

Two Types of Variable Elements in Hmong Anaphora*

David Mortensen
University of California, Berkeley

March 4, 2004

1 Introduction

Hmong¹ is like Thai (Lasnik 1986; Narahara 1995), Vietnamese (Lasnik 1986; Narahara 1995), Quiégolani Zapotec (QZ; Black 2000) and San Luis Quiavini Zapotec (SLQZ; Lee 2003) in that it apparently allows pronouns and R-expressions to be bound with reflexive readings:

- (1) a. Nwgi yeej qhuas nwgi.
3SG always praise 3SG
'He always praises himself.'
- b. Puabi tsuas yog xaav txug puabi.
3PL only is think about 3PL

*Earlier versions of this paper have benefited from comments and suggestions from numerous individuals including Lynn Nichols, Line Mikkelsen, Kenneth Safir, Norbert Hornstein, and Rosemarie Déchaine. I thank them for their generous advice and comments, but retain responsibility for all errors of fact and analysis that remain. A further debt of gratitude is owed to the many Hmong teachers and consultants (too numerous to mention) who provided data, judgments, and guidance relevant to this paper. A special debt is owed to Neeb Hawj (Leena Her), Nchais Laaj Hawj, Tshuv Ntxaij Yaaj (Shawn Yang), Xab Yaj (Sa Yang), Ntaub Muas (Dao Moua), and Iab Hawj (Ia Her).

¹The Hmong data in this paper are from Mong Leng (also called Green Hmong, Blue Hmong, Blue Meo, Hmong Njua, etc.), a dialect of Hmong spoken by around 1,000,000 speakers in Southern China, Vietnam, Laos, Thailand, and various western countries (due to the Hmong diaspora from Laos following the "Secret War" there). Hmong is a member of the Far Western Hmongic group of the Western Hmongic branch of the Hmongic subfamily, which is, in turn, a member of the Hmong-Mien family. The wider genetic affiliations of Hmong-Mien are undetermined.

Like other members of the Hmong-Mien family, Mong Leng is tonal, having 7 contrastive tones. In the RPA orthography used in this paper (and by most Hmong in the West), tones are represented by letters at the end of syllables (-b, -j, -v, -θ, -g, -s, and -m). There are other aspects of this orthography that the naive linguist may not find completely transparent. For example, doubled vowels represent the presence of nasalization or a velar nasal coda. For a good introduction to the sounds and orthographic representation of the closely related Hmong Daw (White Hmong) dialect, see Ratliff (1992) or Heimbach (1969). Mong Leng has several sounds not present in Hmong Daw, and these are represented here in the conventional Mong Leng adaptation of the RPA orthography. The only important differences are as follows: <dl> [tɰ ~ kɰ], <dlh> [t^hɰ ~ k^hɰ], <ndl> [ntɰ ~ ŋkɰ], <ndlh> [nt^hɰ ~ ŋk^hɰ], and <aa> [ãŋ ~ ǎ].

‘They only think about themselves.’

- (2) Pov_i yeej qhuas Pov_i.
Pao always praise Pao

‘Pao always praises himself.’

In Hmong, this type of binding relationship (where the bound element is a pronoun or name identical to its binder) is allowed only between A-positions. A'-binding displays different properties which cannot be easily reduced to the same principles, but which seem to involve some type of lexical competition. Take the following examples of variable binding from a topic:

- (3) a. [Tug twg]_i los nwg_i yeej nyam yawg.
CLF which TOP 3SG always like mister

‘Anybody would surely like that guy.’

Any x : x like y

- b. [Tug twg]_i los yawg yeej nyam nwg_i.
CLF which TOP mister always like 3SG

‘That guy would surely like anybody.’

Any x : y like x

- (4) a. [Tug twg]_i los yawg_i yeej nyam Pov.
CLF which TOP mister always like Pao

‘Anybody would surely like Pao.’

Any x : x like Pao

- b. [Tug twg]_i los Pov yeej nyam yawg_i.
CLF which TOP Pao always like mister

‘Pao would surely like anybody.’

Any x : Pao like x

Hmong has not previously been discussed in the literature on binding. However, the phenomena attested in the data from Hmong present a number of interesting problems for various proposals in binding theory, both classical and recent. These phenomena including seeming Condition B and C violations, chains of semantic dependencies, and apparent competition for dependent reference.

The purpose of this paper is both to explore and describe some of the Hmong data with implications for binding theory and to offer a new proposal to account for these data building upon the earlier work of Lee (2002, 2003) and Safir (2002a,b). The central focus, in terms of empirical phenomena, is upon reflexives of the bound pronoun/name type and upon distributive variables. I will argue that both of these clusters of phenomena can be understood in a model of anaphoric reference based upon lexical competition and dependence as a scalar property of lexical items. Further, I will argue that the bound pronouns and R-expressions of Hmong (and other languages displaying similar phenomena) are actually instances of a single lexical anaphor (α) which is spelled-out as if it were the same lexical item as its antecedent.

2 Copy Anaphora

2.1 Parametric Condition C and the RHC

It has long been recognized that some languages allow names and pronouns to be bound with reflexive readings. Lasnik (1986) recognized that these were a problem for binding theory, since—at the very least—they are apparent violations of Condition C as it is canonically formulated². Lasnik proposed that Condition C must be parametric, since it seems to impose different requirements in some languages than in others. He decomposed Condition C into two parts: a universal Referential Hierarchy Condition (RHC), stating that “A more referential expression must be free from a less referential one,” and a language specific part that could be more or less restrictive (as in English versus Vietnamese) or absent altogether (as in Thai). Black (2000) presented a similar analysis of Quiégolani Zapotec to account for the presence of bound names and pronouns in this language.

However, Lee (2002) pointed out that these hypotheses makes a number of predictions that do not hold true for SLQZ. The predictions of this hypothesis do not hold true for Hmong either, as a brief survey will indicate.

2.1.1 More referential forms should be able to bind less referential forms

Lasnik points out that, in Thai, pronouns cannot bind names, and attributes this fact to the RHC. This is also true in Hmong:

- (5) a. Nwg_i yeej qhuas $Pov_{*i/j}$.
3sg always praise Pao
‘He always praises Pao.’
*‘Pao always praises himself.’
- b. Nwg_i yeej qhuas $Txiv_{*i/j}$.
3sg always praise Father
‘He always praises Father.’
*‘Father always praises himself.’

