Modelling the *ziji* Blocking Effect and Constraining Bound Variable Derivations in MC-TAG with Delayed Locality

Dennis Ryan Storoshenko
University of Calgary
Craigie Hall C205, 2500 University Dr NW
Calgary, AB, Canada
dstorosh@ucalgary.ca

Abstract

MC-TAG derivations using delayed locality (i.e. non set-local composition) have been used to model long distance binding in natural language. In this paper, I explore the possibility within Synchronous TAG that tree sets composing via delayed locality may have many possible syntactic derivations for a single semantic form. Using the Mandarin (Chinese) anaphor *ziji* as a test case, I show that the blocking effect which is described in other generative frameworks as a constraint on covert movement can be modelled as arising from constraints across possible derivation families for a given tree set. Further observing the interaction of *ziji* with other bound variables in the language, I propose that anti-locality effects in variable binding arise when a more specialized bound variable is available. This has consequences for the definition of a typology of bound variables.

1 Variable Binding in STAG

Variable binding in Synchronous TAG (STAG) has been modelled as a phenomenon which exploits delayed locality. Building on the definition of delayed-local derivations presented in Chiang and Scheffler (2008), bound variables can be modelled as multi-component sets (MCSs) in both the syntax and semantics. As formulated in Storoshenko and Han (2015), Figure 2 shows that the bound variable obligatorily composes non-locally, with the α components substituting into the argument position of the predicate selecting the variable, while the β components compose directly with the antecedent. This is most crucial on the semantics side, where the variable's auxiliary tree composes with the "scope part" of a STAG generalized quantifier, as in Figure 3. The bound variable carries a function from $\langle$e,t$\rangle$ to $\langle$e,t$\rangle$ which accomplishes the variable binding, in a form similar to Büring's (2005) Binder Index Evaluation Rule. On the syntax side, the β component acts as solely an agreement check. The English bound variable pronoun *her* is lexically specified for third person feminine φ-features; in this case there is no gender specification on *a student*, so the combination goes through. For nominals with inherent gender specification, for example *every boy*, the resulting feature clash would block the derivation. In this way, agreement is treated as a purely syntactic phenomenon, with no manifestation in the semantics. As written, these trees could compose into a sentence such as *A student loves her father*.

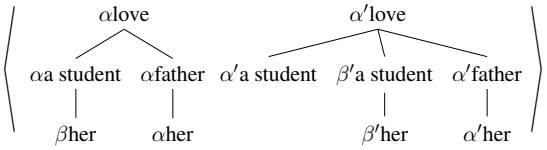

Figure 1: Derivation trees for *A student loves her father*

Following the derivation trees in Figure 1, delayed locality is observed in the composition of the members of the *her* MCS, while the only other MCS present has composed tree-locally[1]. Chiang and Scheffler define a delay as the set of derivation tree nodes including the members of the MCS, and all nodes on a path between members of the MCS, but excluding the common dominating node. Storoshenko and Han extend this notion into a measure for the size of a delay by taking the cardinality of that set. The derivations in Figure 1 will have a delay (d) value of four for the pronoun.

[1]Though the DP anchored by *father* could also be represented as a MCS.

Proceedings of the 12th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+12), pages 67–76,
Düsseldorf, Germany, June 29 - July 1, 2016.

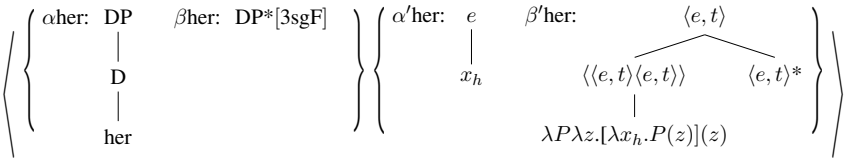

Figure 2: Tree sets for bound variable pronoun *her*

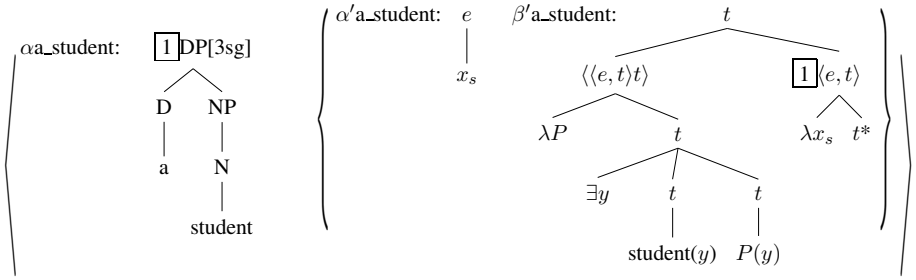

Figure 3: Sample Quantifier Tree Set

Beyond referring to the delay as a whole, they further define the length l of a delayed local MCS combination as $d - n$, where n is the number of MCS members. Thus, for the bound variable pronoun in the derivation trees of Figure 1, l=2. The reason for applying this characterization of a delay becomes clear when considering the quantifier tree sets. As defined, even a tree-local MCS combination will have a trivial delay consisting only of its own derivation tree nodes, yielding a d value of two, though the l value for this MCS combination will be zero. This is a more intuitive match with a singleton tree combination on the syntax side, as it is not immediately obvious at a glance that the two derivations are in fact isomorphic. With derivational isomorphism defined based on the equivalency of l values for any given lexical item's combinations on both sides of the STAG derivation, as well as respecting all linked nodes, the quantifier's combination can be said to be maximally isomorphic (borrowing the term from Storoshenko and Han (2015)), as l=0 in both cases. It is this notion of "maximal" isomorphism that I will assume to underlie all further claims of isomorphism.

