

## The Two Types of Predicate Clefts: Classical Hebrew and Beyond

Daniel Harbour\*  
MIT

Predicate cleft constructions, that is, sentences containing a left-dislocated copy of their predicate, occur in a wide variety of unrelated languages and display a number of intriguing systematic differences. For example, the distance between the cleft form and the 'base' form, the number and type of arguments and adjuncts that may appear with the cleft form, and the positions in which the cleft may occur all vary across languages. However, there appears to be a correlation between the distance between cleft and 'base' form and the number of adjuncts and arguments with which the cleft form may appear: the greater the distance, the greater the number. Work on predicate clefts (Koopman (1983), Lefèbvre (1992), *inter alia*) has therefore generally sought to answer two questions: what is the nature of the predicate cleft construction? What accounts for the cross-linguistic variation? The current squib aims to contribute to this discussion in two ways: by bringing new data to the debate and by using that data to suggest a factor that may be fundamental to the predicate cleft typology.

Classical Hebrew is a language rich in verbal morphology in which the predicate cleft is a well attested construction. On the basis of its morphology and syntax, I argue that the cleft form is a head and that the predicate cleft construction is derived by movement. Supposing the cleft form to be a head accounts for the distributional facts: that it occurs with adjuncts and without arguments, that it can occur in positions through which the verb must have passed on its way into TP, that, being a case of head movement, it is very local. The cross-linguistic component comes when one asks why the cleft form need be a head. One can imagine a language which differed from Hebrew in that its cleft form was not a head but a phrase. Given that the properties of the predicate cleft in Classical Hebrew can be argued to follow from the cleft form's being a head, we could then ask which properties of a predicate clefts in 'phrasal clefting language' would be the same as those of Classical Hebrew and which not. It is easy to predict some differences between two such languages. Moreover, some of these difference hold cross-linguistically. On this basis, I suggest that the difference between head and phrase clefting languages might be fundamental to a typology of the predicate cleft. We return to these comments after discussion

---

\*I am grateful to Ken Hale (practical contributions), Sabine Iatridou (pertinent considerations), Idan Landau (productive chats), Victor Manfredi (passing comments), Wayne O'Neil (precise criticism), David Pesetsky (professional curiosity). The editors' comments have given me more ideas than I have been able incorporate here. Use of Tamar and Abraham Eisenstat's library has been invaluable.

of the Classical Hebrew data, at which point, they will, hopefully, be less abstract and more accessible.

### 1. Basic Facts about Classical Hebrew

Classical Hebrew is a VSO language in which a focused or topicalised element may occur preverbally. A typical sentence is given in (1).<sup>1</sup>

- (1) bre'šitō bārā' 'ālohiym 'eθ haššāmayim w'eθ hā'ārās  
 in-beginning created God ACC the-heaven and-ACC the-earth  
 Gn1': 'In the beginning, God created the heaven and the earth'

Like all Semitic languages, Classical Hebrew has a series of *binyanim* or verb classes. The most common are given in (2) with their basic meanings.

- (2) 1. pā'al basic form of the active verb  
 2. nif'al reciprocal, basic form of some unergatives, common passive of pā'al  
 3. pi'el emphatic, causative, basic active form of some roots  
 4. pu'al passive of pi'el  
 5. hif'iyl causative  
 6. hof'al passive of hif'iyl  
 7. hiθpa'el reflexive

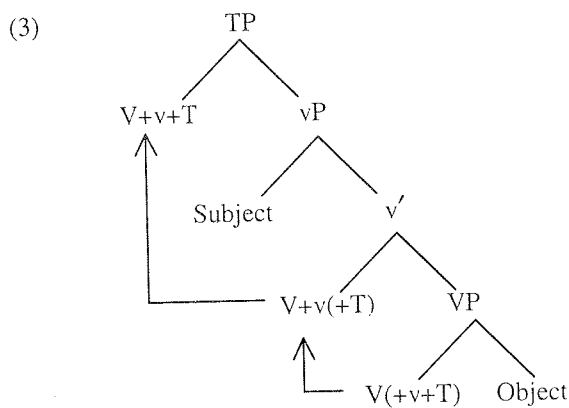
In glosses, the subscripted number gives the binyan. Several binyanim have allied forms. For instance, *hiθpolel* is allied to *hiθpa'el*. As the allied binyanim raise no issues relevant here, they are represented by 1–7, as appropriate.