If this failure to bind is due to the RHC, then it follows that names like *Pov* must be more referential than pronouns like *nwg*. This being the case, referentiality cannot rule out cases where a potential binder is a name and the bound element is a pronoun. Some other factors must rule out local binding in sentences like those in (6):

- (6) a. Pov_i yeej qhuas $nwg_{*i/j}$.
Pao always praise 3sg
‘Pao always praises him.’
*‘Pao always praises himself.’

²The binding of “pronouns” in the Southeast Asian languages Lasnik was discussing might not be seen as a Condition B violation by virtue of the fact that it is very difficult to distinguish between pronouns and R-expressions in these languages. These so-called pronouns display a number of properties of R-expressions, and they should perhaps be expected to be subject to Condition C rather than Condition B. However, in Hmong the distinction between pronouns and R-expressions is much clearer than in Thai or Vietnamese.

- b. Txiv_i tua nwg^{*i/j}.
 Father kill 3sg
 ‘Father killed him.’
 *‘Father killed himself.’

Of course, such interpretations should also be ruled out by Condition B. Problematically, though, we cannot invoke Condition B to rule out these cases since we have plenty of examples of pronouns being bound by other pronouns, as in (1). Thus, the RHC makes the wrong predictions about the local binding of pronouns in Hmong.

2.1.2 Names should be able to bind other names

Furthermore, if there is no Condition C, names should be able to bind other names locally as long as they do not differ in referentiality. This prediction seems to be born out by examples like (2), where the binding name and the bound name are identical, but if we construct a case where the two names are not identical, we get a different result. Hmong culture provides us with just such an opportunity, since adult males can be referred to by either of two names, the *npe hluas* or ‘young-name’, which is given at birth and the *npe laug* or ‘old-name’ which is given later in life. Even if we set up a context where the *npe laug* and the *npe hluas* are known to refer to the same individual, a reflexive reading is not possible:

- (7) a. Kub npe laug hu ua Tshuv-Ntxaij.
 Kou name old call do Chu-Ndzai.
 ‘Kou’s old-name is Chu-Ndzai.’
- b. Kub_i yeej thuum Tshuv-Ntxaij^{*i/j}.
 Kou always criticize Chu-Ndzai
 ‘Kou_i always criticizes Chu-Ndzai.’
 *‘Kou_i always criticizes himself_i.’
- c. Tshuv-Ntxaij_i yeej thuum Kub^{*i/j}.
 Chu-Ndzai always criticize Kou
 ‘Chu-Ndzai_i always criticizes Kou_j.’
 *‘Chu-Ndzai_i always criticizes himself_i.’

Even if there was some difference in the referentiality of an *npe laug* versus an *npe hluas*—a possibility open in more articulated scales of referentiality like that proposed by Narahara (1995)—a RHC-based account does not seem to capture the relevant generalization over these data.

2.1.3 Pronouns cannot bind non-identical pronouns

But perhaps more significant failure of the RHC/parametric Condition C account is that it fails to capture the parallels between names and pronouns (whether in QZ, SLQZ, Thai, Vietnamese, or Hmong). That such a parallel should exist is not surprising in Thai

and Vietnamese, where the line between pronouns and names is difficult to draw, and is probably conceptually unnecessary. It seems equally natural in the case of SLQZ, of which Lee (2003:86) claims “SLQZ pronouns serve purely deictic functions.” She claims that SLQZ pronouns resist non-local A-binding and are always marked for deictic features (either social deixis or social deixis). She proposes, therefore, that SLQZ pronouns are like R-expressions, subject to Condition C rather than Condition B. In Hmong, in contrast, personal pronouns are not marked for deictic features (including social deixis, in contrast to many other Southeast Asian languages). Furthermore, pronouns may be A-bound non-locally:

- (8) Naagmo Pov has tas nwg yog ib tug xibfwb.
 Yesterday Pao say that 3SG COP one CLF teacher
 ‘Yesterday, Pao_i said that he_i was a teacher.’

Given these facts, it seems surprising that Hmong pronoun apparently pattern with names precisely in the area of reflexive anaphora. Not only can pronouns bind identical instances of themselves, as in (1), they cannot bind non-identical pronouns. For example, the exhaustive pronoun *suavdlawg* ‘everybody’ may bind *suavdlawg* locally, and may bind the plural pronoun *puab* ‘3SG’ non-locally, but may not bind *puab* locally:

- (9) a. Suavdlawg_i yeej qhuas suavdlawg_i.
 everyone always praise everyone.
 ‘Everyone (as a group) always praises themselves (as a group).’
 b. Suavdlawg_i yeej khaav tas puab_i zoo xwb.
 everybody always brag that 3SG good only.
 ‘Everyone (as a group) always brags that only they are good.’
 c. Suavdlawg_i yeej qhuas puab_{*i/j}
 everyone always praise 3PL
 ‘Everyone (as a group) praises them (a different group).’
 *‘Everyone praises themselves.’

Lee (2002, 2003) has named this requirement that the bound element in this type of anaphora be identical to its antecedent the Identical Antecedent Requirement (IAR). Undoubtedly the largest shortcoming of accounts of bound names prior to Lee (2002) was the failure to capture this descriptive generalization.

2.2 Bound pronouns and R-expressions as bound copies

Lee (2003) explains the IAR and the other characteristics of the apparent bound pronouns and R-expressions (which Lee reduces to the single case of bound Pro-DPs, as per the typology of Déchaine and Wiltschko 2002) by claiming that the bound elements in these constructions are actually anaphoric. As she states in Lee (2003:109):

Locally bound copies are base-generated and spelled out as copies of their antecedents. Because SLQZ has no distinct reflexive pronouns, this “copying” at Spell-Out serves to reflexive mark the predicate. At LF, the lexical content of the copy is deleted, thus forcing the bound-variable reading.