2 *Ziji* and The Blocking Effect

The key data modelled in this paper all derive from the behaviour of the Mandarin (Chinese) anaphor *ziji*, described extensively in the literature since Cole et al. (1990). In brief, there are five different phenomena which need to be accounted for. The first is the fact that *ziji* can allow both local and long-distance antecedents, as shown in (1).

(1) Zhangsan$_i$ renwei [Lisi$_j$ zhidao [Wangwu$_k$
Zhangsan think Lisi know Wangwu
xihuan ziji$_{i/j/k}$.]]
like self
'Zhangsan thinks that Lisi knows that Wangwu likes self.'

In this respect, *ziji* is not dissimilar to its Korean cognate *caki*, which is discussed in Storoshenko and Han (2015). Where the two diverge is in the manifestation of the so-called blocking effect illustrated in (2):

(2) Zhangsan$_i$ renwei [**wo**$_j$ zhidao [Wangwu$_k$
Zhangsan think **I** know Wangwu
xihuan ziji$_{*i/*j/k}$.]]
like self
'Zhangsan thinks that I know that Wangwu likes self.'

In this sentence where the subject of the middle clause is a first person pronoun, *ziji* is restricted not only to a third person antecedent, but only to the most local one. Descriptively, the presence of an intervening antecedent with mismatched φ-features blocks the anaphor from binding by the highest subject, even though that subject is also third person. However, the situation is different in the local domain:

(3) Zhangsan$_i$ yiwei [**wo**$_j$ hui ba **ni**$_k$ dai
Zhangsan thought I will BA you take
hui ziji$_{*i/j/k}$ de jia.]
back self's DE home
'Zhangsan thought I will take you back to self's house.'

As seen in (3), *ziji* is not inherently restricted to third person antecedents (unlike *caki* and the English *her* from above). In the local domain, *ziji* may be bound by either the first or the second person antecedent. This is a more general property of *ziji*, which can be bound by an antecedent of any φ-feature value, within the constraints of the blocking effect. For example, a version of (1) where all clauses had first person subjects would equally be grammatical. Finally, there is the observation that even when the blocking condition is respected, *ziji* may not take a non-local non-subject antecedent:

(4) Zhangsan$_i$ gaosu Lisi$_j$ [Wangwu$_k$
Zhangsan tell Lisi Wangwu
piping-le ziji$_{i/*j/k}$.]
criticize-ASP self
'Zhangsan told Lisi that Wangwu criticized self.'

Again unlike Korean, the anaphor may not take Lisi as a potential antecedent. However, even though *Lisi* is not considered a possible antecedent, a first person pronoun in that position induces a blocking effect, even though this was not observed in the local domain. Based on these data, the five different properties of *ziji* which need to be modelled are enumerated again below:

1. Local and long distance antecedents are possible
2. φ-feature mismatch in a c-command chain induces blocking across clauses
3. Any antecedent may be viable in the lowest clause
4. The anaphor can in principle take any φ-feature valued antecedent
5. Higher clause antecedents must be subjects

Sohng's (2004) analysis of *ziji* follows the tradition of the original Cole et al. analysis in making use of covert movement to capture all of the relevant facts, accounting for the various differences between Korean *caki* and Mandarin *ziji* as being the result of the difference in φ-features between the two. While the formal mechanisms here will of course be different, STAG not having any notion of covert movement to fall back on, the final analysis will draw on Sohng's key insight. The intricate pattern of data observed will be here be modelled as an interplay between the mechanisms used to capture agreement and a set of well-formedness constraints on STAG derivations.

3 Modelling Core Cases

Above, I showed the agreement checking mechanism that has been previously proposed for bound variables in STAG: a feature value on the degenerate DP node which combines in the syntax with the variable's antecedent. However, this is not the only agreement-checking mechanism which has been proposed in the TAG literature. Working in LTAG, Kallmeyer and Romero (2007) provide an account for reflexive pronouns where agreement is checked via a degenerate T′ node, part of the MCS for the reflexive. The intuition here is that by having agreement check with the functional projection of the verb, a tree-local derivation for the reflexive is ensured. My proposal is simple: the unique φ-feature deficiency of *ziji* versus *caki* is derived by adding this additional degenerate node to the syntax side of the bound variable's MCS, as in Figure 4.

