8

Two Approaches to Mayan Grammar Development in CCG

ELIAS PONVERT

8.1 Introduction

Computational grammar development often tends towards the development of lexemic grammars, i.e. grammars which posit little or no word-internal grammatical structure, or assume such morphological structure is handled by a system other than syntax proper. However, certain morphological processes do have significant syntactic consequences. Verb-internal incorporated pronouns in Mayan languages and others constitute one such phenomenon: as incorporated pronouns saturate one grammatical argument and semantic role, they are basically morphological affixes which alter a verb's subcategorization. This is a particularly acute issue for Categorial Grammar, as syntactic categories—or algebraic characterizations of words' and phrases' sub-categorization—occupy a central role in the theory. As such, incorporated pronouns constitute an appropriate case study in thinking about morphology in computational grammar development.

This paper makes two contributions. First, it presents a CCG for a fragment of the Mayan language Popti' that has been engineered using OpenCCG (Bozsahin et al., 2006). Popti' (Craig, 1977) is a configurational VSO language with an ergative/absolutive nominal system articulated by agreement markers and incorporated pronouns on the verb forms. Relative clause formation and focus constructions are both left-branching phenomena, and there are constraints on what constituents may be raised out of either construction. These constraints are closely related to the CCG slash-modalities introduced by Baldridge (2002) and presented in §8.1.1.

Texas Linguistics Society 10.
Nicholas Gaylord, Stephen Hilderbrand, Heeyoung Lyu, Alexis Palmer and Elias Ponvert eds.

Secondly, it presents an alternate, morphemic presentation of the grammar fragment that analyzes incorporated pronouns and ergative markers directly as syntactic constituents. This simplifies the categorial presentation of verbs in Popti' considerable, and simplifies and generalizes the semantics of the entities under consideration.

8.1.1 CCG

(Multi-modal) Combinatory Categorial Grammar (CCG) (Baldridge, 2002, Baldridge and Kruijff, 2003) is a mathematically constrained radically lexicalist grammatical formalism. In CCG, lexical items are assigned one or more categorial types which are formed from basic categories s, np, ... closed under the the directional slash operators $\{/,\backslash,...\}$. Slashes, in turn, are decorated by modalities $\{*,\diamond,\times,\cdot\}$ which dictate the applicability of the rules in the system to the category formed by that slash.

Syntactic analyses are given in the form of derivation trees, such as:

(8.1)
$$\dfrac{\dfrac{\textit{John}}{\text{np}} \quad \dfrac{\dfrac{\textit{kisses}}{(\text{s}\backslash\text{np})/\text{np}} \quad \dfrac{\textit{Mary}}{\text{np}}}{\text{s}\backslash\text{np}}>}{\text{s}}<$$

First of all *John* and *Mary* have the category np, i.e. they are noun phrases. *kisses* has the category $(\text{s}\backslash\text{np})/\text{np}$, which means that it is basically looking for a constituent with category np on its right to form a constituent of category $\text{s}\backslash\text{np}$, the equivalent of an intransitive verb or verb phrase. (There is no distinction between lexical and phrasal categories in Categorial Grammars.) This occurs with an instance of an Application rule, denoted $>$ in the derivation. The category $\text{s}\backslash\text{np}$, likewise, is looking for an np on the left to form a full sentence, i.e. a constituent with category s. This occurs with an instance of the corresponding Application rule $<$. The application rule schemata, then, are:

$$\begin{array}{rrcl} A/_*B & B & \Rightarrow_> & A \\ B & A\backslash_*B & \Rightarrow_< & A \end{array}$$

These rules by themselves would account for very little significant linguistic phenomena. The other rule schemata for CCG (at least, relevant to the current analysis) are Composition rules (B) and the Crossed Composition rules (B$\times$):

$$\begin{array}{rrcl} A/_\diamond B & B/_\diamond C & \Rightarrow_{>B} & A/_\diamond C \\ B\backslash_\diamond C & A\backslash_\diamond B & \Rightarrow_{<B} & A\backslash_\diamond C \\ A/_\times B & B\backslash_\times C & \Rightarrow_{>B_\times} & A\backslash_\times C \\ B/_\times C & A\backslash_\times B & \Rightarrow_{<B_\times} & A/_\times C \end{array}$$

See Steedman (2000) for extensive discussion.

