In this chapter I look more closely at the domain of phrase structure below and above the E projection.\(^1\) What I have been creating below the E projection is an articulated VP, which encodes parts of the verb meaning that are often not independently realized. For instance, while in Tagalog the intransitive verb *drop* and the transitive verb *drop* are distinguished through morphology, in English they are not. One of the questions, then, that can arise is whether or not this is a matter of syntax or is it something that should be kept in the lexicon. In particular, we will see that many of the phenomena discussed above are quite idiosyncratic in their application suggesting that, indeed, they are part of the idiosyncracies of the lexicon rather than part of the computational system of syntax.

Hale and Keyser (1993) introduce a new level to the grammar by suggesting that syntax may be divided between S-syntax (syntactic syntax) and L-syntax (lexical syntax). As with any innovation, the range of application of this new level must be motivated and constrained. Below I examine the characteristics of L-syntax with the aim of both determining and restricting its use. I will argue that event related categories such as ASP and EVENT play an important role in the representation of event structure in phrase structure and that the event related category E represents the phrase structure boundary between L-syntax and S-syntax. Evidence will come from causatives in Tagalog and Malagasy and from empty anaphors in Tagalog. By investigating these two languages, issues that are obscured in many better studied languages become clearer.

\(^1\) Much of the material in this chapter appeared in Travis (2000).
6.1 BACKGROUND

Hale and Keyser (H&K 1993) observe, following Clark and Clark (1979), that denominal verbs such as *shelve* appear in structures that have a near paraphrase containing the nominal. A typical example of such pairs is given in (265) below.

(265)  a. The librarian put the books on the shelf.

       b. The librarian shelved the books.

They argue that, while traditionally such pairs would only be related through some morphological relationship within the lexicon, in fact the relation can be described using the vocabulary of syntax. Their argument is that, since denominal verb formation displays the same array of constraints as head movement, it is best accounted for through these syntactic notions.\(^2\)

Thus, for example, if established principles of syntax function to constrain denominal verb derivations, then the simplest assumption to make is that these derivations are, in fact, syntactic in nature. (1993: 54)

Given the pair in (265), one could imagine a derivation where the two structures have similar underlying representations, but that in the denominal form shown in (265b) there is head movement from the prepositional object position through the P to the V. The derivation would be as shown in (266) below (H&K 1993: 70).\(^3\)

---

\(^2\) I use a combination of Hale & Keyser (1993) and Hale and Keyser (2002). Because the structures and processes I argue for are in several ways more compatible with the older work, I often use the trees, the terminology, and the account of Hale & Keyser (1993). However, on more minor issues such as the use of DPs rather than NPs, I follow Hale & Keyser (2002). I discuss differences as they come up.

\(^3\) Here I have updated the NPs to DPs. I have chosen not to use the representation in Hale and Keyser (2002) as there is no lower V in that structure (see Hale and Keyser 2002: 7). The morphology in the languages central to my research (e.g. Malagasy and Tagalog) suggest a need for two verbal heads. Further, I follow the earlier work in assuming that there is head movement in these structures rather than conflation (see Hale and Keyser 2002: Chapter 3) for a discussion of the distinction. They do, however, allow that deadjectival verbs may be derived in the syntax via incorporation. It is this type of process that interests me.
In Hale and Keyser (2002), syntactic movement is reserved for d-adjectival verbs such as to thin and they use a different process, conflation, for denominal verbs. In my discussion of causatives, it is the deadjectival type of construction that interests me.

If H&K have succeeded in demonstrating that certain cases of verb formation are created through syntactic means, one might ask why it is not simply syntax. Why does the notion of an L-syntax have to be introduced? Not surprisingly, given that this process is generally considered to be a lexical rule, it is quite easy to argue that denominal verb formation has lexical characteristics. I use four diagnostics for lexical rules: change of category, semantic idiosyncracies, phonological idiosyncracies, and lexical idiosyncracies.
(non-productivity). In what follows, I will be looking at two different processes that may arguably occur in both L-syntax and S-syntax. My aim will be to examine the differences in the L-syntax use of the construction and the S-syntax use to determine whether a principled distinction may be made between the two. My conclusion will be that there is a principled distinction which is not surprising — one appears to happen in the lexicon and is therefore idiosyncratic, while the other arguably happens in the computational system (i.e. syntax) and is therefore productive. Other distinctions, however, must also be accounted for and it is the investigation of these that leads us to interesting results. One distinction involves a consistent difference in morphological realization in certain L-syntax processes and S-syntax processes. I claim that this difference is due to morphology that appears in E. The other is a principled account for what syntactic processes can and cannot be done in the L-syntactic component. This, I argue, follows from a view of event structure and a connected view of phrase structure, which I will elaborate on later in the discussion.

6.2 CAUSATIVES

Causatives provide an obvious place to start looking at a lexical versus productive distinction. I will begin by looking at the two causatives in English arguing that the lexical causative is part of L-syntax and the productive causative is part of S-syntax.

6.2.1 English

A clear example of an L-syntax causative in English comes from de-adjectival verbs such as the verb to thin. H&K use this verb to argue for an L-syntax operation, which, through

---

4 Kiparsky (1997) argues that it is this canonical use of the element and not syntactic structure that predicates possible denominal verbs. His arguments are convincing but I do not see that a syntactic account is precluded. The point would still be that the N → V shift occurs in the syntax.