In other words, the bound elements are not names or pronouns at all. Rather, they are a different type of syntactic entity that is spelled out to be identical to a pronoun or name antecedent. For the sake of clarity and brevity, let us call this type of entity (Lee’s bound variable copies) α .

2.2.1 Referentiality of copies

The analysis of these α -reflexives as bound variables allows Lee to capture a very important generalization in addition to the IAR: α -copies of R-expressions are not fully referential. This can be seen in ellipsis contexts, where elided α -reflexives produce obligatory sloppy readings in SLQZ, Thai, and Hmong:

- (10) a. B-gwi’ih Gye’eihlly lohoh Gye’eihlly zë’cy cahgza’ Li’eb.
PERF-look Mike at Mike likewise Felipe
 ‘Mike looked at himself, and Felipe did too.’ (i.e. Felipe looked at himself/*Mike) (SLQZ; Lee 2003:89)
- b. John konnuad khong John lae Peter ko muankan.
John shave of John and Peter the same
 ‘John shaved himself, and Peter did too.’ (Thai; Lee 2003:93)
- c. Pov yeej qhuas Pov; Maiv los kuj ua le.
Pao always praise Pao; Maiv TOP also do as
 ‘Pao always praises himself, and so does Maiv (praise herself).’ (Hmong)

The obligatory nature of these sloppy readings suggests that α copies are anaphoric, or—to use the terminology of Safir (2002a,b)—the relationship between the first instance of the name and the α -copy is one of dependent reference rather than independent coreference. That is to say, α -copies are acting as “variables”.

2.2.2 α -copies and A’-positions

There is another interesting generalization, mentioned by Black (2000) for QZ and by Lee (2002, 2003) for SLQZ and Thai: quantified expressions may not bind α -copies. Take the following examples from SLQZ:

- (11) a. R-yu’lààa’z ra bxuuhahz ra bxuuhahz.
HAB-like PL priest PL priest
 ‘The priests like themselves.’ (Lee 2003:84)
- b. * B-guhty cho’nn ra bxuuhahz cho’nn ra bxuuhahz.
PERF-kill three PL priest three PL priest
 ‘Three priests killed themselves.’ (Lee 2003:89)

When not quantified, the noun meaning ‘priest’ can bind a copy of itself. On the other hand, quantification makes this kind of binding impossible. It is also true that quantified expressions cannot bind α -copies in Hmong:

- (12) [Ob tug xibfwb]_i yeej qhuas [ob tug xibfwb]_{*i/j}.
 two CLF teacher always praise two CLF teacher

‘The two teachers always praised the (other) two teachers/*themselves/*each other.’

In Hmong, however, it is not obvious that it is quantification that is ruling these sentences out. Typically, bare nouns cannot bind α -copies either:

- (13) * Xibfwb yeej qhuas xibfwb.
 teacher always praise teacher

This is unsurprising, since, in Hmong, bare nouns (those lacking classifiers) denote properties rather than entities and cannot typically antecede or be replaced by pronouns. It would be odd if such nouns could bind an anaphor, and if it did, it would be hard to say what the meaning of the resulting construction could be.

But if a noun is individuated by adding a classifier (a necessary precondition for quantification and possession) it is still incapable of binding a copy of itself:

- (14) [Tug xibfwb]_i yeej qhuas [tug xibfwb]_{*i/j}.
 CLF teacher always praise CLF teacher

‘The teacher praise the teacher/*himself.’

This is somewhat harder to explain. In (14), there is no overt marker of quantification, although the semantics of these DPs are arguably quantificational, both in that their referents are obligatorily singular, and because they are definite (and definiteness has sometimes been viewed as having quantificational properties). Since indefiniteness is marked by the numeral *ib* ‘one’, it may be that definiteness is marked by a null quantifier that still undergoes quantifier raising at LF. If this is not the case, then it would appear that quantified expressions cannot bind α variables in Hmong simply because they are not names, which along with pronouns have a monopoly on α -binding. However, if this null quantifier speculation is true, it would allow us to reduce the prohibition on binding by quantified expressions to a prohibition against binding α from A’-positions.

We can enforce this hypothesis with the observation that DPs in (structurally and morphologically defined) topic positions cannot bind α variables either:

- (15) a. Pov_i mas nwg_i tsi nyam tuabneeg.
 Pao TOP 3SG NEG like people

‘Pao, he doesn’t like people.’

- b. * Pov_i mas Pao_i tsi nyam tuabneeg.
 Pao TOP α NEG like people

Intended: ‘Pao, he doesn’t like people.’

This fact is predicted by the hypothesis that α is never bound by A' antecedents. Desirably, this hypothesis also predicts that phrases raised by Wh-movement should not be able to bind α copies. This is true of SLQZ, where Wh-expressions (which undergo overt raising) must be reflexivized via an alternative strategy (Lee 2003:92). In Hmong, which lacks overt Wh-movement, α -copies are also not a valid means of reflexivization:

- (16) [Tug twg]_i tua [tug twg]_{*i/j}?
 CLF which kill CLF which
 'Who_i killed who_j/*himself_i?'

If Hmong displays Wh-raising at LF, then this example could be explained by the A'-hypothesis. However, this explanation is probably unnecessary, since the Wh-expressions that would be candidates for reflexivization are classified and are therefore ruled out on other grounds.

Since this hypothesis allows these restrictions against binding by a Wh-expression, a quantified expression, and a topic to be reduced to a single structural condition, I will assume for the remainder of the paper that these effects are due to a prohibition against the A'-binding of α . This position differs markedly from that of Lee (2003), who attributes the non-participation of quantified arguments in copy anaphora to type-theoretic factors.

2.2.3 α -copy as a strategy for marking reflexive predicates

Lee employs a notion of reflexivity influenced by Reinhart and Reuland (1991). However, she notes that the SLQZ data seem inconsistent with R&H's assumption that reflexive predicates must be either overtly marked (as with SELF reflexives) or else inherently reflexive. She argues that the Spell-Out of α as a copy of its antecedent serves the same purpose as SELF anaphors. Since "SLQZ lacks a distinct series of anaphoric pronouns," Lee assumes (2003:109):

... that bound local copies serve the same function as SELF pronouns in SLQZ: like SELF pronouns, they are DPs, they lack independent reference (as seen from their bound-variable status), and they should (obviously) the person/number features of their antecedents. Given that bound copies are often the only way to express local reflexivity in SLQZ (even in predicates that aren't inherently reflexive), this suggests that bound copies, too, can serve as potential reflexivizers.