I suppose that Classical Hebrew is underlyingly SVO and that V raises to T in the overt syntax. S generally precedes O, when both are overt. O may precede, but only under slightly special circumstances: OVS occurs where the object is prominent in the discourse and VOS, where O is pronominal. The normal narrative order is VSO. Therefore, it is likely that S precedes O in the underlying structure. Assuming VO to be an underlying constituent, the question is whether Classical Hebrew is underlyingly SVO or SOV. However, in all infinitival clauses and gerunds, the word order is VO. This can be accounted for assuming either SVO or SOV to be underlying. However, SOV requires the additional assumption that V obligatorily moves in such constructions. For reasons of simplicity, then, I assume SVO to be underlying.

<sup>1</sup>A note on verse numbers and transcriptions. References for Classical Hebrew are given as BookChapter<sup>Verse</sup>, with books abbreviated as follows: *Ecc* = Ecclesiastes, *Ex* = Exodus, *Ez* = Ezekiel, *Gn* = Genesis, *Is* = Isaiah, *Jrm* = Jeremiah, *Ju* = Judges, *1/2K* = Kings I/II, *Lev* = Leviticus, *Nu* = Numbers, *Pro* = Proverbs, *1/2S* = Samuel I/II, *Song* = The Song of Solomon, *Thr* = Threnode, *Zac* = Zachariah. All consonants in the Hebrew are rendered in the transcription, even if not pronounced (e.g.: some instances of *y*, *h*, ' ). When *h* is pronounced syllable finally, it is transcribed as *h̄*. The pharyngeal fricatives are ' (voiced) and *h̄*. Pharyngealisation is indicated by a superscripted ' , as in *s'*. The glottal stop is ' . The voiceless palatal fricative is *š*. Schwa is never transcribed.

## Predicate Clefts

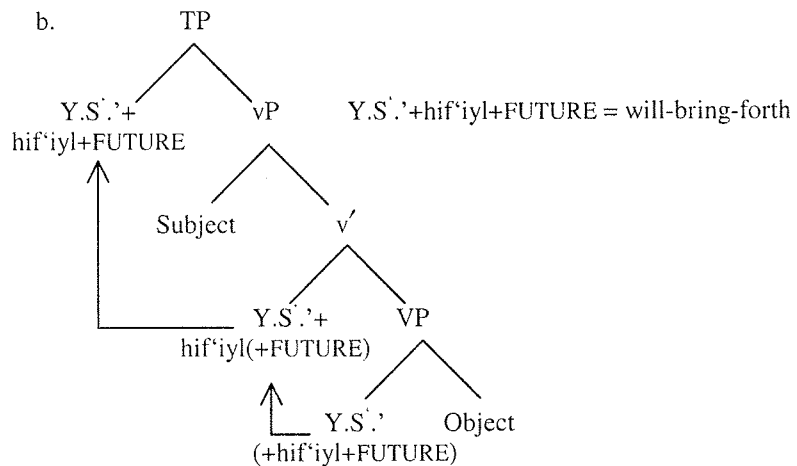
In the derivation of a typical VSO sentence, I propose the following stages. The bare verbal root is merged as the head of VP. It then raises to  $v'$ , where it picks up binyan. (I will refer to this as 'checking binyan'. The reader is referred to Arad (1999) for arguments connecting binyan to  $v$ .) The inflected verb then raises higher until it reaches T where it checks tense features. So, a typical VSO sentence would look something like (3), omitting various functional projections and the lexical items. The features in parentheses have not been checked: so, 'V+v(+T)' refers to the fully inflected verb at the stage when it has checked binyan but not tense.



So, for a typical VSO sentence, such as that in (4a), we would have a structure such as that shown in (4b).

- (4) a.   wattos'e'           hā'ārās   dāšā'  
 &-will-bring-forth, the-earth grass  
 Gn1<sup>2</sup>: 'And the earth brought forth grass'<sup>2</sup>

<sup>2</sup>The gloss 'will bring forth' is no error. Curiously, in Classical Hebrew, when *and* combines with a verb, the meaning of the morphology switches: past means future and future, past. For example, *and + she'll-bring-forth* (future) means *and she brought forth* (past). This is the only environment in which the language has preserved the Proto-Semitic forms for the past/perfective and future/imperfective, compared with, for instance, Arabic or Akkadian. In all other environments, the morphology has the reverse semantics. The phenomenon of the 'wāw conversum' ('wāw' is the name of the letter *w* which prefixed in various allophonic guises means "and") is of some relevance to the predicate cleft. Since wāw conversum must always be clause initial and must be followed immediately by the verb (or else it is not a wāw conversum), it always results in the order 'wāw conversum V Λ'. More commonly, however, one finds 'wāw copulativum Λ V'. Note that the wāw conversum induces a phonological change from *tos'iy'* to *tos'e'*. Note also the occurrence of Everyone's Favourite Segholate.