Based on these MCS definitions, it is clear where the notion of delay length will become relevant, as there are asymmetric sets on the syntax and semantic sides. Ultimately, the semantics will define the minimum l value for any given derivation, as the delay needed for semantic composition will be directly determined by the intended meaning of the sentence: α'ziji will substitute at the appropriate argument position while β'ziji will adjoin to the scope part of the desired antecedent[2] Between the two MCSs in Figure 4, I assume the following symmetric compositions: αziji and α'ziji will need to substitute at linked nodes, while βziji$_{DP*}$ and β'ziji will also be required to adjoin into the tree sets of the antecedent. With no matched component to force a fully isomorphic derivation via a linked node, βziji$_{T*}$ is essentially free to attach at any node with a matching label, constrained only by the φ-feature value. I will begin with the assumption that both adjoining sites must have no incompatible features, with the α variables valued via unification with features at the adjoining sites. However, to capture all the facts within a single account, it will emerge that the equivalence of the two feature unifications is relaxed under certain structural relationships between the two agreeing members of the *ziji* MCS. Moving forward, I will not be presenting any information on the semantic derivations, choosing to focus exclusively on the syntax. Full STAG

[2]Proper names are assumed to also have generalized quantifier scope parts to provide the necessary binding.

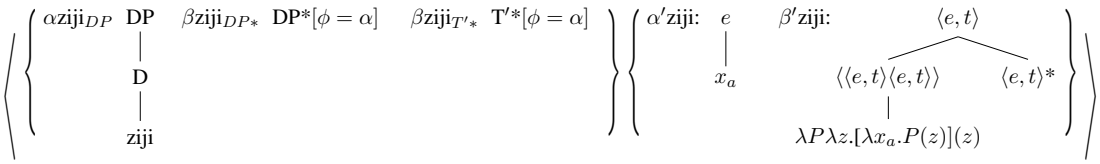

Figure 4: Tree Sets for *ziji*

derivations in the syntax and semantics for many of the predicates used here can be found in existing work, and no changes in the mechanics of deriving variable binding are being proposed in this paper.

For the sake of illustration, one syntactic predicate tree appears in Figure 5. Predicates such as *renwei* "think" and *zhidao* "know" will be simpler transitive verbs, still recursive on C′, with all of these able to either recursively join to each other, or a CP-rooted embedded clause such as *xihuan* "like" or *piping* "criticize" having a DP terminal node at the object position which will be the substitution site for αziji. The ditransitive *gaosu* "tell" is represented using a standard VP-shell structure. The only crucial detail to note is that in a ditransitive predicate, the T′ node is c-commanded by the subject, but not the indirect object.

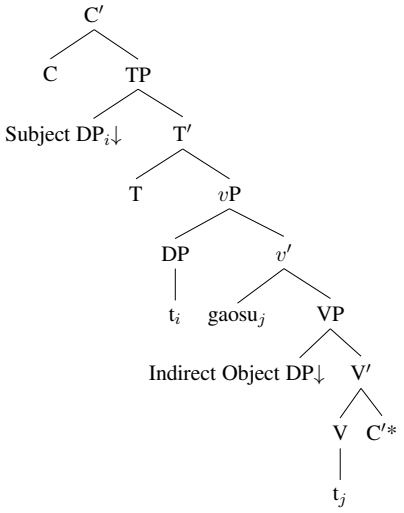

Figure 5: Sample Embedding Predicate *gaosu* "tell"

The three readings for (1) will emerge from the DPs into which βziji$_{DP*}$ and its associated semantic tree adjoin. Multiplying this out by the three clauses each with their own T′ nodes yields nine possible derivation trees, three for each reading. These are given in full in Figure 6, though all further examples will be summarized as in Table 1. In these tables, the l value for the semantic combination of *ziji* will correspond to the cell matching the predicate of which *ziji* is an argument with the intended antecedent for that reading; the leftmost column in all cases.

Table 1: Tabular representation of Figure 6

T′* at:	xihuan	zhidao	renwei
DP* at: Wangwu	δ1: l=1	δ2: l=2	δ3: l=3
Lisi/Wo	δ4: l=2	δ5: l=2	δ6: l=3
Zhangsan	δ7: l=3	δ8: l=3	δ9: l=3

Based on this table, it is immediately clear that not all possible derivations are licit. For the situation where *Wangwu* is the antecedent, placing the βziji$_{T'*}$ in any clause higher than the one containing αziji will create a longer delay in the syntax than the semantics, making δ2 and δ3 not (maximally) isomorphic STAG derivations. Conversely, when *Zhangsan* is the antecedent, all placements of βziji$_{T'}$ yield a formally isomorphic derivation, under the terms defined above. I label this set of isomorphic derivations as a derivational family: a set of isomorphic derivations (as defined by sharing equal l values for all MCS compositions) which yield identical semantic outputs. Removing spurious derivations yields the final set of possible derivations for (1) in Table 2. The lowest antecedent is derived via a singleton derivation, while the other two are represented with derivational families.