5 In Marantz’s work, there is a discussion about apparent lack of productivity. Marantz (2001: section 3) suggests that, in some cases, lack of productivity does not point to anything deep, but rather accidental gaps.

6 Shibatani (1976) provides a nice overview of the distinction.

7 In fact, it may be that examples like shelf have become lexical. Note that we can say ‘Shelve the books on the windowsill.’ In such a construction, it is difficult to see what the exact structure would be if it were to be derived through syntactic movement. This is one of the reasons that in Hale & Keyser (2002:71), denominal verbs are derived by a different process, conflation. Deadjectival verbs, however, do not have
head-movement, incorporates an adjective into a verb as shown in (267) below (Hale and Keyser 1993: 72).

(267) a. The cook thinned the gravy.

b. The structure for (267b) is given in (268c).

c. The cook (CAUSE) the gravy (BECOME) thin.

We can see the transitive verb to thin as containing sub-lexical items meaning something like ‘cause’ and ‘become’ and it is through these two empty V heads that the adjective thin moves. We can also assume that this process is used to account for transitivity alternations such as the one given in (268) below. The structure for (268b) is given in (268c).

(268) a. The vase broke.

b. The child broke the vase.

draft: 8/10/06
d. The child (CAUSE) the vase break.

The fact that these causatives are lexical (or are part of L-syntax) is clear from their characteristics. As we can see in examples (269)-(271) below, they can change category, they are semantically idiosyncratic, and they undergo lexical phonological changes.

(269) \[ [A \text{ thin}] \Rightarrow [v \text{ thin}] \]

(270) The make-up artist reddened the movie star’s cheeks.
  ≠ The make-up artist caused the movie star’s cheeks to redden.

(271) The chef softened the butter.
  = sofnd
  ≠ softnd

Further, it is not a productive process. As we will see later, only unaccusative verbs in the sense of Perlmutter (1978) can undergo lexical causativization. But as (272) below shows, not even all unaccusatives can be causativized in this way.⁹

---

⁸ An alternative is that roots are category neutral and all lexical category information is added syntactically (see Marantz 2001). I have been convinced by Demirdache and Matthewson’s (1995) work on Salish and my own work on Malagasy (Travis 2001) that roots do have categorial information.

⁹ One way of looking at this, however, is simply that ‘kill’ is a suppletive realization of CAUSE-die.
(272) a. *They will fall the children. (=They will cause the children to fall.)
b. *They died the plants. (=They killed the plants.)

Productive causatives in English are formed by the addition of a causative verb such as make.¹⁰ Unlike the lexical causative, the productive causative cannot trigger a change in category nor lexical phonological processes. The semantics is fairly predictable and the process is productive. So, for example, while the lexical causative could not apply to the unaccusative verbs given above, the productive causative may.

(273) a. They made the children fall.
b. They made the plants die.

In English, then, we have the two types of causative that we want to study. The lexical causative is idiosyncratic in the expected ways and may be indicated by no overt morphology. The productive causative is always indicated by a separate lexical item and is productive with predictable results.

6.2.2 Tagalog and Malagasy

Tagalog and Malagasy also have two types of causative (we briefly looked at Malagasy causatives in Chapter 3, section 3.3.1.2xx), but they are more instructive than the English equivalent as they use affixation in both cases. In fact, I will argue that the morpheme used is exactly the same not only in Malagasy as we have seen, but also in Tagalog. The difference in appearance is due, I argue, to a morpheme that always co-occurs with the productive causative morpheme but not with the lexical causative morpheme. The difference in behavior between the productive causative and the lexical causative, I argue, has to do with where on the tree the morpheme is generated since the position on the tree will determine whether the morpheme is part of L-synta or S-syntax.

¹⁰There are other causative verbs such as cause and have which show the productivity of make but which have characteristics of their own as described, for instance, by Ritter and Rosen (1993).
In Tagalog, the lexical causative is formed by adding the prefix *pag-* to the root. Some examples of the alternation already seen in Chapter 3, section 3.2.1xx, are repeated in (274) below.

(274) Alternations (Maclachlan, 1992)

\[\begin{align*}
\text{a. tumba} & \quad \text{X fall down} & \text{b. pagtumba} & \quad \text{Y knock X down} \\
\text{sabog} & \quad \text{X explode} & \text{pagsabog} & \quad \text{Y scatter X} \\
\text{luwas} & \quad \text{X go to into the city} & \text{pagluwas} & \quad \text{Y take X to the city} \\
\text{sabit} & \quad \text{X be suspended} & \text{pagsabit} & \quad \text{Y hung X} \\
\text{sali} & \quad \text{X join} & \text{pagsali} & \quad \text{Y include X}
\end{align*}\]

Note that the morpheme *pag-* may be used even when there is no alternation, i.e. when there is only the transitive form of the root. Therefore, even though there is no form *luto* meaning something like ‘X be cooked’, there is a form *pagluto* meaning ‘Y cook X’.