This seems to imply that languages that employ copy anaphors do so because they lack other anaphoric options—they do not have SELF reflexives and therefore need other means of marking reflexive predicates. This cannot be exactly true, however. SLQZ does have another means of marking reflexive predicates that can be employed in certain environments that copy anaphors may not. For examples, Wh-expressions cannot bind α , so the reflexive marker *-ag* is suffixed to a pronoun instead (Lee 2003:92):

- (17) Tu b-guhty laa-g-ih.
 who PERF-kill PRON-REFL-3SG.DIST
 'Who killed him/herself?'

Lee does not discuss this marker in detail in her paper, so it is difficult to know how its distribution relates to that of α reflexives.

Hmong, on the other hand, clearly has both α reflexives and pronoun-SELF reflexives. Though it is difficult to make a definitive statement about the morphological structure of Hmong SELF anaphors, they appear to consist of a pronominal possessor and a classified noun, *tug kheej* ‘CLF self’³. Thus, *kuv tug kheej* ‘myself’ seems on the surface to have the same structure as *kuv tug kwv* ‘my younger brother’. However, *tug kheej* seldom occurs without an overt pronoun, and even in these cases it seems to behave as a reflexive or logophor rather than a prototypical noun⁴. The set of environments in which *tug kheej* reflexives may occur is a superset of those in which α may occur. α anaphors, in Hmong, seem to occur only with local c-commanding antecedents in A-positions. *Tug kheej* reflexives, on the other hand, can be bound from A'-positions, as well as A-positions:

- (20) a. *Pov_i yeej qhuas Pov_i.*
 Pao always praise Pao
 ‘Pao always praises himself.’
- b. *Pov_i yeej qhuas [nwg tug kheej]_i.*
 Pao always praise 3SG CLF self
 ‘Pao always praises himself.’
- (21) a. *[Txhua tug tub hluas]_i yeej qhuas [nwg tug kheej]_i.*
 every CLF boy youth always praise 3SG CLF self
 ‘Every boy praises himself.’
- b. *[Txhua tug tub hluas]_i yeej qhuas [txhua tug tub hluas]_{*i/j}.*
 every CLF boy youth always praise * α
 ‘Every boy praises every (other) boy.’
 *‘Every boy praises himself.’

Since α can not be bound by A'-elements including quantifier phrases, α cannot receive distributive readings directly. Reflexive predicates that are distributed over the individuals of a set can be formed with *tug kheej*, as in (21a)—a strategy functionally

³Hmong nouns must normally take a classifier in order to be possessed (Bisang 1993:29-30)

⁴In fact, it can be argued that these cases are actually instances of *pro tug kheej* where there is a pronoun, but it is not overt. They typically mean something like ‘oneself’ or ‘one’s own’, as in the following example:

- (18) *pro_i tsi yog [tug kheej kev txav-txim]_i*
 NEG is CLF self way decide

‘It is not one’s own decision’

Occasionally, though, one can find contexts where *tug kheej* means something like identity:

- (19) *pro_i maam-le qha [pro_i tug kheej] rua koj.*
 eventually tell CLF self to 2SG

‘(He’ll) eventually reveal (his) identity to you.’

identical to that described for Thai by Lee (2003:94), or can be mediated by the distributive pronoun *nyas*, the whole function of which is to act as a variable distributing over the set delineated by a topic:

- (22) [Txhua tug tub hluas]_i mas nyas_i yeej qhuas nyas_i.
 every CLF boy youth TOP DIST always praise α .

‘Every boy, each praised himself.’

This is very similar to the distributive reflexive construction that Lee (2003:90) describes for SLQZ, where a topicalized QP binds a third person distal pronoun⁵.

Tug kheej-reflexives can also function as “long distance anaphors”:

- (23) a. [Txhua tug tub hluas]_i yeej khaav tas [nwg tug kheej]_i zoo xwb
 every CLF boy youth always brag that 3SG CLF self good only

‘Every boy brags that he alone is good.’

- b. Yug_i cov rov qab qhuas yug_i los [yug tug kheej]_i rov qhuas yug_i
 one CLF return back praise one or one CLF self return praise one
 xwb.
 only

‘Only one’s group praises one or only *oneself* praises one.’

- c. Nwg_i nyam thaam txug lwm tug tej teeb-meem kws [nwg tug kheej]_i
 3SG like talk about other CLF PL problem REL 3SG CLF self
 tub tsi paub qhov tag.
 really NEG know thing real

‘He likes to talk about other people’s problems regarding which he doesn’t know the truth.’

The fact that the distributions of these two types of forms is so different seems to indicate that the effects of SELF-marking and α copy are not equivalent in their syntax. Their semantics do not seem to be equivalent, either, especially in emphatic or quasi-reflexive contexts. Many sentences containing α have an alternate emphatic reading, where α marks the predicate as having taken place of the agent’s own volition or agency:

- (24) a. Nwg_i tua nwg_i.
 3SG kill α

‘He killed himself.’

‘He killed (him/her/it) of his own volition.’ (No one told him to do it).

⁵The fact that this SLQZ distal pronoun can act as a distributive variable is surprising in light of the fact that it is marked for deictic features and that, according to Lee’s (2003:86) account, SLQZ pronouns are pro-DP pronouns (as described by Déchaine and Wiltschko 2002).

- b. Pov_i paub Pov_i.
 Pao know α
 ‘Pao knows himself.’
 ‘Pao knows it of himself.’ (No one told him or taught him).

In contrast, *tug kheej* forms cannot provide these types of ‘own volition’ or ‘own agency’ readings. Instances of *tug kheej* forms that do not mark classically reflexive predicates are most often logophoric, as in (23c), but may also contribute a restriction such as ‘alone’ or ‘only’ as in (23a) and (23b). In other words, it does not seem that α and *tug kheej* can be reduced to cases of exactly the same predicate marking phenomenon. This does not mean, however, that α -copying does not function to mark predicates as reflexive, but simply that it is not equivalent to the SELF marking of predicates.