Table 2: Isomorphic Derivations for (1)

T′* at:	xihuan	zhidao	renwei
DP* at: Wangwu	δ1: l=1		
Lisi	δ4: l=2	δ5: l=2	
Zhangsan	δ7: l=3	δ8: l=3	δ9: l=3

Taking these same values for the blocking effect example in (2) yields the derivations in Table 3. However, two of the isomorphic derivations, δ4

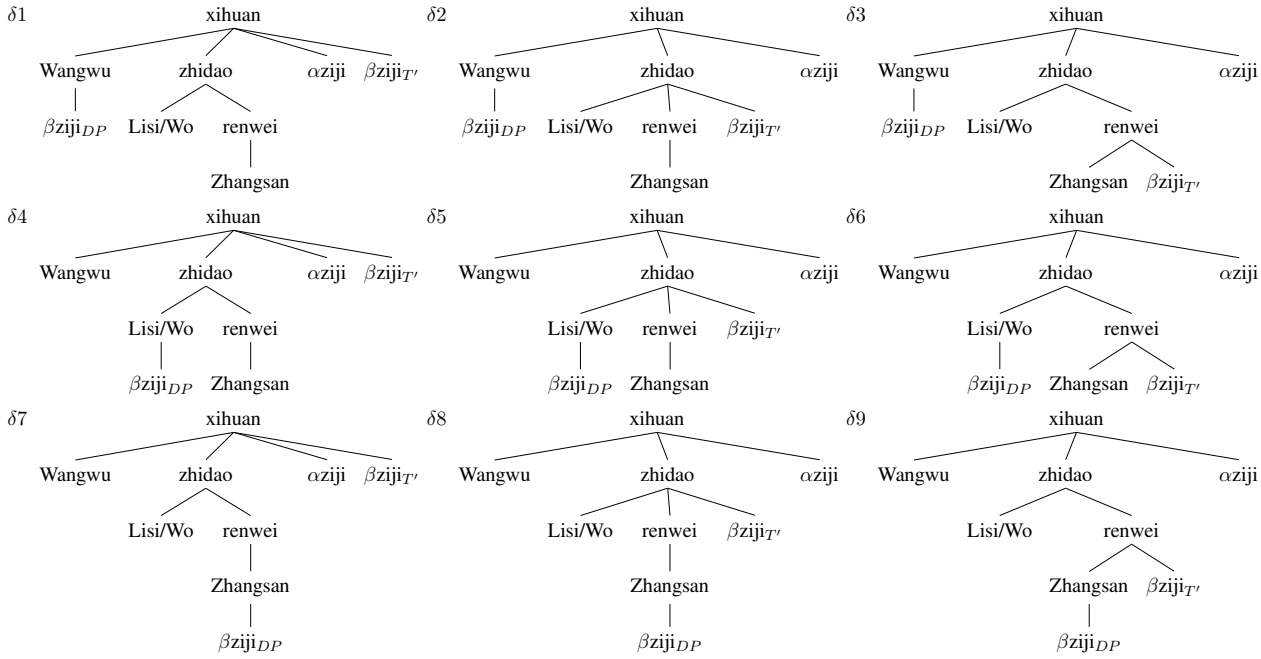

Figure 6: Derivation Trees for (1) and (2)

and δ8 (indicated in bold in the table), will result in a feature clash as the βziji components will adjoin at locations with different φ feature values.

Table 3: Isomorphic Derivations for (2)

T′* at:	xihuan	zhidao	renwei
DP*: at Wangwu	δ1: l=1		
Wo	δ**4:** l=2	δ5: l=2	
Zhangsan	δ7: l=3	δ**8:** l=3	δ9: l=3

This clash is what I propose to be at the root of the blocking effect. Following on the definition of a derivational family, I propose the Fully Functional Family (FFF) constraint:

(5) FULLY FUNCTIONAL FAMILY (FFF): If one member of a derivational family violates a syntactic constraint, all members of the family are rendered infelicitous.

Applying this constraint to the set of derivations in Table 3, only the derivation belonging to the most local antecedent survives.

(3) presents a more complex situation. Here, the lowest clause is a ditransitive, while *ziji* is further embedded inside of a possessive DP structure. The partial derivation for this this example is presented in Figure 7. With three possible antecedents and

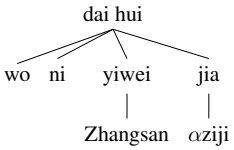

Figure 7: Partial Derivation Tree for (3)

two clauses, here there are six possible derivations, presented in Table 4. Non-isomorphic derivations have already been removed, and those creating a feature clash are again bolded.