Within sentences, the forms given in (274) above are combined with another morpheme. In (275a) the other morpheme is *-um-* and in (275b) it is *n-* which I assume to be an Actor Topic morpheme and a perfective morpheme respectively, following e.g. Maclachlan (1989) and the discussion in Chapter 3.

(275) a. Tumumba ang bata t-um-umba
\[\text{AT-PERF-tumba NOM child um=AT; 0=PERF} \]
‘The child fell.’

b. Nagtumba ng bata si Rosa. n-pag-tumba
\[\text{AT-PERF-pagtumba ACC child NOM Rosa 0=AT; n=PERF} \]
‘Rosa knocked the child down.’

We have already seen that in Malagasy we can find similar alternations, also mediated by morphology. Again, my assumptions here, following Hung (1988), are that the inchoative form is *-i-* and the lexical causative is formed by the addition of *-an-. The analysis of the *m-* found in both members of each pair I analyze as an Actor Topic morpheme such as the *-um-* found in Tagalog. Further, like in Tagalog, there are some
forms that do not have an unaccusative counterpart such as manome ‘Y give X to Z’. The forms are given in (276) with an exemplifying sentence in (277).

(276) Alternations (Malagasy)

a. mihisatra X move slowly      b. manisatra Y move X slowly
   milahatra X be in order       mandahatra Y arrange X
   milona X soak                mandona Y soak X
   misitrika X hide             manitrika Y hide X

no alternation:

   manome Y give X to Z

       PST.AT.hide PST.there ACC-house NOM.3SG
       ‘He hid in the house.’

       b. Nanitrika ny vola tao an-trano izy
          PST.AT.hide the money PST.there ACC-house NOM.3SG
          ‘He hid the money in the house.’

It is immediately clear that these causative alternations are lexical in nature. It can be argued, in Malagasy at least, that they always change category. Typically the roots are either nominal or adjectival in nature. Some examples of typical roots are given in (278) below.\footnote{See Phillips (2000) for an extensive discussion of the use of roots in Malagasy. In the main Malagasy-French Dictionary (Abinal&Malzac 1988), which was first published in 1888, all roots are given a category label even though the root might never be used on its own.}

---

\footnote{See Phillips (2000) for an extensive discussion of the use of roots in Malagasy. In the main Malagasy-French Dictionary (Abinal&Malzac 1988), which was first published in 1888, all roots are given a category label even though the root might never be used on its own.}
(278) Malagasy

a. hisatraN action of slowly moving
b. lahatraN organization
c. lonaN action of putting in a liquid
d. sitrakaN action of hiding

Also, there are clear cases of semantic drift. In Tagalog, the bare root sabog means ‘to explode’ while the causativized form means ‘to scatter’. This has the result that one form may be used in situations where the other one would produce a semantically odd sentence.

(279) Tagalog

a. Sumabog sa Boston ang bomba

   AT-PERF-sabog in Boston NOM bomb

   ‘The bomb exploded in Boston.’

b. # Nagsabog ng bomba sa Boston ang terorista

   PERF-pag-sabog ACC bomb in Boston NOM terrorist
can’t mean: ‘The terrorist exploded the bomb in Boston.’

   get odd reading: ‘The terrorist scattered the bomb.’

Further, it is clear in Malagasy that the causativizing suffix triggers a lexical rule of phonology rather than a post-lexical rule. In a post-lexical rule, such as reduplication, a nasal preceding a consonant triggers prenasalization. In the lexical rule that is triggered by the lexical causative affix, the result is fusion, the voiceless consonant will drop.

(280) Malagasy

POST-LEXICAL (pre-nasalized consonant)

\[ n + p \rightarrow mp \quad \text{pentsona} + \text{pentsona} \quad \text{pentson}^m \text{pentsona} \quad \text{N. ‘chatter’} \]

\[ n + s \rightarrow pts \quad \text{m+an+sampona} + \text{sampona} \quad \text{manampo}^n \text{tsampona} \quad \text{V. ‘to stop’} \]
Finally, the lexical causative, while more productive than in English, is not completely predictable. Some forms may appear with or without the prefix with no change in meaning as example (281a) shows; some, where we expect the prefix because the verbs have external theta-roles (i.e. are not unaccusative), do not have it as (281b) shows.

(281)a. Tagalog
   
   hiwa or paghiwa  X cut/slice Y

   b. Malagasy
   
   mividy  X buy Y

All of these characteristics simply confirm that this is a lexical rule of causative with all its expected idiosyncracies. The end result, however, is that there is a morpheme that is used in both languages that in many cases indicates a transitivity alternation (i.e. causativization) between two forms. Also, in both languages, this morpheme may be used simply to indicate a transitive (or agentive) structure even if there is no intransitive counterpart. In Tagalog the lexical causative morpheme is pag- and in Malagasy it is an-.

(282) INTRANSITIVE TRANSITIVE (agentive) (generally)
    (UNACCUSATIVE) (LEXICAL CAUSATIVE)

Tagalog  0 pag-
Malagasy   i- an-

Both Tagalog and Malagasy also have another causative which is much more productive and predictable. As we saw in Chapter 3, in Malagasy, the productive causative appears to be formed by attaching the causative morpheme, amp-, to the stem
and we reanalyzed this as an $f$. The examples in (283) below show that the stem may either take the form of the lexical causative verb ($anitrika$) or the unaccusative verb ($isitrika$). It is clear that the productive causative morpheme can be attached to a form containing the lexical causative morpheme. This is shown in (283b').