The root Y.S.' is merged as the head of VP. Leaving a copy in V<sup>0</sup>, it raises to v' where it checks binyan (*hif'iyl*). Then it continues to raise, via head-to-head movement not shown in (4b), until it reaches T<sup>0</sup>, where it checks tense features. The lower copies then delete, resulting in the VSO sentence in (4a).

The basic properties of the predicate cleft are as follows. It serves to emphasise the predicate by 'repeating' a reduced version of it. In its most usual form, it consists of a VSO sentence with a preverbal cleft form. I shall denote this form by 'Λ'. In examples, Λ is written **bold** and normal tensed V's are *italicised*. An example of ΛVSO is given in (5).

- (5) **šāxov** *yīškav* 'iyš 'oθāh  
**lie<sub>1</sub>** *he-will-lie<sub>2</sub>*, man ACC-her  
 Lev15<sup>24</sup>: '[If] a man lie with her at all'

Λ need not be sentence-initial; for example, nearly all of Classical Hebrew's complementisers are attested before Λ. The most common are 'if' (*'im*) and 'because' (*kiy*). Others include rarer forms of 'if' (*lu'*), the interrogative (*ha/halo'/hā*), the conjunction (*w*), the exclamatory (*'ax*).

- (6) 'im-**himmāse'** *θimmāse'* vyāðo haggnevāh ... šnayim yšallem  
 if-**find<sub>1</sub>** *will-be-found<sub>2</sub>* in-hand-his the-theft ... two he-will-pay<sub>3</sub>  
 Ex22<sup>3</sup>: 'If the theft be certainly found in his hand ..., he shall restore it  
 double'

Full XPs are also attested and may be a subject or object DP, a PP, or a CP. The example shows the first of these.

- (7) kāl-'omsāyhā **šārot'** *yīššāret'u*  
 all-burden<sub>1</sub>-PART.-PL.-her **cut<sub>1</sub>** *they-will-be-cut<sub>2</sub>*  
 Zac12<sup>3</sup>: 'All that burden themselves with it shall be cut into pieces'

However, nothing may occur between  $\Lambda$  and V except negation (*lo'*). I assume NegP to be lower than TP, as appears to be the case cross-linguistically. As negation occurs preverbally (8a), the negative particle presumably prefixes to the verb as it raises to T<sup>0</sup>. The resulting word order is  $\Lambda$  NEG-V (8b).

- (8) a.    mayim rabbiym lo' yuxlu    lxabboθ    'äθ-hå'ahavåh  
           waters many not will-be-able<sub>1</sub> to-quench<sub>3</sub> ACC-the-love  
           Song8<sup>7</sup>: 'Many waters cannot quench love'
- b.    wayyåsem 'äθ-hakkna'aniy lāmås  
           &-he'll-put<sub>1</sub> ACC-the-Canaanite to-the-tribute  
           w'horeyš                            lo'-horiyšo  
           and-ethnically-cleanse<sub>3</sub> not-ethnically-cleanse<sub>3</sub>-him  
           Jul<sup>28</sup>: 'And [Israel] put the Canaanites to tribute, and did not utterly  
           drive them out'

(The meaning of  $\Lambda$  NEG-V is approximately, e.g., 'drive out, they didn't drive them out utterly' or 'drive them out, they didn't utterly'.) Given that negation is a prefix collected by the raising verb and only it may intervene between V and a preverbal  $\Lambda$ , we may conclude that  $\Lambda$  is in the nearest position above V.

As I shall argue that  $\Lambda$  is a head, this position must be a head position. However, before we can consider in detail the syntax of the predicate cleft, it is necessary first to understand what the cleft form is. This question concerns the...