Table 4: Isomorphic Derivations for (3)

T′* at:	dai hui	yiwei
DP* at: Ni	δ**1:** l=2	
Wo	δ3: l=2	
Zhangsan	δ**5:** l=3	δ6: l=3

Recall that the possible antecedents here were *ni* and *wo*. The correct prediction is made for *Zhangsan*, whose derivational family is eliminated by virtue of the FFF, and for *wo*, where the first person subject values T′ in the only isomorphic derivation. Note that a hypothetical δ4 would exist, where βziji$_{T'*}$ attaches to the higher clause with a third person subject. However, that derivation is not isomorphic, and does not survive into the (singleton) family, avoiding the feature

clash. A similar $\delta 2$ would exist, but is also non-isomorphic. The unexpected result here is that $\delta 1$ yields a feature clash, and yet the reading is grammatical.

Refining the definition of how a feature clash results resolves this issue. The distinction to make here is in the relationship of the elements in the final derived syntax tree; I assume the higher clause to be a transitive verb auxiliary tree recursive on C′, while the embedded clause is a CP-rooted ditransitive-like structure. In $\delta 5$, βziji$_{DP*}$ is adjoined to the subject in the higher clause, at a position c-commanding βziji$_{T'*}$ adjoined to the T′ node of the lower clause. This c-command relationship, I claim, is required for the feature clash to be triggered. Looking at $\delta 1$, an opposite relationship occurs, where βziji$_{T'*}$ dominates βziji$_{DP*}$. Here, the features do not enter into the necessary configuration to be checked, and $\delta 1$ is not actually a clash. To fully implement this would require a revision to the original *ziji* trees such that it is not a formal part of the MCS that the two feature unifications must match: βziji$_{DP*}$ would receive a value for its α, while βziji$_{T'*}$ would have a distinct value β for its φ features, and the constraint described here would be that if βziji$_{DP*}$ c-commands βziji$_{T'*}$, then $\alpha = \beta$.

This c-command relationship comes into play when it comes to discussing (4) and the closely-related (6):

(6) Zhangsan$_i$ gaosu wo$_j$ Wangwu$_k$
Zhangsan tell I Wangwu
piping-le ziji$_{*i/*j/k}$.
criticize-ASP self
'Zhangsan told me that Wangwu criticized self.'

Recall that (4) indicated that the non-subject is not a possible antecedent, but having *ziji* bound by *Zhangsan* was still possible. (6) has the same structure, but replaces the indirect object with a first person pronoun. Note that even though it is not considered a viable antecedent, it still induces a blocking effect. The partial derivation tree for (4) and (6) is as in Figure 8. Again, with three potential antecedents, and two clauses, there are six possible derivations to consider. Table 5 presents only the isomorphic families for (6).

In this case, the lowest antecedent is correctly predicted to be fine, and the first person indirect object is ruled out because in $\delta 3$ βziji$_{DP*}$ will c-command βziji$_{T'*}$ with mismatched φ-features.

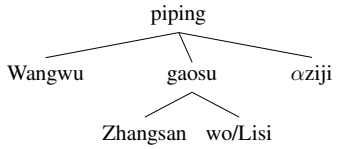

Figure 8: Partial Derivation Tree for (4) and (6)

Table 5: Isomorphic Derivations for (6)

T′* at:	piping	gaosu
DP* at: Wangwu	$\delta 1$: *l*=2	
Wo	δ**3:** *l*=2	$\delta 4$: *l*=2
Zhangsan	$\delta 5$: *l*=2	$\delta 6$: *l*=2

With one derivation ruled out, FFF will block all derivations in the same family. What is not immediately obvious is how $\delta 5$ and $\delta 6$ are also ruled out. To fully lay out the issue, I also present the derivations for (4) in Table 6. In this case, the data conform with the prediction for *Zhangsan*, who is a viable antecedent, but there is no obvious reason that *Lisi* is ruled out. Given that $\delta 3$ is well-formed and satisfies the c-command requirement for feature checking, $\delta 4$ should be examined for the cause of the unacceptability. The intuition here is that as (4) most clearly shows the subject orientation of *ziji*, that effect should fall out from the one derivation where βziji$_{T'*}$ ends up in a position to dominate βziji$_{DP*}$. However, fully ruling out this relationship would block exactly the configuration in the unexpectedly grammatical $\delta 1$ from (3). So, just the right constraint which predicts the grammaticality of (3) and (4) with the reported readings, and possibly offer some insight into the so-far unexplained blocking effect in (6) needs to be found.

Looking at (3) and (4), the most obvious difference is that the first case dealt with a singleton family, while the unavailable *Lisi* reading of (4) has two isomorphic derivations. An additional structural constraint on the two degenerate nodes of the syntactic *ziji* set obtains the correct result: βziji$_{T'*}$ may only dominate βziji$_{DP*}$ if there is no other isomorphic derivation where the c-command relationship which enables feature checking obtains. Framing the constraint in this way allows an escape hatch for *ziji* to have different properties in the most local domain, as the lowest clause attachment of βziji$_{T'*}$ will always reduce any derivation with local binding of *ziji* to a singleton derivation family, and any higher antecedent will always have

Table 6: Isomorphic Derivations for (4)

T* at:	piping	gaosu
DP* at: Wangwu	δ1: l=2	
Lisi	δ3: l=2	δ4: l=2
Zhangsan	δ5: l=2	δ6: l=2

a larger family of isomorphic derivations. This constraint isolates exactly the derivations where subject orientation and blocking appear to be suspended. Applying this constraint to the derivations in Table 6, δ4 will be ruled out, which in turn eliminates the derivational family associated with the antecedent *Lisi*, following the FFF.