2.2.4 α -copies as long distance anaphors

In Hmong, the evidence seems to suggest that α -anaphors must be locally bound. This fact is interesting in light of Lee’s (Lee 2002, 2003) claim that α -copies can act as long-distance anaphors in SLQZ. On examination, the argument for long-distance α anaphors in both SLQZ and Hmong seems to rely largely upon the presumption of Condition C. It is questionable whether the copies in the relevant examples really do act as variables, and without Condition C, there seems little reason for assuming that they do.

Lee claims that bound copies of R-expressions can occur as objects of embedded clauses. In Hmong, it is also true that two identical R-expressions can be co-construed non-locally within the same sentence, as show by example (25):

- (25) Pov_i has tas Maiv nyam Pov_{i/?j}.
 Pao say that May like Pao
 ‘Pao said that May likes him.’

Such sentences are acceptable in Hmong, but the coreferent reading is pragmatically strange. In ellipses contexts, both strict and sloppy readings are available, as Lee (2003:95) also concedes in the case for SLQZ, but the sloppy reading is marginal:

- (26) Pov_i has tas Maiv nyam Pov_i; Tub los kuj ua le hab.
 Pao say that May like Pao Tou TOP also do as too
 ‘Pao said that May likes him, and so did Tou (say that May likes Pao)’
 ??‘Pao said that May likes him, and so did Tou (say that May likes Tou)’.

Lee also points to apparent copies in adjunct clauses as evidence for bound copies as long distance anaphors. Equivalent sentences occur in Hmong:

- (27) Thaus Pov_i ua teb Pov_{i/j} has lug-txaj.
 when Pao do fi eldPao speak poetry
 ‘While Pao did farm work, Pao sang poetry.’

In ellipses of sentences of this type, sloppy readings are favored, but the strict reading is also available:

- (28) Thaus Pov ua teb Pov has lug-txaj; Tub los kuj ua le.
While Pao do field Pao speak poetry Tou TOP also do as

‘While Pao did farm work, Pao sang poetry, and so did Tou (sing poetry while doing farm work).’

‘While Pao did farm work, Pao sang poetry, and so did Tou (sing poetry while Pao did farm work).’

Lee (2003:95) minimizes the significance of the fact that strict readings are available in these contexts in SLQZ, asserting that these expressions “show the same interpretive behavior as local and long-distance anaphors crosslinguistically.” However, it is significant that the same type of readings can be obtained for R-expressions in English. Some speakers of English (including the author) accept (29) as grammatical (if awkward):

- (29) Jenny sings while Jenny works.

For these speakers, a sloppy reading is favored in ellipses:

- (30) Jenny sings while Jenny works, and so does Dawn.

‘... Dawn sings while Dawn works.’

? ‘... Dawn sings while Jenny works.’

If the mere availability of a sloppy reading indicates that a syntactic element is functioning as a variable, we are led to the problematic conclusion that, for some speakers of English, in some contexts, names can serve as anaphors. If we wish to avoid this conclusion, we must concede that, contra Koopman and Sportiche (1989), the availability of a sloppy reading is a necessary but not a sufficient condition for classifying a syntactic elements as a variable. If we don’t consider these English names to be long distance anaphors, we should probably not consider the analogous expressions in Hmong to be long-distance anaphors either, since their properties are quite similar (and contrast markedly with local α -copies in Hmong). If there were no Condition C, we would have little motivation for making either assertion.

In addition to the lack of evidence that non-local repetitions of names in Hmong have the same status as local α copies, there is an interesting piece of positive evidence that these two types of entities are different: local α copies are unstressed, and their tones are slightly depressed. In contrast, coconstrued iterations of a name within a sentence are all stressed, and do not display the same tonal perturbation as the local copies. Take the following example, where stressed instances of the impersonal singular pronoun *yug* are printed in capitals but the unstressed, tonally depressed instance is not:

- (31) Xob ca YUG_i rov qhuas yug_i; yuav-tau ca luag qhuas YUG_i xwb.
don’t let IMP.SG return praise α should let IMP.PL praise IMP.SG only.

‘One ought not oneself to praise oneself; one should only allow others to praise one.’

The same patterns holds for bound name copies:

- (32) POV_i yeej qhuas Pov_i .
Pao always praise α
'Pao praises himself.'

However, coconstrued non-local iterations of names have the same prosodic characteristics as names that are not coconstrued with any other name in the sentence:

- (33) a. MAIV nyam POV.
May like Pao
'May likes Pao.'
- b. POV_i xaav has tas MAIV nyam POV_i .
Pao think say that May like Pao
'Pao thinks that May likes Pao.'

This fact is suggestive. Some earlier researchers have suggest that, in languages like English, deictic pronouns are stressed but bound pronouns are unstressed. In this case, only the local copies seem to behave like bound pronouns. Certainly, Lee's analysis does make a distinction between local copies and non-local copies. The local copies are base generated, while the non-local copies are the residue of illicit movement. While this difference could be accounted for by the difference between the mechanisms that produce them, this set of assumptions seems to predict the opposite state of affairs: that local copies should be stressed and non-local copies should be unstressed.

2.3 Proposal

Lee's (2002; 2003) work brings us at least half way to a insightful analysis of copy anaphora in Hmong. My proposal differs from Lee's in the following respect: while Lee treats α -copies as base-generated copies—DPs that the syntax treats as anaphors—I propose that α is actually a lexical item on par with *PRO* or *pro* which the syntax treats as a fairly typical anaphor, except that it can only be licensed by an antecedent in an A-position. The remaining conditions on A-binding can be reduced to a single competitive algorithm, following Safir's (2002a; 2002b) Form to Interpretation Principle (FTIP), that only regulates dependent reference.

This renders the binding conditions, and other principles of grammar that target pronouns or R-expressions as classes, superfluous. Instead of talking about anaphors, pronouns, and R-expressions as discreet categories, and assuming that any complementarity in their distribution is accidental, we can simply distinguish pronominal elements that are subject to special antecedent licensing conditions from those that are not and capture the remaining distributional facts in terms of competition along a scalar axis that I will call, following Safir, dependence.