## 2. Morphology of the Cleft Form

On the basis of morphology,  $\Lambda$ , the cleft form, can be argued to be a head. An alternative explanation, that  $\Lambda$  is a cognate noun, can be discounted on the same basis. Further evidence for this view comes from syntactic properties of the predicate cleft construction, presented in the next section. An interesting connection between the morphology and the syntax arises when we examine the ways in which the binyan of  $\Lambda$  may differ from that of V. The relationship is examined in this section.

A fully inflected verb consists of a root, a binyan, tense information, and possibly clitics.  $\Lambda$ , by contrast, consists of a root and binyan only. (Alternatively, it can be said to consist of a root with binyan optionally, if one views *på'al* as supplied as the default binyan. For more on this point, see the discussion of how the binyan of the  $\Lambda$  may vary from that of V.) There is no binyan in which  $\Lambda$  cannot occur, its form is not dependent on the tense of the verb, and it never bears clitics. Any clitics occur 'on' the finite verb, V, that is, they are phonological dependents of the verb. (Syntactically, they may head their own projections, as has been argued to be the case cross-linguistically). A typical example is shown in (9).

- (9)    'åsor ya-'asr-u-niy  
       bind<sub>1</sub> will-bind<sub>1</sub>-they-me  
       Jul<sup>15</sup>: 'They will bind me fast'

The minimal hypothesis is that  $\Lambda$  is a 'verb' lacking tense information and incapable of bearing clitics. In that case, it is as head just as much as *ya-'asr-u* (they will bind) is.

An alternative explanation is that  $\Lambda$  is a cognate noun, on a par with the English *to dance a dance, to sleep the sleep of the innocent*. Such an explanation suggests itself immediately to Semitists, as cognate object constructions are very common across the family. However, there are at least three reasons not to regard  $\Lambda$  as one. Cognate nouns behave quite differently from cleft forms. Only the former can host possessive clitics and can also be preceded by the accusative marker *'äθ-/eθ*. Moreover, there are sentences that contain a verb, a cleft, and an etymologically related object.

- (10) a. **s'ānof** *yis nāfxā*      s'nefāh kaddur 'äl-'ärās' rahavaθ yāðāyim  
**turn**<sub>1</sub> *he-will-turn, you* turn ball to-land width two-hands  
 Is22<sup>18</sup>: 'He will surely violently turn and toss thee like a ball into a large country'
- b. **riyv** *yāriyv*      'äθ-riyvām  
**plead**<sub>1</sub> *he-will-plead,* ACC-cause-their  
 Jrm50<sup>34</sup>: 'He shall thoroughly plead their cause'

Furthermore,  $\Lambda$  is often distinct from cognate nouns. As an example, consider the root S'.Q., which, in *pā'al*, means 'cry'. Its cleft form, *s'ā'oq*, in (11a), is clearly different from the cognate noun, *s'āqāh*, in (11b).

- (11) a. **s'ā'oq** *yis 'aq*  
**cry**<sub>1</sub> *he-will-cry,*  
 Ex22<sup>22</sup>: '[If] he cry at all'
- b. wayyis 'aq      s'āqāh gðolāh umārāh 'að-m'oð  
 and-he-will-cry, cry great and-bitter excessively  
 Gn27<sup>34</sup>: 'And [Esau] cried a great and excessively bitter cry'

Secondly, the cleft form has other uses, such as naming an activity.

- (12) 'āxol dvaš harboθ lo'-tov  
 eat<sub>1</sub> honey much not-good  
 Pro25<sup>27</sup>: '[It is] not good to eat too much honey'

In these contexts, the cleft form seems ambiguous between an infinitive and a gerund. That is, *'āxol* in (12) can be translated as 'to eat' or 'eating'. If it were a gerund, one would expect it to be usable in construct nominals, like any other non-pronominal nominal. However, a different form is used in such contexts.

- (13) 'eθ s(\*ā)foð w'eθ r(\*ā)qoð  
 time mourn, and-time dance,  
 Ecc3<sup>4</sup>: 'A time to mourn and a time to dance'

## Predicate Clefts

(The asterisks mean that the form is unattested in this context.) Likewise, it is the form in (13), not the cleft form, that takes possessive clitics. Lastly, cognate objects share the same binyan as the verb they are paired to. Cleft forms are not constrained in that way (see the next paragraphs). I conclude that  $\Lambda$  is some sort of verbal head, not a nominalisation.