There is an additional consequence for (6). Following this new well-formedness constraint on the combination of the βziji components, δ4 in Table 5 is also blocked. This leads to another situation which has so far not been seen: one in which all members of an isomorphic family of derivations are separately ill-formed. My final proposal is that this triggers another constraint between derivations: the Poisoned Path Principle, or PPP.

(7) POISONED PATH PRINCIPLE: If one isomorphic family consists entirely of ill-formed derivations, no other family of the same or greater l value is licit.

This principle captures the distinction between (4) and (6) where the same derivations are licit in one case but not the other. The distinction lies not in those derivations, but in the status of another family of derivations with the same l value. In (4), it is actually the well-formedness of δ3 which allows the δ5/δ6 family to survive.

This set of constraints yields the desirable result of capturing the distinction between (4) and (6), but other analytic options are available. Indeed, (Huang et al., 2009) describe the distinctions between local and long-distance *ziji* as so intractable that they suggest treating the two as distinct forms with different properties. A reviewer points out that following such a path would allow for a simpler characterization of subject orientation in the long distance domain as falling out from a requirement that βziji_{DP*} must c-command $\beta\text{ziji}_{T'*}$. It is also suggested that the blocking effect in (6) can be captured by adding another component to the *ziji* MCS which will check for agreement at some position in the predicate tree where the indirect object's φ features are checked. While this certainly has a precedent in the literature (Kallmeyer and Romero use a similar device to check for agreement with an indirect object), its use in this context would be hard to stipulate. Assuming that all three of βziji_{DP*}, $\beta\text{ziji}_{T'*}$, and the proposed $\beta\text{ziji}_{V'*}$ would have to match in φ features, then a clash would result if both $\beta\text{ziji}_{T'*}$ and $\beta\text{ziji}_{V'*}$ adjoined at the highest clause in the derivation of the ungrammatical matrix subject reading of (6). Recalling that Storoshenko and Han (2015) uses a minimum l value to force long distance binding, then a dedicated tree set for long distance *ziji* would need a minimum l based on the distance between βziji_{DP*} and αziji in the derivation tree, but the position of $\beta\text{ziji}_{V'*}$ does not change the l value. As such, there is no way to force $\beta\text{ziji}_{V'*}$ to attach to the higher clause rather than the lower one (which could potentially avoid the clash) short of writing in a stipulation that $\beta\text{ziji}_{V'*}$ and $\beta\text{ziji}_{T'*}$ must have a relative l of zero with respect to each other (tree local) and a relative l of one with respect to βziji_{DP*} (in the same clause as the subject), all within an overall l of two for the whole MCS (the anaphor is in an embedded clause). This is just to say that no treatment of these data will be devoid of stipulations somewhere in the account; it is left to further debate which are more likely. Furthermore, the local domain is no less complex, as (3) has a close counterpart in which subject orientation *does* obtain in the local domain:

(8) Wangwu$_i$ shuo Zhangsan$_j$ zengsong gei
Wangwu say Zhangsan give to
Lisi$_k$ yipian guanyu ziji$_{i/j/*k}$ de wenzhang.
Lisi one about self DE article
'Wangwu says that Zhangsan gave an article about self to Lisi.'

Unlike in (3), the lower clause indirect object of (8) cannot be an antecedent. So, just as a distinct long-distance account would require some derivational stipulations, a dedicated local treatment of *ziji* would need to be sensitive to the distinction between a "true" ditransitive as in (8) and the so-called *ba*-construction of (3).

In closing this discussion of the core analysis, it's worth pointing out that all of the cases discussed under the initially proposed analysis do scale to longer sentences than those used here. Looking at the core blocking example (2), it's clear to see how any reading where the antecedent c-commands a φ-feature mismatched element will incur a feature mismatch in one of its isomor-

phic derivations. Similarly, even though the constraints obviate blocking in the lowest clause, any clause higher with a mismatched antecedent will be blocked. In the case of (6), a mismatched non-subject at any embedding level will have no viable derivations: all attachments of βziji$_{T'}$ in lower clauses will incur feature mismatches because they are c-commanded by the antecedent, while attachments to the local clause or higher will all be ruled out because βziji$_{T'}$ will c-command its own antecedent DP within a family where other isomorphic derivations are available. By the PPP, this will block all antecedents in the same or higher clauses. Conversely, in (4), only derivations where βziji$_{T'}$ attaches to a position c-commanding a matched non-subject in an embedding clause will be ruled out, while the lower attachments will be felicitous derivations. While the FFF is triggered to rule out the non-subject antecedent, the existence of some well-formed derivations will obviate application of the PPP, allowing other longer derivation families to go through.