Note that the slashes in the rule schemata above are decorated with the modalities $*$, $\diamond$ and $\times$. These control which rules may apply to what classes

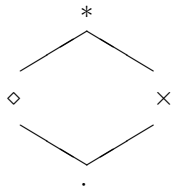


FIGURE 1 CCG Modality Hierarchy

of categories. The relationship between the modalities is hierarchical, as illustrated in Figure 1. Basically, categories whose slash is decorated with $*$ may only enter instances of Application, those decorated with $\diamond$ may enter into instances of Harmonic Composition or Application, those decorated with $\times$ may enter into instances of Crossed Composition or Application, and those decorated with $\cdot$ may enter into instances of any of the rules above. As a methodological priority, the fewest number of categories possible are assigned to a given lexical item. Also, for simplicity, the English example above (8.1) was not presented with slash-modalities. In general, the default modality is $\diamond$, i.e. the default assumption is that the Application and Harmonic Composition rules are applicable to most categories.

8.1.2 Grammar Engineering with CCG

The Popti' grammar fragment discussed in this article was implemented in OpenCCG. The grammar was developed using the CCG grammar specification language developed by Ben Wing[1]. The fragment is relatively small: 21 CCG categories were specified for 41 lexical items, plus 6 ergative morphemes, described below. The test bed comprises of 23 clauses, many of which are discussed below. The data and examples are taken from a linguistic grammar of Popti' (Craig, 1977).

8.2 Lexemic Analysis

This section provides a lexemic presentation of a grammar of a fragment of Popti'. In particular, it posits a unified syntactic account of relative clause formation and agent-focus constructions based on derivational control via slash-modalities.

[1] Cf. http://comp.ling.utexas.edu/ccggui/

8.2.1 Basic Word Order

Popti' features a rich set of determiner-like NP classifiers. Some examples are:

(8.2) *naj* *winaj*
cl man
the man

(8.3) *ix* *malin*
cl Mary
Mary

(8.4) *naj* *pel*
cl Peter
Peter

The generic term cl is used to gloss classifiers in what follows. Classifiers serve both as pronouns (of type np) and as determiners. A possible analysis of the determiner role could be given by a determiner category like np/n, that is a category which provides a noun phrase given a common noun (with category n). However as Popti' nouns are predicative (8.5), we opt instead for s/np for nouns and np/(s/np) for classifiers. This analysis of predicative nominals follows Baldridge (2002)'s of nominals in Tagalog.

(8.5) *winaj* *hach*
man you
You are a man

$$\dfrac{\dfrac{\textit{winaj}}{\text{s/np}} \quad \dfrac{\textit{hach}}{\text{np}}}{\text{s}}\ >$$

(8.6) *x'apni* *naj* *winaj*
arrived cl man
The man arrived

$$\dfrac{\dfrac{\textit{x'apni}}{\text{s/np}} \quad \dfrac{\dfrac{\textit{naj}}{\text{np/(s/np)}} \quad \dfrac{\textit{winaj}}{\text{s/np}}}{\text{np}}\ >}{\text{s}}\ >$$

Also, we put forward the straightforward right adjunct category s\s for the adverbial *ewi*, 'yesterday':

(8.7) *x'apni* *naj* *winaj* *ewi*
arrived cl man yesterday
The man arrived yesterday

$$\dfrac{\dfrac{\dfrac{\textit{x'apni}}{\text{s/np}} \quad \dfrac{\dfrac{\textit{naj}}{\text{np/(s/np)}} \quad \dfrac{\textit{winaj}}{\text{s/np}}}{\text{np}}}{\text{s}} \quad \dfrac{\textit{ewi}}{\text{s\s}}}{\text{s}}$$

Relative clauses are formed by appending the clause to the right of the relativized noun directly without a relativizer per se:

(8.8) *naj* *winaj* *x'apni*
cl man arrived
the man who arrived

(8.9) *naj* *winaj* *x'apni* *ewi*
cl man arrived yesterday
the man who arrived yesterday