(283) Malagasy ($amp-$, or $an-$ + $f-$)

<table>
<thead>
<tr>
<th>STEM</th>
<th>PRODUCTIVE CAUSATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>$misitrika$ ‘X hide’</td>
</tr>
<tr>
<td>b.</td>
<td>$manitrika$ ‘Y hide X’</td>
</tr>
<tr>
<td>b’</td>
<td>$m + an + f + an + sitrika$</td>
</tr>
<tr>
<td></td>
<td>$M + PC + F + LC + ROOT$</td>
</tr>
</tbody>
</table>

In Tagalog the productive causative prefix is, for the most part, $pagpa$- which I will argue is $pag + pa$ parallel to the Malagasy $an + f$.\textsuperscript{12} What makes it different from Malagasy is the effect that the addition of this morphology has on the realization of the stem. Once the productive causative morpheme has been added, the lexical causative morpheme drops. This has the end result of collapsing the unaccusative form with the lexical causative form making the productive causative ambiguous between the two. In other words, when the productive causative morpheme $pagpa$- is attached to the stem $pagsama$, instead of getting $pag-pa-pag-sama$, the form is $pag-pa-0-sama$, homophonous with the productive causative form of the unaccusative.

(284) Tagalog (Actor Topic: -$pagpa$-)

<table>
<thead>
<tr>
<th>STEM</th>
<th>PRODUCTIVE CAUSATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>$sumama$ ‘X be with Z’</td>
</tr>
<tr>
<td>b.</td>
<td>$magsama$ ‘Y combine X’</td>
</tr>
</tbody>
</table>

\textsuperscript{12}The forms of the verb which do not surface as $pagpa$- will be very important to the discussion of the morphological analysis of the causative and will be looked at in more detail below.
Chapter 6

Above we have been looking solely at Actor Topic forms. In the Theme Topic form of the productive causative, where the Causee appears as the Subject/Topic, however, we get two interesting results. One is that the form of the productive causative morpheme changes from *pagpa* to *pa*-. The second is that the lexical causative morpheme reappears on the transitive stem. This is shown in (285) below. Sentences are given showing the use of each form. With the morphological change of the verb comes a change of the Subject/Topic, realized as the *ang* DP.

(285) Tagalog (Theme Topic: *pa*-)

a. pasamahin 'X be made to be with Z' (see (284a))
b. papag*pag*samahin 'Y be made to combine X' (see (284b))

b'. pa + pag + sama + in

PC? + LC + ROOT + ThemeTopic

The first observation suggests that the productive causative morpheme, in fact, consists of two morphemes, *pag*- and *pa*-, and the *pag*- drops in the Theme Topic form. Independent evidence for this analysis comes from the fact that the lexical causative morpheme *pag*- also drops in the Theme Topic form of the simple lexical causative. The relevant forms are given in (286) below.

(286) Tagalog

a. Actor Topic of lexical causative: *pagsama* 'X combines Y'
b. Theme Topic of lexical causative: *samahin* 'Y is combined by X'

While I save the account of why the 'top' *pag*- drops in the Theme Topic form till section 6.4.2xx, we can now at least make the observation that *pag*- drop occurs with both the productive causative and the lexical causative in particular forms.
This brief excursion into Tagalog morphology has left us with two results. One result is that, underlyingly, Tagalog and Malagasy are quite similar. The productive causative morpheme may be attached to a form containing the lexical causative morpheme. In Tagalog, this fact is obscured, though, by the null realization of the lexical causative when the productive causative is overt. The second result is that, because we are forced to reanalyze the productive causative morpheme in Tagalog into pag- and pa-, we now can see that at least part of the productive causative morpheme is identical to the lexical causative morpheme (in both instances pag-). In fact, this is very similar to the analysis of the Malagasy productive causative, which I argued in Chapter 3 also consists of two morphemes, an- and f-. As proposed earlier for the morpheme f- in Malagasy, we will assume that the extra Tagalog morpheme pa- is generated in E. What distinguishes the productive causative from the lexical causative is where the causative morpheme is generated on the syntactic tree — productive causatives are generated above E and lexical causatives are generated below E.

My account for these morphemes is basically a development of the one presented in Hung (1988), but I have used her results as a starting point to investigate the differences between L-syntact and S-syntact. We have seen that causatives divide nicely between l-syntact and s-syntact, but we would expect this distinction to show up in other areas of the grammar. We will see in the next section that Tagalog offers another phenomenon that shows the same split in properties.