4 Interaction with Other Forms

When blocking occurs, and a bound variable use of *ziji* is impossible, Mandarin is similar to English in that it allows third person pronouns to act as bound variables:

(9) Mei-ge ren$_i$ dou renwei [wo$_j$ piping-le ta$_{i/*j}$.]
every-CL man all think I criticize-ASP he
'Every man thought that I criticized him.'

I assume that bound variable *ta* will have tree sets similar to English. Lacking the extra T′* component, no blocking effect obtains, and the derivation goes through. However, in the local domain, *ta* may not be used:

(10) * Mei-ge ren dou piping-le ta$_i$.
every-CL man all criticize-ASP he
'Every man criticized himself.'

Even though there should be nothing formally constraining the use of *ta* in (10), native speakers report *ziji* to be preferable in this context[3]. *Ta* is also dispreferred in a non-blocking version of (9):

(11) Mei-ge ren$_i$ dou renwei ta$_j$ piping-le ta.$_{?i/*j/k}$
every-CL man all think she criticize-ASP he
'Every man thought that she criticized him.'

For the long distance bound variable reading in (11), *ziji* is the clear choice, even though it would be ambiguous between a local or long distance reading. While it would be simple to characterize *ta* as an anti-local bound variable, examining the full set of data suggests this to be an overstatement. It is not merely that *ta* cannot be used as a bound variable locally, the facts are that *ta* is only ever used when no other form is available. One could make the same claim about bound variable uses of the English third person pronouns, which are not used in local reflexive contexts. Keeping in mind that TAG accounts of English reflexive pronouns have treated them as tree-local derivations (see for example Kallmeyer and Romero (2007) and and Frank (2008), among others), one could imagine a derivational economy account restricting uses of *her*: a derivation that yields the same meaning without resorting to delayed locality is available. Similarly, while isomorphic derivations using either *ta* or *ziji* will have identical l values in (10) and (11), the ambiguity introduced by the fact *ta* also has a non-bound referential use may block its use: the obligatorily bound *ziji* wins out. Given that this paper argues for constraints over sets of derivations, it is not unreasonable to propose that the same types of comparisons come into play when selecting the optimal expression of a desired meaning.

This reasoning calls for a re-examination the typological conclusions reached in Storoshenko and Han (2015), where three types of bound variable are defined: ones similar to *ziji* or *caki* which can be used locally or long distance, ones restricted to non-local uses formalized as an l value greater than one, and finally ones restricted only to the most local use where l=1. Firstly, that the present paper explains the nonexistence of any bound variables with minimum l values of two or greater. These would be hyper-anti-local bound variables whose antecedents need to be two or more clauses

[3]The bimorphemic *taziji* is also acceptable, though restricted only to local uses similar to English *himself*. A reviewer correctly points out that the example could be ruled out by Condition B of the standard binding theory; the same does not however apply to the bound *ta* in (11), which suggests more than just Condition B would be at play. Furthermore, the only implementations of Condition B for TAG that I am aware of (Champollion, 2008; Nesson, 2009) each make different assumptions about the semantics of binding than are made in this paper. Defining the full set of Chomskyan binding conditions in this particular version of STAG remains a matter for further research, though see Storoshenko et al. (2008) for discussion of Condition A as epiphenomenal.

removed. As seen here, this would amount to ruling out entire families of derivations with smaller l values, which would trigger the PPP. The account for why non-subjects are not antecedents and yet induce blocking with a feature mismatch is captured by exactly the derivational constraint which would rule out such a bound variable. However, one may wonder whether the anti-local category exists at all, or should rather just be re-cast as an epiphenomenon falling out from the fact that better local derivations are available?

Still, there is no compelling reason to eliminate the category of bound variable where l=1. While it could be the case that apparent local bound variables may warrant closer examination to see whether a tree-local derivation is available, there may be other cases where a certain MCS is forced to make use of delayed locality, but is constrained to using only as much as possible for a felicitous composition. One possibility for such a case outside the realm of binding would be subextraction in *wh*-movement:

(12) What$_i$ did you take a picture of t$_i$?

While it has been compellingly argued that *wh*-extraction should be treated as tree-local movement (Frank, 2002), such an analysis is untenable for (12), as the trace is in a different lexical projection than the *wh*-word's landing site. Treating this *wh*-movement as a MCS consisting of the *wh*-word and its trace would need a derivation with an l value of one, but the use of such a MCS would need to be held to that value, or else numerous subjacency violations would become possible. This analysis is held over for future work.