These contrast with (8.6) and (8.7), from which we conjectured that *x'apni* has the type s/np. To accommodate both this and (8.8) we posit a type-changing rule of the form

(TC) $s/np \Rightarrow (s/np)\backslash(s/np)$

That is, a rule that changes a predicate into a noun-adjunct[2]. This rule is illustrated here with *naj winaj x'apni* 'The man who arrived':

$$\dfrac{\dfrac{naj}{np} \quad \dfrac{\dfrac{winaj}{(s/np)} \quad \dfrac{\dfrac{x'apni}{s/np}}{(s/np)\backslash(s/np)}\,\mathrm{TC}}{s/np}\,>}{np}\,>$$

However, the simple application rules $>, <$ will not allow the construction in (8.9) to go through, since there is no derivation for *x'apni ewi* as we have it so far ($***$ indicates that a derivation cannot go through):

$$\dfrac{\dfrac{x'apni}{s/np} \quad \dfrac{ewi}{s\backslash s}}{}\,***$$

The construction will go through, however, if we allow the Backwards Crossed Composition rules $B_\times$ to apply, which would suggest that *x'apni* is minimally $s/_\times np$ and *ewi* minimally $s\backslash_\times s$. Then, the following is a derivation for (8.9):

$$\dfrac{\dfrac{naj}{np/(s/np)} \quad \dfrac{\dfrac{winaj}{(s/np)} \quad \dfrac{\dfrac{x'apni}{s/_\times np} \quad \dfrac{ewi}{s\backslash_\times s}}{s/_\times np}\,<B_\times}{(s/np)\backslash(s/np)}\,\mathrm{TC}}{s/np}}{np}$$

[2]There is more than one way to do it: other analyses can handle this without the type-changing rule, but that does not alter the central argument here.

This provides an elegant distinction between predicative nominals (*winaj*) and intransitive verbs (*x'apni*):

(8.10) **Nominals** $s/_{\diamond}np$
Intransitive Verbs $s/_{\times}np$

One consequence of this analysis is that the adverb *ewi* can shift to the left of the subject in finite clauses, as in

(8.11) *x'apni* *ewi* *naj* *winaj*
arrived yesterday cl/the man
The man arrived yesterday

$$\dfrac{\dfrac{\dfrac{\textit{x'apni}}{s/_{\times}np}\quad\dfrac{\textit{ewi}}{s\backslash_{\times}s}}{s/_{\times}np}<B_{\times}\quad\dfrac{\textit{naj winaj}}{np}}{s}$$

Indeed, this prediction seems to be correct. (Nora England, p.c.).

8.2.2 The focus operator *ha'*

The focus operator *ha'* extracts either the object of transitive verbs or the subject of intransitive verbs (i.e. the absolutive constituents) from the VSO verbal nucleus. The categorial type for *ha'* is given by $(s/(s/_{\times}np))/np$, as in

(8.12) *ha'* *naj* *smak* *ix*
focus him hit she
It's him who she hit

$$\dfrac{\dfrac{\dfrac{\textit{ha'}}{(s/(s/_{\times}np_1))/np_1}\quad\dfrac{\textit{naj}}{np_1}}{s/(s/_{\times}np_1)}\quad\dfrac{\dfrac{\textit{smak}}{(s/_{\times}np_1)/np_2}\quad\dfrac{\textit{ix}}{np_2}}{s/_{\times}np_1}}{s}$$

It follows immediately from the categorial analysis of fronting, here, that only the subject of intransitives or the object of transitives may be focused, as indicated in the derivation for (8.12). In (8.12) np_1 must be the object and np_2 the subject of the clause. Also, the particular type provided for *ha'* prevents multiple instances of focus-extraction.