2.3.1 Dependence and the FTIP

Safir's (2002a; 2002b) lexical competition model of anaphora assumes that languages have fixed pronominal inventories. Within such an inventory, the members vary in their ability to support dependent reference (inversely correlated with their ability to independently refer). Presumably, this property of dependence is a function, at least in part, of the internal morphological or syntactic structure of a pronominal item. In general, the following hierarchy holds:

- (34) anaphor » pronoun » R-expression

The FTIP is a principle that governs dependent reference:

(35) **Form to Interpretation Principle (FTIP)**

If x c-commands y and z is not the most dependent form available in position y with respect to x , then y cannot be directly dependent on x (Safir 2002b:16)

This condition replaces Conditions B and C. Dependent reference for a given form is not ruled out because of some principle that makes direct reference to its features, but because there is some other form that could occupy its place that is more dependent. Dependent reference is prevented by virtue of “pragmatic obviation”:

(36) **Pragmatic Obviation**

If FTIP does not permit y to be interpreted as dependent on x , then x and y form an obviate pair.

Obviate pairs are expected not to be coreferent, but “independent coreference” between members of obviate pairs can be induced by other factors. Such coreference between members of an obviate pair always indicates a semantically or pragmatically marked construction—one that differs in content from the equivalent construction containing the most dependent available form.

Condition A is the only of the binding principles that is retained. It is restated as a condition called Local Antecedent Licensing.

(37) a. **Local Antecedent Licensing (LAL)**

An anaphor must be c-anteceded in Domain D.

b. **X c-antecedes Y** if X covaries with Y and X c-commands Y.

However, not all anaphors are subject to LAL. The distributions of so-called “long-distance anaphors” (in Safir's terminology, unbounded dependency or UD forms) are not governed by this principle.

On superficial examination, a competitive reformulation of binding theory seems rather ill-suited to account for the Hmong binding data. After all, the most obvious interest of α -reflexives lies in the fact that the bound elements are apparently highly referential. Furthermore, two other areas of interest notes—the apparent overlap in environments between α reflexives and *tug kheej* forms, and the non-local instances that Lee treats as long-distance anaphors, which appear to compete with pronouns, both seem to be cases of non-complementarity. To the extent that it can accommodate

these data, it might seem that a competitive theory would fail to make any predictions at all. But in fact, given the right assumptions about α , this lexical competition theory gives a subtle but reasonably insightful account of all these phenomena.

The bound elements in α reflexives are a problem for a condition like the FTIP only if they are evaluated as if they are pronouns and names instead of anaphors. The obvious solution is to treat α as an anaphoric lexical item (universally available, but appearing parametrically, or par with *PRO* and *pro*) that enters into competition with the other members of the Hmong pronominal inventory. To the interpretive interface, α looks like an unexceptional reflexive anaphor. It is only at Spell-Out that this element takes on the guise of its antecedent. At that point, it is exposed as if it were the same lexical item as the form on which it depends for reference⁶.

In Hmong, α is the most dependent form. However, it is subject to rather strict distributional constraints. It is subject, at least, to LAL, and perhaps to a stricter condition that requires it to be licensed by an A-binder (thus accounting for its inability to be bound by quantificational or topicalized antecedents). In contexts where it cannot occur, the pragmatically unmarked reflexives are pronoun-*tug kheej* forms. In contexts where it can occur, pronoun-*tug kheej* forms can appear, but always with special (object focus) pragmatics. The FTIP also prevents pronouns from being locally bound, accounting for the data in (6) without referring to pronouns as a class or stipulating their domain in any way. The pragmatic oddness of cases like (25), where an R-expression is coconstructed with a non-local R-expression, is also easy to account for in these terms, since a more dependent form (a pronoun) is always available in these environments. All of this is done without principles like Conditions B and C and only in terms of properties of lexical items such as availability (whether a derivation can converge with the relevant lexical item in its enumeration) and dependence (the relative ability of the lexical item to behave as a variable).

3 Topic-Bound Variables

So far, an attempt to characterize the properties of variable elements in Hmong, we have mentioned three important types of variables: α variables that are always locally bound from A-positions and give reflexive readings, *tug kheej* forms that may be bound either locally or non-locally from A or A'-positions for reflexive, distributive reflexive, or logophoric readings, and the distributive pronoun *nyas*, which is always bound from A'-positions and serves only to distribute the predicate over the individuals referred to by the topic or within the scope of a quantified expression. There is a complete

⁶In its earliest formulation, I conceived of α as a morpheme rather like the *RED* morpheme that features in various accounts of reduplication both classical (Marantz 1982) and contemporary (McCarthy and Prince 1995). This morpheme would simply copy its phonological content from its antecedent. Such an analysis works fine for Hmong, where the copies never differ in their phonology (aside from the subphonemic lowering that can be noted in the tones of the α -copies) but cannot account for SLQZ sentences where the bound copy shares all of the lexical features of its antecedent, but differs in case:

(38) R-yu'laàà' z-éng la'anng.
 HAB-like 3SG.PROX 3SG.PROX
 'He/she like himself/herself.'

complementarity between the environments that license α and those that license *nyas*, as well as between the functions of these two items.

However *nyas* is not the only Hmong pronoun that can function as a variable of this kind. Third person pronouns and zero anaphors can be bound as distributive variables of a sort. Take the following examples, where the dependent forms are in boldface:

- (39) a. [Tug twg]_i los **nwg**_i nyam nwg_i.
 CLF which TOP 3SG like α
 ‘Anybody would like themselves.’
 Any x : x would like x
- b. [Tug twg]_i los **nwg**_i nyam nwg*_{i/j}.
 CLF which TOP 3SG like 3SG
 ‘Anybody would like him.’
 Any x : x would like y
- c. [Tug twg]_i los nwg*_{i/j} nyam **nwg**_i.
 CLF which TOP 3SG like 3SG
 ‘He would like anybody.’
 Any x : y would like x

Example (39) gives three possible interpretations of the same (surface) sentence. This sentence features a common pattern where a Wh-expression is marked with the “topic” marker *los*. Wh-expression in indicative sentences induce “any” quantification. Such predicates are interpreted with irrealis modality, even when there is no overt marker of modality (as in the examples here).