6.3 EMPTY ANAPHORS IN TAGALOG

In Tagalog there is evidence for an empty category that is obligatorily bound. Because it is empty but in a position that can be filled, I assume that this empty category is pro. Because it is obligatorily bound, I assume that it is anaphoric. del Pilar (1993) argues that this anaphoric pro appears in productive (syntactic) causatives and has very particular characteristics which suggest that it has syntactic status (and is not simply

---

13This analysis of causatives in Tagalog is not universally accepted. Many simply see the causative morpheme as unanalyzable. Schachter and Otanes (1972) simply list it as magpa- in the AT form and pa- in the TT form. Carrier (1979) breaks magpa- down into mag- and pa- but treats pa- as the causative morpheme and mag- as the AT morphology. Rackowski (2002) analyzes pag- as an anti-EPP marker found in voiceP following ideas of Pylkkänen (2002). I comment on Rackowski’s analysis in section 6.4.3xx.
pragmatically determined). As she points out, anaphoric pro also appears in some non-causative structures. In the next section I develop her observations and conclude that anaphoric pro may be licensed by morphology that is added either in the L-syntax or the S-syntax. If it is licensed by the L-syntax morpheme, it shows the idiosyncrasies expected at this level. If it is licensed by morpheme added in the S-syntax, however, its behavior is predictable.

6.3.1 S-syntax anaphoric pro

del Pilar (1993) begins by introducing the productive causative in Tagalog which we have already seen above. As we can see in (287a,b) below, the productive causative turns a two argument predicate into a three argument predicate with the addition of an Agent.\textsuperscript{14}

\begin{align*}
(287) & \quad \text{a. Sumundo} \quad \text{si A ng B} \\
& \quad \text{AT-PERF-fetch NOMA ACCB} \\
& \quad \text{‘A fetched a B.’} \\

& \quad \text{b. Nagpasundo} \quad \text{si A kay B ng K} \\
& \quad \text{AT-PERF-pagpa-fetch NOMA OBLB ACCK} \\
& \quad \text{‘A caused B to fetch a K.’} \\
\end{align*}

She notes further, however, that in the productive causative, one can leave out the third argument and get obligatory binding with the causer.

\begin{align*}
(288) & \quad \text{Nagpapuri} \quad \text{si A kay B} \\
& \quad \text{AT-PERF-pagpa-praise NOMA OBLB} \\
& \quad \text{‘A caused/let B to praise A’s self’} \\
\end{align*}

This is unlike pro-drop in that the antecedent may not come from outside the sentence. So, for instance, (288) cannot mean that A caused/let B praise C. Further, the

---

\textsuperscript{14} I use del Pilar’s data which indicates arguments with letters.
empty category must exist in the syntax since it is able to control PRO in a control structure as in (289b) below.\(^\text{15}\)

\[(289)\text{a. Nagpahikayat} \quad \text{si A kay B ng K na bumili ng bahay}\]
\[
\begin{array}{llllllllllll}
\text{AT-PERF} & \text{cause-persuade} & \text{NOM} & \text{A} & \text{OBL} & \text{B} & \text{ACC} & \text{K} & \text{COMP} & \text{AT-buy} & \text{ACC} & \text{house} \\
\end{array}
\]

‘A caused/let B (to) persuade K to buy a house.’

\[
\text{b. Nagpahikayat} \quad \text{si A kay B na bumili ng bahay}\]
\[
\begin{array}{llllllllllll}
\text{AT-PERF} & \text{cause-persuade} & \text{NOM} & \text{A} & \text{OBL} & \text{B} & \text{COMP} & \text{AT-buy} & \text{ACC} & \text{house} \\
\end{array}
\]

‘A caused/let B (to) persuade A PRO to buy a house.’

\[
\text{c. [IP nagpahikayat [DP si A]_i kay B [DP pro]_i [CP na [IP bumili PRO_i ng bahay ]]]}
\]

In (289a) the third argument, K, controls the empty subject of the embedded clause. In (289b) this third argument is not lexically realized and the empty embedded subject appears to be controlled by the highest argument, A. del Pilar assumes that the control facts are captured by assuming a syntactically active but not lexically realized third argument. This argument is an anaphoric pro which takes as its antecedent the highest argument, A. This is shown structurally in (289c) where A binds pro and pro controls PRO.

What is interesting for my purposes, however, is that these forms cannot passivize (i.e. but put in the Theme Topic form) as (290) below shows.

\[(290)\text{Pinapagpuri} \quad \text{ni A si B}\]
\[
\begin{array}{llllllllllll}
\text{TT-PERF} & \text{papag-praise} & \text{GEN} & \text{A} & \text{NOM} & \text{B} \\
\end{array}
\]

≠ ‘B was caused by A to praise B/A’s self’

\(^{15}\) Some speakers find this construction odd not because of the binding, but because they do not get object control structures. (289a) is also not possible for them. In dialects, then, that do not have object control, this cannot be tested. I am relying on data provided in del Pilar (1993). I am grateful to R. Mercado for discussion of these data.
At this point we can make the following observations. With the productive causative morpheme, we can license an empty category that behaves like an anaphor in that it must be bound, and its antecedent must be the Agent in an Actor Topic construction.