5 On Trans-Derivational Constraints

The analysis presented in this paper relies crucially on a set of trans-derivational constraints (TDCs) which run the risk of extending the generative capacity of the grammar. Here, I address this concern by comparing the present analysis with that presented in Freedman (2012). In that paper, two constraints are presented, along with a proof that a TAG with restricted TDCs is no more expressive than one without those constraints. The first constraint Freedman proposes is the DERIVATIONAL COMPLEXITY CONSTRAINT ON SEMANTIC INTERPRETATION or DCCSI: "A derivation d producing meaning m is ruled out if another shorter derivation d' also produces m." The maximal isomorphism constraint I have assumed to act as an initial filter on possible derivations has essentially this effect, militating against derivations with identical semantic outputs but larger l values in the syntactic derivation. This seems to be a reasonable comparison class for derivations where an uneven syntactic and semantic MCS pairing is allowed to combine using delayed locality. The FFF proposed above uses a subset of the same comparison class, the family of maximally isomorphic derivations. The second constraint Freedman proposes is a means of limiting the comparison class for TDCs. His LOCALITY CONSTRAINT ON COMPETITION dictates that any two derivation trees may only be compared if the are identical save for the daughters of one node, which may be either excised or replaced. The comparisons within families required by the FFF are similarly restricted in that they are identical in all measures save for the position in the tree of the βziji$_{T'}$ component. Applying the constraint symmetrically to the pair δ4 and δ5 in Table 2, each derivation differs from the other by the daughters of one node. Moving on to the PPP, there is a similar constraint on comparison classes when looking at the whole families being compared. Examining the relevant derivations in Table 5, the δ3/δ5 and δ4/δ6 pairs differ only by the positions of αziji. While TDCs are definitely tools to be used sparingly, restricting these constraints to applying within families or at most between adjacent families (allowing the PPP to apply transitively) provides similar restrictions as those discussed by Freedman.

6 Conclusion

In this paper I have presented a model of the behaviour of the Mandarin long distance anaphor *ziji*. By combining two well-formedness constraints on the relationship between elements in the *ziji* MCS with two more global constraints on derivation families, the blocking effect and certain asymmetries between local and long distance domains are captured. Furthermore, while the more global constraints provide motivation for the existing claim that hyper-anti-local bound variables should not exist, the existence of trans-derivational constraints poses a new question of whether anti-local bound variables are a *bona fide* category, or merely epiphenomenal.

Acknowledgments

I would like to the TAG+ reviewers for their comments and suggestions, all which have improved the quality of this paper, and also the audiences at Yale University and the University of Delaware who heard preliminary versions of the work contained here. This work was started at Yale under SSHRC Postdoc Fellowship 756-2010-067, and has been completed under a Faculty Start-Up Grant from the University of Calgary. My thanks also to Peng Han for providing some new data and confirming judgements. All errors are my own.

References

Daniel Büring. 2005. *Binding Theory*. Cambridge, UK: Cambridge University Press.

Lucas Champollion. 2008. Binding theory in LTAG. In Claire Gardent and Anoop Sarkar, editors, *Proceedings of the 9^{th} International Workshop on Tree Adjoining Grammars and Related Formalisms*, pages 1–8.

David Chiang and Tatjana Scheffler. 2008. Flexible composition and delayed tree-locality. In Claire Gardent and Anoop Sarkar, editors, *Proceedings of the 9^{th} International Workshop on Tree Adjoining Grammars and Related Formalisms*, pages 17–24.

Peter Cole, Gabriella Hermon, and Li-May Sung. 1990. Principles and parameters of long-distance reflexives. *Linguistic Inquiry*, 21(1):1–22.

Robert Frank. 2002. *Phrase Structure Composition and Syntactic Dependencies*. Cambridge, MA: MIT Press.

Robert Frank. 2008. Reflexives and TAG semantics. In Claire Gardent and Anoop Sarkar, editors, *Proceedings of the 9^{th} International Workshop on Tree Adjoining Grammars and Related Formalisms*, pages 97–104.

Michael Freedman. 2012. Scope economy and TAG locality. In *Proceedings of the 11^{th} International Workshop on Tree Adjoining Grammars and Related Formalisms*, pages 223–231.

James C.T. Huang, Yen-hui Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge University Press.

Laura Kallmeyer and Maribel Romero. 2007. Reflexives and reciprocals in LTAG. In Jeroen Geertzen, Elias Thijsse, Harry Bunt, and Amanda Schiffrin, editors, *Proceedings of the Seventh International Workshop on Computational Semantics*, pages 271–282.

Rebecca Nesson. 2009. *Synchronous and Multicomponent Tree-Adjoining Grammars: Complexity, Algorithms, and Linguistic Applications*. Ph.D. thesis, Harvard University.

Hong-Ki Sohng. 2004. A minimalist analysis of X^0 reflexivization in Chinese and Korean. *Studies in Generative Grammar*, 14(3):375–396.

Dennis Ryan Storoshenko and Chung-hye Han. 2015. Using Synchronous Tree Adjoining Grammar to derive the typology of bound variable pronouns. *Journal of Logic and Computation*, 25(2):371–403.

Dennis Ryan Storoshenko, Chung-hye Han, and David Potter. 2008. Reflexivity in English: An STAG analysis. In Claire Gardent and Anoop Sarkar, editors, *Proceedings of the 9^{th} International Workshop on Tree Adjoining Grammars and Related Formalisms*, pages 149–157.