That the relationship between the pronoun and the topic is actually dependent reference is shown by ellipses examples:

- (40) a. Tug twg los nwg nyam nwg; Maiv los kuj ua le hab.
 CLF which TOP 3SG like α May TOP also do as too
 ‘Anybody likes themselves; May does too (likes herself/*him/*her/*anyone).’
 (Any x : x would like x) \wedge (May : May would like May)

The availability of the sloppy reading in this case suggests that the first *nwg* is truly functioning as a variable, rather than referring to some person in the real world. Thus, the set {*Maiv*} may also be distributed over this variable, which binds the α variable, and thus licenses the reflexive reading in the elided clause.

A zero anaphor may replace *nwg* in either the subject or the object position:

- (41) a. [Tug twg]_i los **pro**_i nyam nwg*_{i/j}.
 CLF which TOP like 3SG
 ‘Anybody would like him.’
 Any x : x would like y

- b. [Tug twg]_i los nwg*_{i/j} nyam **pro**_i.
 CLF which TOP 3SG like
 ‘He would like anybody.’
 Any x : y would like x

The astute reader should immediately object to the analysis of the zero elements implied in (41). In these sentences, it is always the zero element that depends upon the topic for reference, so the null hypothesis should be that the zero anaphor identified as *pro* in the sentences is actually a gap left by the topicalization of *tug twg* ‘which person’. What justifies the *pro* analysis?

First, in sentences like (39), under the movement analysis, we would have to either posit that the topic was base-generated (unlike that in [41]) or would have to treat the dependent pronouns in (39) as resumptive pronouns. It is difficult to explain why such a resumptive pronoun can be inserted when the sentence is perfectly grammatical—and less ambiguous—without it.

Second, it is necessary to posit the existence of *pro* anyway, since the *pro* type of zero anaphora is very common in Hmong, with surprisingly few restrictions on its distribution and it is very common for *pro* to appear as a deictic pronoun that cannot reasonably be analyzed as the residue of movement⁷. In the same paradigm of sentences that we have been exploring, we can find the following possibilities:

- (42) a. ? [Tug twg]_i los **pro**_i nyam **pro**_i.
 CLF which TOP like α
 ‘Anybody would like themselves.’
 Any x : x would like x
- b. [Tug twg]_i los **pro**_i nyam *pro**_{i/j}.
 CLF which TOP like
 ‘Anybody would like him.’
 Any x : x would like y
- c. [Tug twg]_i los *pro**_{i/j} nyam **pro**_i.
 CLF which TOP like
 ‘He would like anybody.’
 Any x : y would like x

In both (42b) and (42c) there is one argument position that is not occupied by a variable. The parsimonious assumption, in these cases, is that the zero anaphor occupying these positions is *pro*. If we assume that *pro* is available in both the subject and object positions in this type of construction, there is no *a priori* factor excluding it from being the variable element in (41). Furthermore, this interpretation of the data has the benefit of capturing the parallelism between these sentences, those in (39), and the more surprising examples in (3).

The sentences in (3) reflect the same odd facts as those in (43):

⁷In fact, *pro* is the only pronoun aside from demonstratives that can be used to refer to most non-human entities. The only options in referring to such entities are to use *pro* or to employ a classifier with a demonstrative.

- (43) a. [Tug twg]_i los **nwg**_i nyam yawg.
 CLF which TOP 3SG like mister
 ‘Anybody would like the guy.’
- b. [Tug twg]_i los yawg nyam **nwg**_i
 CLF which TOP mister like 3SG
 ‘The guy would like anyone.’
- c. ?[Tug twg]_i los **yawg**_i nyam Maiv.
 CLF which TOP mister like May
 ‘Any guy would like May.’
- d. ?[Tug twg]_i los Maiv nyam **yawg**_i
 CLF which TOP May like mister
 ‘May would like any guy.’

Yawg is a kinship term used for older male relatives and can be used, pronominally, to refer to adult males. When *yawg* occupies one argument position and *nwg* occupies the other argument position, it is always *nwg* that is interpreted as a variable. That is, it seems that the relationship between *nwg* and *yawg* is comparable to that between the zero anaphor and *nwg*. To further the parallel, when *yawg* occurs in one argument position and something very referential, like a name, occurs in the other argument position, *yawg* is able to function as a variable (though the acceptability of these sentences is somewhat marginal).

3.1 Another scale

In other words, there seems to be a sort of scale in operation. The zero anaphor is the “best” variable of group, followed by third person pronouns, which are followed in turn by kinship pronominals. Names are not able to function as variables at all, and so come at the least-variable like end of the scale:

- (44) a. *pro* » *nwg* » *yawg* » *Maiv*
 b. *pro* » pronouns » kinship pronominals » names

The generalization is that a pronominal element cannot be interpreted as a variable unless there is no element in the sentence that is higher on the scale and that could be bound as a variable. The existence of this competitive scale complicates the movement analysis considerably. If the topic is never base generated, and if the variable reading is always due to movement, then the form the resumptive variable pronoun can take is constrained, but not directly determined, by the other elements in the sentence, according to this scale. In sentences like (43c) and (43d), the variable pronoun has features with which a potential referent must agree ([+masculine] and [+adult], for example) which are not specified in the topic. But we can capture the descriptive generalization directly if we assume that the topic and all of the pronominal elements, including the zero anaphors, are base generated.

The potential variable elements in a sentence can be computed directly from the overt form of the sentence, allowing at most one element to be bound as a variable by each operator, and ensuring that there is no better candidate for binding, based on the scale. The scale, of course, bears a remarkable resemblance to Safir's scales of dependence. However, the algorithm that licenses possible binding scenarios for these variables cannot be identical to the FTIP: the FTIP is only concerned with what could have taken the place of a potential dependent element, not with what other potential variables are present in the domain. Nevertheless, the algorithm that is needed is similar to the FTIP in that its outcome is determined competitively, rather than in terms of absolute structural conditions, as in classical binding theory.