What I suggest in this section and hope to confirm in the following section is that the anaphoric pro of Tagalog is similar to the long-distance subject oriented anaphors of languages like Icelandic and Chinese. I argue that two conditions must hold in order for the anaphor to be licensed — the pag- morpheme must be overt, and the antecedent must be in subject position. Before turning to the account of Tagalog, I give a brief introduction to one of the first accounts of long distance anaphora.\footnote{The phenomenon of long-distance anaphora has been the topic of many papers. Some of the relevant references can be found in Cole et al. 2001. Pica’s account is sufficient for my needs.}

Pica (1987), taking data from Scandinavian languages, investigates the problem of long-distance anaphora. His observation is that long-distance anaphors are monomorphemic while local anaphors are compound. This is very clear in a language like Chinese where the long-distance anaphor is \textit{ziji} and the local anaphor is \textit{ta ziji}. In the examples below, we see in (291a, b) that the long-distance anaphor may be bound by a DP which is outside of a small clause in (291a) and outside of an embedded (subjunctive) sentence in (291b). Example (291c) shows that this anaphor in Danish cannot take an object as its antecedent.\footnote{These examples appear in Pica’s paper without glosses.}

\begin{tabular}{l}
(291)a. & Han\textsubscript{i} betragter patienten som farlig for sig\textsubscript{i} & Danish (Pica 1987: 484) \\
& ‘He considers the patient as dangerous for himself.’ \\

b. & Jón\textsubscript{i} upplýsti að María elska sig\textsubscript{i} & Icelandic (Pica 1987: 484) \\
& ‘Jon says that Mary loves himself.’ \\

c. & * Jeg fortæller Hans\textsubscript{i} om sig\textsubscript{i} & Danish (Pica 1987: 485) \\
& ‘I told John about himself.’ \\
\end{tabular}
Pica’s account links the monomorphemic shape of the anaphor with the fact that it may be bound long-distance and that it is subject-oriented. He assumes that monomorphemic anaphors are $X^0$s rather than XPs. As $X^0$s they move to INFL\(^\text{18}\) at LF and in this position take the closest c-commanding DP, which is the subject, as their antecedent. In this account, the structure for (291a) would be as in (292).\(^\text{19}\) (291c) is ungrammatical since the object will not c-command the anaphor at LF and therefore cannot act as its antecedent.

(292) \[ \text{IP} \]
\[ \begin{array}{c}
\text{DP} \\
\text{Han} \\
\text{I}' \\
\text{I} \\
\text{VP} \\
\text{X}^0 \\
\text{sig} \\
[+\text{tense}] \\
\end{array} \]

Along the lines of Pica, one could propose that the empty anaphor in Tagalog is an $X^0$, it must move to INFL to be licensed, and in the position of INFL it can take only the subject as its antecedent. The question remains, however, why these forms cannot passivize (i.e. appear in the Theme Topic form) as in (290). Recall from our discussion of productive causatives above that the $pag$- morpheme in a productive causative disappears in the Theme Topic form. I hypothesize for the moment that it is this morpheme that licenses the empty anaphor in INFL and if this morpheme is not lexically realized, then the empty anaphor cannot be licensed. This issue will be brought up again in the next section.

6.3.2 L-syntax anaphoric pro

What is interesting about the anaphoric pro found in the causative construction is that a similar phenomenon occurs in structures that do not contain a productive causative. del Pilar points to a few examples in her paper such as the one in (293) below.

\[^\text{18}\text{In fact, the } X^0\text{ anaphors must move to INFL to be saturated. I refer the interested reader to Pica’s article for more details. Whether or not this is the appropriate way to account for long distance anaphora is actually not crucial to my analysis. It is only important to note that Tagalog anaphoric pro behaves like a long distance anaphor in being subject sensitive.}\]
Other verbs which allow this alternation are: wash, shave, dress, clean, shoot, cure, hit/whip, shut in, blame, force, lose.\textsuperscript{20} Note that the Actor Topic form of the verb changes from the paradigm which shows -\textit{um}- insertion (\textit{b-um-aril}) to the paradigm which contains \textit{n+pag} (\textit{n+pag+baril}). In these cases of anaphoric pro, the fact that there is a missing argument seems to have to do as much with the meaning of the verb as with the construction in which the verb appears. In fact, many of the Tagalog verbs which allow a reflexive reading also allow a reflexive reading in English when the second argument is dropped, such as \textit{wash} and \textit{shave}.\textsuperscript{21} On closer examination, however, we can see that what is crucial for the anaphoric reading of these verbs is not only the choice of verb but also the syntactic configuration that it appears in. To show this more clearly I turn to an article by Carrier-Duncan (1985), which discusses the issue in more depth.

Carrier-Duncan, using lexical rules, sets out to collapse two phenomena. She starts by describing Rule 1 and Rule 2. With Rule 1, the second argument of a verb appears to be bound to the first argument. In (294a,b) below we can see her description of the facts. She assumes that the verb form remains the same, but that the choice of topic paradigm changes. The verb in (294a) with no binding chooses the -\textit{um}- form for the Actor Topic while the verb in (294b) with the argument binding chooses the \textit{mag}-form of the Actor Topic. The paradigm choice is indicated by the morphemes placed above each of the arguments in a theta-grid. For example, with the root \textit{\textcircled{h}iwalay}, if the

\textsuperscript{19}In (291b) the anaphor would have moved to the higher INFL.
\textsuperscript{20} Del Pilar does not give the Tagalog equivalents of these verbs. Since this process is so variable across speakers, it is difficult to know exactly which forms she had in mind.
theme (highest argument) becomes the subject, the morphology that appears on the verb is the infix -um-. If the source argument becomes the subject, the relevant affix on the verb is -an. In the form of the root that shows the binding effect, the form of the verb which surfaces when the highest argument (Theme) is the subject is a mag-form. Since the source argument in this form is always null (i.e. bound by the theme argument), it never appears as the subject so there is no morpheme required.