Whilst on the morphology, it is important to understand as how the binyan of  $\Lambda$  related to the binyan of V. In most of the examples given so far, both  $\Lambda$  and V have been in *pā'al*. However, as mentioned, other combinations are possible and give an important clue to the syntactic properties and derivation of the predicate cleft construction. The main cooccurrence possibilities are given in table (14). The binyan of  $\Lambda$  is given on the top row, that of V, in the left hand column.

(14)

		$\Lambda$						
		pā'al	nif'al	pi'el	pu'al	hif'iyl	hof'al	hiθpa'el
V	Pā'al	✓						
	Nif'al	✓	✓					
	Pi'el	✓		✓				
	Pu'al	(✓)			✓			
	Hif'iyl	✓				✓		
	Hof'al	✓					✓	
	Hiθpa'el	✓						✓

Several comments are required at this point. The first concerns the semantics. There is no obvious difference between having  $\Lambda$  in the same binyan as V and having it in *pā'al*. (*Pā'al* can be regarded as the default binyan as it is the most common and as its 3<sup>rd</sup> masculine singular past involves no alterations to the root, other than such general as processes spirantisation.) One finds such minimal pairs as (15).

- (15) a. w'attāh hu' nāqoh tinnāqāh lo' θinnāqāh  
 and-thou he *be-clean<sub>1</sub>*, *you-will-be-clean<sub>1</sub>*, not you-will-be-clean,  
 Jrm49<sup>12</sup>: 'And art thou he that shall altogether go unpunished?  
 Thou shalt not go unpunished'
- b. w'attem hinnāqeh θinnāqāh lo' θinnāqu  
 and-you *be-clean<sub>2</sub>*, *you-will-be-clean<sub>2</sub>*, not you-will-be-clean,  
 Jrm25<sup>29</sup>: 'And should ye be utterly unpunished? Ye shall not be  
 unpunished'

Though the King James translation obscures it somewhat, they are clearly very similar. Other examples point in the same direction, such as passages containing many predicate clefts where the verbs are in *nif'al* but where  $\Lambda$  varies between *pā'al* and *nif'al* without any evident difference in meaning.

The obvious generalisation to be drawn from (14) is that  $\Lambda$  must have the same binyan as V or else must be *pā'al*. However, there are rare combinations not mentioned in the table. Some are regarded by scholars as scribal errors. Others are more interesting. For instance, in Ez16<sup>4</sup>, we find

<  $\Lambda$ , V > = < hof'al, pu'al >. These two conjugations are semantically similar, both common as passives of causatives. A possible explanation for the divergence, then, is that both  $\Lambda$  and V are identically specified in the syntax, modulo tense and agreement information, but are treated differently by the morphophonology. Like comments are applicable to 1S2<sup>16</sup>, where we find <  $\Lambda$ , V > = < pi'el, hif'iyl >, both these binyanim being common as causatives; and for Lev19<sup>20</sup>, where <  $\Lambda$ , V > = < hof'al, nif'al >, both binyanim having a passive-of-causative sense.

- (16) a. molðoθayix... **hāmleah** lo' *humlaḥat*                      **whāhttel**  
 nativity-your... **salt**<sub>6</sub>      not *you-will-be-salted*<sub>6</sub>      and-**swaddle**<sub>6</sub>  
                                  lo' *huttālt*  
                                  not *you-were-swaddled*<sub>6</sub>  
 Ez16<sup>4</sup>: 'As for thy nativity, ... thou wast not salted at all, nor swaddled at all'
- b. **qat'er** *yaq'i iyrun*                      kayyom *haḥelāv*  
**burn**<sub>3</sub>      *let-them-burn*<sub>3</sub>      presently the-fat  
 1S2<sup>16</sup>: 'Let them not fail to burn the fat presently'
- c. **hāfdeh**      lo' *nifdāθāh*  
**redeemed**<sub>6</sub>      not *she-was-redeemed*<sub>2</sub>  
 Lev19<sup>20</sup>: '[A bondsmaid who is] not at all redeemed'

(Note that the cleft forms and the first verb in (16a) are all in the same binyan. Some kind of stylistic parallelism may, therefore, be involved in the binyan-mismatch between *hāhttel* and *huttālt*. However, this would not affect the basic point that the mismatch is possible owing to similarity of the binyanim: passive and causative are expressed by both binyamin.)