3.2 An alternate proposal

The account I have offered of these binding phenomena has included some rather unattractive stipulations, especially concerning the domains for various types of binding. It would be advantageous to make these facts flow from more general principles, rather than having to stipulate them directly. One possible approach to this problem would be that advocated by Hornstein (2001), who argues that all construal relations can be reduced to movement relationships. Hornstein argues, for instance, that reflexive anaphors are the residue of A-movement, and that pronouns are expressions that are inserted where, without them, the derivation would not converge. In the spirit of this analysis, we might propose that α -reflexives are the unadulterated results of A-movement and that the A'-variables we have discussed more recently are the residue of A'-movement. Under such an analysis, many of the distributional properties of these two types of variables could be reduced to conditions on movement. Likewise, the difference in the form of the two types (copy of the antecedent versus pronouns/zero-anaphor) could be attributed to the kind of movement that left it behind.

However, as mentioned above, predicting the form that the residue will take is not as simple as it might appear. For example, under this analysis, we would have to treat locally bound *tug kheej* forms as a residue of the same type of movement as α -variables. This same problem appears cross-linguistically in the fact that some languages use α -type reflexives while other languages used dedicated reflexive forms. It is not immediately evident what principle should decide whether the copy left by movement is spelled out as a copy or as a reflexive anaphor. A formalism that simply evaluates the possible semantic relationships between base generated pronominal forms does not have to deal with this problem directly.

A related, but more complicated, problem is that of variables bound from topics and other A'-positions. If we want to treat the co-reference in these cases as a byproduct of movement, and refuse to treat the A'-expressions as base-generated, then we must explain why the resumptive pronouns left by this movement are able to vary in the way that they do. Either the form of this resumptive pronoun is constrained by the forms of deictic pronouns that have already been inserted, or the forms of the deictic pronouns that will be inserted are constrained by the form of this pronoun. In other words, this movement-based analysis would apparently have to, in some respect, duplicate the stipulative competitive algorithm described in my analysis.

4 Conclusion

At first blush, the type of copy-anaphora seen in Hmong seems to be the bane of theories of anaphora that attempt to derive the distribution of anaphors and other pronominal elements from scalar lexical competition. However, we have shown that this problem can be surmounted relatively easily if we assume that the bound element is—from the grammar’s point of view—a rather unexceptional anaphor that is spelled out to be the same lexical item as its antecedent. What proves more difficult to account for, in Hmong, is the apparent non-complementarity of α -anaphors and *tug kheej* forms and the distribution of dependent readings for A’-bound variables. However, while determining which elements in a sentence can be potentially A’-bound seems to require a different competitive algorithm than the FTIP, they provide additional evidence for the existence of scalar relationships among pronominal items that are strikingly similar to the dependency scales proposed by Safir.

The means by which binding domains are enforced in the account I have defended are probably more stipulative than is necessary, and it would be desirable to find a means of reducing these facts to some independently motivated principle. One possible means of doing this is to move towards an analysis in which dependent reference is wholly the result of movement. This solution obviates the need for α as an anaphoric lexical item, and does provide an elegant way of dealing with the asymmetry between A and A’-bound variables in Hmong. However, it raises some additional problems and does not seem able to directly account for the apparently competitive conditions on the forms of A’-bound variables.

References

- BISANG, WALTER 1993. “Classifiers, quantifiers, and class nouns in Hmong”. *Studies in Language* 17(1):1–51.
- BLACK, CHERYL A. 2000. *Quiégolani Zapotec Syntax*. Dallas, Texas: SIL and the University of Texas at Arlington.
- DÉCHÂINE, ROSE-MARIE AND MARTINA WILTSCHKO 2002. “Decomposing pronouns”. *Linguistic Inquiry* 33(3):409–442.
- HEIMBACH, ERNEST E. 1969. *White Meo–English Dictionary*. Ithaca: Cornell Southeast Asia Program.
- HORNSTEIN, NORBERT 2001. *Move! A Minimalist Theory of Construal*. Malden, MA: Blackwell.
- KOOPMAN, HILDA AND DOMINIQUE SPORTICHE 1989. “Pronouns, logical variables, and logophoricity in Abe”. *Linguistic Inquiry* 20(4):555–589.
- LASNIK, HOWARD 1986. “On the necessity of the binding conditions”. In Lasnik (1989), 149–167.
- 1989. *Essays on Anaphora*. Boston: Kluwer Academic.

- LEE, FELICIA 2002. "Anaphoric R-expressions as bound variables". In "Proceedings of the Berkeley Linguistics Society", 28. Berkeley, CA: Berkeley Linguistics Society.
- 2003. "Anaphoric R-Expressions as bound variables". *Syntax* 6(1):84–114.
- MARANTZ, ALEC 1982. "Re reduplication". *Linguistic Inquiry* 13:435–482.
- MCCARTHY, JOHN AND ALAN PRINCE 1995. "Faithfulness and reduplicative identity". Number 18 in University of Massachusetts Occasional Papers in Linguistics. Amherst, Mass.: Graduate Linguistic Student Association, 249–384.
- NARAHARA, TOMIKO 1995. "Alternatives to reflexives in Thai and Vietnamese: Binding theory and language variations". In Mark Alves (ed.), "Papers from the Third Annual Meeting of the Southeast Asian Linguistics Society", Tempe, Arizona: Arizona State University, Program for Southeast Asian Studies.
- RATLIFF, MARTHA 1992. *Meaningful Tone: A Study of Tonal Morphology in Compounds, Form Classes, and Expressive Phrases in White Hmong*. Number 27 in Monograph Series on Southeast Asia. DeKalb Illinois: Northern Illinois University Center for Southeast Asian Studies.
- REINHART, TANYA AND ERIC REULAND 1991. "Anaphors and logophors: An argument structure perspective". In J. Koster and E. Reuland (eds.), "Long Distance Anaphora", Cambridge: Cambridge University Press, 283–321.
- SAFIR, KENNETH J. 2002a. "The syntax of anaphora". Ms Rutgers University.
- 2002b. "The syntax of (in)dependence". Ms Rutgers University.