Popti' has a means to focus the subject of transitives, however. (8.13) highlights the anti-passive *-ni* form of *smak* from above. That is, there is a morphologically realized lexical rule that transforms *smak* to *xmakni* and that, in this analysis, basically switches the arguments of the verb. So:

(8.13) *ha'* *naj* *xmakni* *ix*
focus cl/he hit cl/her
It's he who hit her

$$\dfrac{\dfrac{\dfrac{\textit{ha'}}{(s/(s/_{\times}np_1))/np_1}\quad\dfrac{\textit{naj}}{np_1}}{s/(s/_{\times}np_1)}\quad\dfrac{\dfrac{\textit{xmakni}}{(s/_{\times}np_1)/np_2}\quad\dfrac{\textit{ix}}{np_2}}{s/_{\times}np_1}}{s}$$

This points to a strong generalization about Popti' verb categories. In general, intransitive and transitive verb categories share a core component of the form $s/_{\times}np$:

(8.14) **Intransitive Verbs** $s/_{\times}np_1$
Transitive Verbs $(s/_{\times}np_1)/_{\diamond}np_2$

Note that, given that Popti' is an ergative/absolutive language, np_1 must always correspond to the the absolutive argument.

8.2.3 Complements and quotations

Verbs of reporting, such as *xal* ("said") take a complement clause object, as in

(8.15) *xal* *naj* *jet-an* *tato* *x'apni* *ya'* *cumi*
said he to-us that arrived cl/the lady
He said to us that the lady arrived

This is handled straightforwardly with a new atomic category cp:

$$\dfrac{\dfrac{\dfrac{\textit{xal}}{(s/_{\times}cp)/_{\diamond}np}\quad\dfrac{\textit{naj}}{np}}{s/_{\times}cp}>\quad\dfrac{\dfrac{\textit{jet-an}}{s\backslash_{\times}s}}{s/_{\times}cp}<B_{\times}\quad\dfrac{\dfrac{\textit{tato}}{cp/_{\diamond}s}\quad\dfrac{\textit{x'apni ya' cumi}}{s}>}{cp}}{s}>$$

Under certain conditions, reporting verbs such as *xal* undergo a morphologically realized transformation into a quoting term, *yalni*, that accompanies quotative inversion. That is,

xal $(s/_{\times}cp)/_{\diamond}np \rightsquigarrow$ *yalni* $(s\backslash_{\times}s)/_{\diamond}np$

For example:

(8.16) *x'apni* *ya'* *cumi* *yalni* *naj* *jet-an*
arrived cl/the lady said he to-us
He said the lady arrived

$$\dfrac{\dfrac{\textit{x'apni ya' cumi}}{s}\quad\dfrac{\dfrac{\dfrac{\textit{yalni}}{(s\backslash_{\times}s)/_{\diamond}np}\quad\dfrac{\textit{naj}}{np}}{s\backslash_{\times}s}>}{s}<\quad\dfrac{\textit{jet-an}}{s\backslash_{\times}s}}{s}<$$

8.2.4 *mac* and long-distance extraction

The Wh pronoun *mac* ("who") occurs to the left of the verb and induces a question, roughly like English:

(8.17) *mac* *xul* *ewi*
who arrived yesterday
Who arrived yesterday?

$$\dfrac{\dfrac{\dfrac{\textit{mac}}{\mathrm{s}/_{\diamond}(\mathrm{s}/.\mathrm{np})} \quad \dfrac{\textit{xul}}{\mathrm{s}/_{\times}\mathrm{np}}}{\mathrm{s}}> \quad \dfrac{\textit{ewi}}{\mathrm{s}\backslash_{\times}\mathrm{s}}}{\mathrm{s}}$$

mac can enter into (somewhat) long-distance dependencies as in:

(8.18) *mac* *xawa'* *ha* *melyu* *tet*
who you-gave your money to
Who did you give your money to

$$\dfrac{\dfrac{\textit{mac}}{\mathrm{s}/_{\diamond}(\mathrm{s}/.\mathrm{np})} \quad \dfrac{\dfrac{\dfrac{\textit{xawa'}}{(\mathrm{s}/_{\diamond}\mathrm{pp})/_{\times}\mathrm{np}} \quad \dfrac{\dfrac{\textit{ha}}{\mathrm{np}/_{\diamond}(\mathrm{s}/.\mathrm{np})} \quad \dfrac{\textit{melyu}}{\mathrm{s}/_{\diamond}\mathrm{np}}}{\mathrm{np}}>}{\mathrm{s}/_{\diamond}\mathrm{pp}}> \quad \dfrac{\textit{tet}}{\mathrm{pp}/_{\diamond}\mathrm{np}}}{\mathrm{s}/_{\diamond}\mathrm{np}}>\mathrm{B}}{\mathrm{s}}>$$