\[(294)\]
\[
\begin{array}{c|c|c}
\text{-um-} & \text{-an} \\
\hline
\text{a. HIWALAY:} & (\text{theme source}) \\
\text{‘x separate from y’} & \\
\text{b. HIWALAY:} & (\text{theme}_{\text{i}} \text{source}_{\text{rec, i}}) \\
\text{‘x and y separate from each other’} & \\
\end{array}
\]

Rule 1, which binds the second argument with the first argument, can be used with other verbs such as fight with, meet, see, converse, and triggers a reciprocal reading as shown in the Actor Topic constructions below.

\[(295)a. \text{H-um-iwalay sa kaibigan ang bata (adapted from C-D)} \]
\[
\begin{array}{llll}
\text{AT-PERF-separated} & \text{SA} & \text{friend} & \text{NOM child} \\
\end{array}
\]
\‘The child separated from his friend.’

---

21 It is equally important to note, however, that other Tagalog verbs are quite different from their English counterparts such as hit, cure, blame, etc. *The child bit. (with the meaning the child bit himself.)

22 Carrier-Duncan assumes that this promotion of arguments via verbal morphology is a process of topicalization and not a process of promotion to subject. So as not to confuse the reader, I describe and gloss the Tagalog data in a way that is consistent with my view of this process. In doing so, however, I depart from Carrier-Duncan's original characterization of these facts. Further, while I continue to call the -um- and mag- forms as Actor-Topic forms, and the -in forms as Theme-Topic, as is done in the Austronesian literature.

23 Carrier-Duncan also only gives the English translations and not the relevant Tagalog roots.
b. Nag-hiwalay ang mga kaibigan
   AT-PERF-PAG-separated NOM PL friend
   ‘The friends separated from each other.’

C-D’s conclusion, then, is that by changing from the -um- AT paradigm to the mag-AT paradigm, the verb triggers the binding of the second of its arguments by the first of its arguments. The result is that a two-argument verb becomes a one-argument verb with a reciprocal interpretation.

Rule 2 applies to three argument verbs and it binds the third argument to the second argument. Once again, according to Carrier-Duncan, the rule does not add morphology to the verb, but it does affect the choice of paradigm for topic morphology. Without the binding, the Theme Topic morphology is i-, but with the binding, the Theme Topic morphology is -in (again shown by the morphemes listed over the respective theta-grids).

(296) mag- i- -an

a. SAMAH (agent theme goal)
   ‘x puts y with z’
   (mag-) -in

b. SAMAH (agent theme goal_{1} goal_{rec,1})
   ‘x put y and z together’

Other verbs which can undergo this process are join (x joins y and z to each other), paste (x pastes y and z to each other) and put (x puts y and z near each other). As shown in the examples below, the resulting meaning again is reciprocal.

(297)a. I-sasamah ang karne sa gulay ng magluluto
   TT-IMP-put NOM meat SA vegetables GEN cook
   ‘The cook will put the meat with the vegetables.’ (adapted from C-D: )
b. Pag-sasamah-in ang karne at gulay ng magluluto

\textit{PAG-IMP-put-TT NOM meat and vegetables GEN cook}

‘The cook will put the meat and vegetables together (with each other).’

The verb forms undergoing Rule 2 are the most interesting at this point because of the restrictions placed on them, and a morphological quirk that they show. Note first that the morphological paradigm given for these verbs has the Actor Topic form (\textit{mag-}) in parentheses in (296). It is in parentheses is because this verb form never appears in an Actor Topic construction, but only in a Theme-topic construction. Carrier-Duncan explains this in the following way:

For subclass 2 [verbs undergoing Rule 2], the [\textit{AT}] form is not used in a sentence, a quirk shared by a few non-derived verbs as well. However, the [\textit{AT}] form exists since it serves as the stem to which the [\textit{TT}] suffix -\textit{in} is added (causing \textit{mag-} to show up as \textit{pag-}) (emphasis mine: LDT) (Carrier-Duncan 1985: 15)

It is strange that the derived verb form is part of a small idiosyncratic verb class when the underived form behaves normally. Another way to explain this fact, however, is to say that there is a syntactic restriction on the binding relation and that the antecedent must always be in the subject position. This is central to my account.

There is a further oddity to be noted. A strange morphological fact about these verb forms is that the \textit{pag-} from the Actor Topic form remains in the Theme Topic form (and C-D gives this as her reason for knowing that the AT form is \textit{mag-} even though it is unattested). This retention of \textit{pag-} in the TT form is unlike both the productive causative \textit{pag-} and the lexical causative \textit{pag-}. It is certainly unlike any other paradigm of topic morphology. To see this more clearly, let us look more closely at the paradigm choice for the bound forms as described in (296b). The paradigm chosen by these verbs is \textit{mag-} for AT (which, in fact, never surfaces) and -\textit{in} for TT. While most non-derived verbs that take -\textit{in} for TT take -\textit{um-} for AT, there are some verbs that do take \textit{mag-} as AT and -\textit{in} as TT. When they do, however, the \textit{pag-} predictably disappears in the TT form. A few examples of this are given below.
The paradigm for the Rule 2 verbs, then, is odd for two reasons. The AT form never surfaces, and the TT form retains the pag- morpheme. In fact, though it is not as clear in the case of the Rule 1 verbs, we can make a generalization across both sets that the antecedent will always be the subject (forcing the Rule 2 verbs to appear in the TT form), and the pag- must always be present (forcing the unexpected TT form of the Rule 2 verbs).

