’t get generated. The representation of (17) is as follows:

Similarly, when the type wh-argumental is-int-cl interacts with the type definite-NP, contradiction also arises. This explains the anti-definiteness effect. As the following shows, the presence of a demonstrative determiner marks the head NP as a [DEF+] NP, however, percolation of WH-value from the relative clause daughter requires the head NP to bear a [DEF-] value. The competing requirements cannot be accommodated, and therefore this structure is also ruled out.

6 Conclusion

This paper argues that the CNPC in Chinese are explained by semantico-pragmatic mechanisms. Structural explanations fail to address the fact that the idiosyncratic semantics of why gives rise to different interpretations compared to other wh-phrases. I show that once this distinction is made, the interpretation differences underlie the judgment contrasts w.r.t. island effects. This solution is readily accommodated within a constraint-based HPSG framework.

This proposal suggests a simpler grammar, since there is no need to specify structural constraints on in situ wh-questions. Also, the removal of structural stipulations renders void the grounds for positing covert movement at the LF level. Thus wh-licensing mechanisms need not be sensitive to the syntactic categories of wh-phrases.

One consequence of this is on the evaluation of island theories. If we assume that in situ island effects are treated on a par with overtly displaced island effects, island theories that are formulated on overt displacement cannot be extended to in situ cases, thus suffering from an empirical disadvantage (Lasnik, 1999; Sprouse et al, 2012; Boeckx, 2012). For example, in processing-based theories, overt displacement is crucial because the dependency it creates imposes taxing
burden on the processing resources of a cognitive agent (Deane, 1991; Kluender, 1998; Hofmeister & Sag, 2010). As such this approach doesn't have an explanation for Chinese island effects. However, according to my theory, it is not necessary to make this extension. Therefore, although the current claim doesn’t in principle favor nonstructural theories over structural ones, it enables a level playing ground by rendering the evidence from in situ islands irrelevant.

Notes

1 Other strong island constraints can be similarly obviated, for example, adjunct islands and subject islands, illustrated as follows

(i) a. Ta [yinweishejme jiegu yuangong] yihou bei laoban piping=le?  
   ‘For what reason, was he criticized by the boss after he sacked
b. [Ta yinweishejme eizhi] zui hao?  
   ‘For what reason, will that he resigned _ be the best?’

These phenomena can follow from the analysis laid out in this paper. However, I will leave their exact formulation to future work.

2 The same asymmetrical pattern holds also for Japanese and Korean.

3 One can possibly accept this sentence in a reprise reading: where the *wh*-word serves as an anaphora that refers to a previously pronounced linguistic entity in prior discourse.

References