However, combining *mac* with the reporting verb *xal* from §8.2.3 does not go through straightforwardly. Crucially, harmonic composition is blocked by the $\times$ modality on the verbal absolutive nucleus $\mathrm{s}/_{\times}\mathrm{np}$. Consider this attempt for "Who did Peter say hit Mary?":

(8.19) **mac* *xal* *naj* *pel* *chubil* *xmakni* *ix* *malin*
who said cl Peter that hit cl Mary
(attempted:) Who did Peter say hit Mary?

$$\dfrac{\dfrac{\textit{chubil}}{\mathrm{cp}/_{\diamond}\mathrm{s}} \quad \dfrac{\dfrac{\textit{xmakni}}{(\mathrm{s}/_{\times}\mathrm{np})/_{\diamond}\mathrm{np}} \quad \dfrac{\textit{ix malin}}{\mathrm{np}}}{\mathrm{s}/_{\times}\mathrm{np}}>}{***}$$

(The *** indicates that there is no rule to carry the derivation forward.) In Craig (1977)'s treatment, (8.19) is marked with a (?). This may be because, since *xmakni* is a subject-inverted pseudo-passive, some speakers may treat it without the $\times$ modal verbal nucleus. Nevertheless, using the quotative construction produces a clean reading:

(8.20) *mac* *xmakni* *ix* *malin* *yalni* *naj* *pel*
who hit cl Mary say cl Peter

Who did Peter say hit Mary?

$$\dfrac{\dfrac{\textit{mac}}{\mathrm{s}/_{\diamond}(\mathrm{s}/_{\diamond}\mathrm{np})} \quad \dfrac{\dfrac{\dfrac{\textit{xmakni}}{(\mathrm{s}/_{\times}\mathrm{np})/_{\diamond}\mathrm{np}} \quad \dfrac{\textit{ix malin}}{\mathrm{np}}}{\mathrm{s}/_{\times}\mathrm{np}}> \quad \dfrac{\dfrac{\textit{yalni}}{(\mathrm{s}\backslash_{\times}\mathrm{s})/_{\diamond}\,\mathrm{np}} \quad \dfrac{\textit{naj pel}}{\mathrm{np}}}{\mathrm{s}\backslash_{\times}\mathrm{s}}>}{\mathrm{s}/_{\times}\mathrm{np}}<\mathrm{B}_{\times}}{\mathrm{s}}>$$

In fact, the grammar implementation gets both readings, as in

- "Who did Peter say hit Mary?"
- "Peter said 'Who hit Mary?"'

8.2.5 Summary

Summing up the analysis so far, we have to following general category assignments:

	Nominals	*winnaj*	$\mathrm{s}/_{\diamond}\mathrm{np}$
	Classifiers	*ix*, *naj*	$\mathrm{np}/_{\diamond}(\mathrm{s}/_{\diamond}\mathrm{np})$
	Intrans. Verbs	*x'apni*	$\mathrm{s}/_{\times}\mathrm{np}$
	Transitive Vbs	*smak*	$(\mathrm{s}/_{\times}\mathrm{np})/_{\diamond}\mathrm{np}$
(8.21)	**Temporal Adv.**	*ewi*	$\mathrm{s}\backslash_{\times}\mathrm{s}$
	Focus	*ha'*	$(\mathrm{s}/_{\diamond}(\mathrm{s}/_{\times}\mathrm{np}_1))/_{\diamond}\mathrm{np}_1$
	Complement Clause	*tato*	$\mathrm{cp}/_{\diamond}\mathrm{s}$
	Reporting Verbs	*xal*	$(\mathrm{s}/_{\times}\mathrm{cp})/_{\diamond}\mathrm{np}$
	Quotative	*yalni*	$(\mathrm{s}\backslash_{\times}\mathrm{s})/_{\diamond}\mathrm{np}$

8.3 Issues with the Lexemic Analysis

A lexemic approach such as is posited in §8.2 poses several methodological issues. In practice, for a language with significant morphology, the grammar writer must generally specify a larger number of lexical items, and with reduplicated effort comes the increased likelihood of error.