The appeal to similarity of the semantics of the binyanim in order to explain the mismatches suggests a reason why a *pā'al*  $\Lambda$  may occur with V of any binyan. Suppose that the syntax requires only that  $\Lambda$  and V be 'compatible', perhaps in the sense of having non-conflicting features. Suppose also that V must contain all semantically relevant information: i.e., root, binyan, tense, agreement, clitics. We have observed that  $\Lambda$  may have, at most, root and binyan information. So, it seems that  $\Lambda$  has a subset of the information that V bears. As a limiting case, therefore,  $\Lambda$  could be a bare root that receives default morphology. As *pā'al* is the most common and semantically neutral binyan, it is reasonable to suppose that it would be the default. If the notion of 'feature containment' between  $\Lambda$  and V is approximately correct, then we would expect to see *pā'al*  $\Lambda$  paired with V of any binyan.

*Pu'al* provides interesting evidence that forces the idea of feature containment to be made more precise. Since it is the passive of *pi'el*, one could expect to find < pi'el, pu'al > pairs. Likewise, one could expect to find the bare root alongside a *pu'al*, that is, < pā'al, pu'al >; and, of course, < pu'al, pu'al >. In actual fact, < pi'el, pu'al > is unattested. So, (17a) is a straightforward instance of < pu'al, pu'al >, and (17b), an instance < pā'al, pu'al >.



### Predicate Clefts

- (17) a. **gunnov** *gunnavtiy* me'ārās' hā'ivriym  
**steal-away**<sub>4</sub> *I-was-stolen-away*<sub>4</sub> from-land the-Hebrews  
 Gn40<sup>15</sup>: 'Indeed I was stolen away out of the land of the Hebrews'
- b. **t'ārof** *t'orāf* wlo' r'iyθiyw 'aḏ—hennāh  
**tear**<sub>1</sub> *he-is-torn-apart*<sub>4</sub> and-not I-have-seen,-him until—here  
 Gn44<sup>28</sup>: 'He is torn to pieces, and I have not seen him since'
- c. **bāšel** *mvuššal* bammāyim  
**cook**<sub>3</sub> *cooked*<sub>4</sub>-PART. in-the-water  
 Ex12<sup>9</sup>: '[Eat not of it raw, nor] sodden at all in water'

These data present a queer asymmetry. We often find cases of < pā'al, nif'al > with *nif'al* as the passive of *pā'al*. However, we never find cases of < pi'el, pu'al > with *pu'al* as the passive of *pi'el*. Compounding the oddity, all instances of < pā'al, pu'al > involve roots which either do not occur in *pā'al*, or do so only with a meaning different from their *pu'al* form. The nearest we come to < pi'el, pu'al > is (17c), where V is a *pu'al* participle and 'Λ', an adjective bearing some resemblance to a *pi'el* cleft form; and the example is, to the best of my knowledge, the only of its kind.

Though not an uncommon binyan, clefts with V in *pu'al* are too scarce to permit a definitive analysis of these facts. However, the same facts hold for *hif'iyl* and its passive *hof'al*: there are instances of < hof'al, hof'al > and of < pā'al, hof'al >, yet none of < hif'iyl, hof'al >. These point to a stronger restriction than feature containment. The absence of a 'real' *pā'al* form of the root T.R.P., &c., suggests that, when *pā'al* morphology occurs on Λ but not on V, Λ is a bare root with default morphology. The restriction, then, is that Λ must be a root, or must bear the same features as V, where those features can be spelt out differently in the morphonology.

The morphological facts reviewed suggest that the predicate cleft construction involves a reduced duplicate of the finite verb. In particular, Λ is a bare head, consisting of a verbal root and optionally of binyan with the same semantic or featural content as the verbal binyan.

The next question is what accounts for these restrictions. Allied to this question is how the predicate cleft is generated. The two obvious options are that Λ is derived by copy and movement operations, and that Λ is an independent lexical item, base-generated or merged where it is found. These issues belong to the ...

### 3. Syntax of the Predicate Cleft

In this section, I argue the Classical Hebrew predicate cleft to be derived by movement, specifically, by head movement, given the cleft form is a head. The argument is based on the different word orders attested in predicate clefts. The cleft form, Λ, is not constrained to occur preverbally. In particular, there are three interesting conditions under which Λ does not occur preverbally: when the usual preverbal landing site of Λ is occupied, when Λ is not a bare head, and when Λ could only move by violating a constraint on movement. Moreover, the