Incorporated pronouns in languages such as Popti' provide an acute test case of this. Popti' has two sets of incorporated pronouns, for each the absolutive and ergative verbal arguments. To accommodate this, the lexemic grammar written for the fragment of Popti' alluded to above must specify the redundant Transitive Verb category: ones highlighted in (8.21):

(8.22) • SubjectEmbeddedTransitiveVerb $\mathrm{s}/_{\times}\mathrm{np}$

Moreover, aspectual markers and the subject-focus markers were embedded in the lexical items, leading to lexical entries akin to

```
hit: TransitiveVerb { smak; }
hit-subject-focus: TransitiveVerb { xmakni; }
```

In fact, morphologically *smak* is a fairly complex form:

(8.23) *x-* *s-* *mak*
asp erg hit

where asp is completive aspect and erg is the 3rd person ergative marker. A more general morphemic system that can also handle word-internal structure is proposed in the following section.

8.4 Morphemic Approach

A morphemic approach (Bozsahin, 2002) differs significantly from a lexemic approach in that it posits categorial types for word-internal morphemes and combinatory processes for word-formation akin to the syntactic processes themselves. Here, the analysis does not in fact distinguish between morphological and syntactic categories and processes.

To enable this in part, the analysis below makes use of *inert slashes* $/^!, \backslash^!$. A category specified with an inert slash, say, $s/^!np$, may not combine its argument (here, np) as usual, may be regarded as an atomic category decorated with structural and semantic information.

The morphemes under consideration, then, are:

Aspect Aspect markers combine with a non-finite, inert VP to the right to produce a finite, aspectually specified VP. Here is the category for the completive aspect marker:

- *x-*: $(s/_\times np_1)/_\diamond(s/^!_\times np_1)$

Ergative (incorporated) pronouns Incorporated pronouns consume the outer (first) argument of a transitive verb. Leaving out the details, their category is $(s\ /^!_\times n\ p)\ /_*\ ((s\ /^!_\times\ np)\ /^!_\diamond\ np)$. That is, they take a non-finite transitive verb to the immediate right and render an intransitive verb. Specifically:

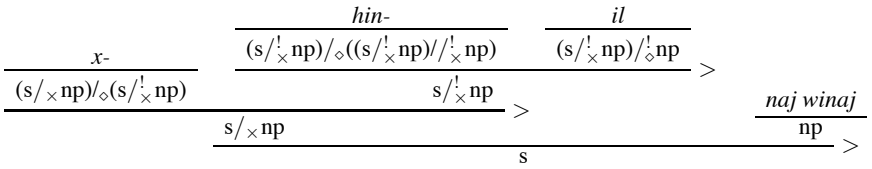
- *hin-* 'I', *ha-* 'you': $(s/^!_\times np)/_*((s^!_\times np)/_\diamond np)$

Transitive Verbs Transitive verbs, without aspect, agreement or embedded pronouns are simply assigned a category identical to that specified before, only now with inert slashes:

- *il* 'to see', *mac* 'to hit': $(s/^!_\times np)/^!_\diamond np$

The aspect and incorporated pronouns enable derivations of sentences that exhibit the shifting arity of verbs at work. In the following is example, the transitive verb *il* 'to see', is rendered as a pseudo-intransitive *x-hin-il*, 'I saw ...':

(8.24) *x-* *hin-* *il* *naj* *winaj*
asp pro see cl man
I saw the man



Ergative agreement Ergative agreement takes a transitive verb to the immediate right and basically "turns-on" the verb's ergative argument:

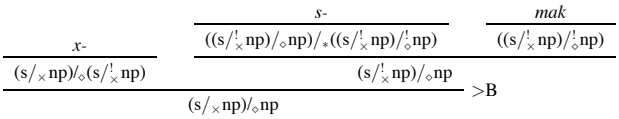
• *s-*: $((s /^{!}_{\times} np_1) /_{\diamond} np_2) /_{*} ((s /^{!}_{\times} np_1) /^{!}_{\diamond} np_2)$

This category also checks that the first NP in its delcarative sentence is of ergative voice.

By contrast with incorporated pronouns, the 3rd person ergative marker does not change the arity of a transitive verb, as in:

(8.25) *x- s- mak ix malin naj pel*
asp erg hit cl Mary cl Peter
Mary hit Peter

As in this and example (8.2.2) above, the desired category for *x-s-mak* is $(s/_{\times}np)/_{\diamond}np$. This is given by:



Subject focus The anti-passive or subject-focus morpheme *-ni* takes a transitive verb to the immediate left and, in effect, switches the order of the arguments, so as to make the absolutive argument to the verb accessible to the focus operator *ha'*, as sketched in §8.2.2:

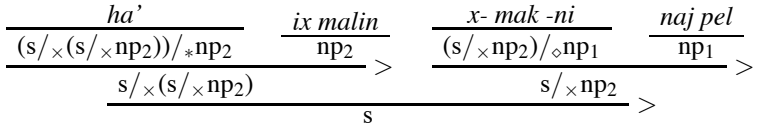
• *-ni*: $((s /^{!}_{\times} np^2) /_{\diamond} np^1) /_{*} ((s /^{!}_{\times} np^1) /^{!}_{\diamond} np^2)$

Note: np^1 is shorthand for $\left[np\ \langle 1 \rangle\right]$.

With this suffix we can now specify a full analysis of an instance of subject focus with the focus operator *ha'*:

(8.26) *ha' ix malin x- mak -ni naj pel*
foc cl Mary asp hit subj-foc cl Peter
It was Mary who hit Peter

The basic derivation is:



The indices indicate the reversal of the verbal arguments, specific to the analysis of subject extraction above. The complex verbal form *x-mac-ni* together with the derivation reversing its arguments is given here:

$$\dfrac{\dfrac{\textit{x-}}{(s/_{\times}np_2)/_{\diamond}(s/^{!}_{\times}np_2)} \quad \dfrac{\dfrac{\textit{mak}}{(s/^{!}_{\times}np_1)/^{!}_{\diamond}np_2} \quad \dfrac{\textit{-ni}}{((s\,/^{!}_{\times}\,np_2)\,/_{\diamond}\,np_1)\backslash_{*}((s\,/^{!}_{\times}\,np_1)\,/^{!}_{\diamond}\,np_2)}<}{(s/^{!}_{\times}np_2)/_{\diamond}np_1}}{(s/_{\times}np_2)/_{\diamond}np_1}\,{>}\mathrm{B}$$

8.4.1 Results

The morphemic analysis presented here allows for a simplification of how verbal forms are specified in general. Few lexical entries are required for verb forms, and fewer lexical families are required.

The morphemic grammar parses, produces semantics and is able to realize all the interpretations of the sentences the lexemic grammar was. One construction, involving gapping ('John ate a mango and Mary an orange' in Popti'), eludes both grammars. And in fact, whereas the lexemic grammar was able to parse the focus constructions involving *ha'* (§8.2.2), but not realize the original form from the semantics, the smaller and simpler morphemic grammar improved on the analysis of *ha'* and produce the right realizations.

8.5 Conclusions

The morphemic grammar for Popti' has several nice properties with respect to the lexemic alternative, especially with respect to the depth, generality and compositional consistency of its semantic analyses. However, what it gains in generality and elegance, it loses in efficiency in parsing, as it has more work to do parsing the same material. Perhaps an integrated strategy, the makes reference to a morphemic system at "compile-time" to capture the desired generalities where relevant, but does actual processing and realization with a purely lexemic system, could take advantage of both techniques.

References

Baldridge, Jason. 2002. *Lexically Specified Derivational Control in Combinatory Categorial Grammar*. Ph.D. thesis, University of Edinburgh.

Baldridge, Jason and Geert-Jan M. Kruijff. 2003. Multi-modal categorial grammar. In *Proceedings of EACL*.

Bozsahin, Cem. 2002. The combinatory morphemic lexicon. *Computational Linguistics* 28:145–186.

Bozsahin, Cem, Geert-Jan Kruijff, and Michael White. 2006. *Specifying Grammars for OpenCCG: A Rough Guide*.

Craig, Colette Grinevald. 1977. *The Structure of Jacaltec*. Austin: UT Press.

Steedman, Mark. 2000. *The Syntactic Process*. Cambridge, Mass.: MIT.