These characteristics now make C-D’s reciprocal verbs (e.g. (295b) and (297b)) look very similar to del Pilar’s productive causative reflexive constructions (e.g. (288)) and the lexical reflexive verbs (e.g. (293b)). In all of these cases, the antecedent must be the subject and the pag- must be lexicalized. To try to relate these data with Pica’s analysis of long distance anaphora, I will assume that the empty anaphor in Tagalog is a head (perhaps non-head anaphors in Tagalog must be lexically realized) and it moves to an INFL (or T position) that contains a visible pag-. It may be the case that the empty anaphor must have a lexically realized morpheme to identify it. This would explain the obligatory presence of pag-. In the T position, the anaphor may have only the subject as its antecedent. We can then propose the structures below for anaphoric binding in the four types of verb we have been discussing: productive causatives, lexical reflexives, lexical reciprocals (Rule 1), and lexical reciprocals (Rule 2).24

---

24 An alternative would be the one presented in Reinhart and Reuland (1993). What I have been calling anaphoric pro would be an SE in their terms. They would specify the L-syntax cases as being reflexive roots but I’m not sure how they would ensure the right binding relation.
(299)a. PRODUCTIVE CAUSATIVES (del Pilar)

25 I have not indicated whether the productive causative pag- is a $V_1$ or a $V_2$. We will see shortly that it acts like a $V_1$ in terms of morpheme deletion, and it has the same form as the lexical causative $V_1$. Then the question arises as to whether there is a $V_2$ that selects for EP. I’m assuming that there is but have no firm reasons at this point and therefore do not include it. I will leave this V unspecified throughout.
b. LEXICAL REFLEXIVES (del Pilar)

```
TP

anaphor binding

DPᵢ

T'

T

EP

Xₒ

proj


DP

V₁P

V₁'

AspP

Asp

V₂P

DP

V₂'

tj

Asp

V₃P

DP

V₃'

tj

Asp

Xₒ

proj


```

c. LEXICAL RECIPROCAL VERBS (Rule 1: Carrier-Duncan)

```
TP

anaphor binding

DPᵢ

T'

T

EP

Xₒ

proj


DP

V₁P

V₁'

AspP

Asp

V₂P

DP

V₂'

tj

Asp

V₃P

DP

V₃'

tj

Asp

Xₒ

proj


```
This analysis is an attempt to account for what these phenomena have in common. There are ways in which they differ, however. In the spirit of this chapter, I feel that the most appropriate way to categorize the binding cases we have seen is to put the productive causative on one side (299a) and the lexical type binding on the other side (299b,c,d). This means putting together the Rule 1 and Rule 2 cases of Carrier-Duncan with the lexical reflexive examples given by del Pilar. These would all be cases of the L-syntax use of the anaphoric pag-. The productive causative cases would be s-syntactic uses of the anaphoric pag-.

It is easy to see that the L-syntax examples show L-syntax characteristics. Not all verbs can undergo this process. In other words, only certain verbs may add a pag- to their stems and thereby bind one argument with another and the list of verbs varies from speaker to speaker.

Further, it must be determined verb by verb which argument is the bindee and which the binder. In some cases the Agent binds the Theme, in other cases, the Theme
Finally, the verb must determine whether the anaphoric pro will be a reciprocal (as in C-D’s examples) or a reflexive (as in del Pilar’s examples).

The productive causative form of anaphoric pag- shows none of these idiosyncracies. All productive causatives can license the empty anaphor, and in every case the anaphor will be interpreted as a reflexive.

Once again, like with the causative, we have the same morpheme creating essentially the same effect. The differences are determined solely by the position that the morpheme is placed in. If the morpheme is above E (as in (299a)), it is acting like a lexical item on its own which has consistent properties. If it is below E (as in (299b,c,d)), it may be determined by the lexical item of which it is a part, accounting for its idiosyncratic nature. What is important to note, however, particularly with respect to the licensing of the empty anaphor, is the close connection with syntax. Causative formation creates complex words with complex argument structure — both processes that can arguably be kept within the lexicon. The setting up of anaphoric relations has much more of a syntactic flavor to it, however, as it relates to XP and is sensitive to the grammatical relation (subjecthood) of the antecedent. This provides further support for the syntactic side of L-syntax. In the next section, I look more closely at the nature of L-syntax.

### 6.4 WHERE AND WHAT IS L-SYNTAX?

L-syntax is assumed to have some characteristics of the lexicon (category changing, idiosyncracies, etc.) and some characteristics of syntax (head movement), but the question remains as to where it is located in the grammar.\(^{27}\)

#### 6.4.1 Syntax in the lexicon

H&K (1993) appear to want at least a bit of the syntax to appear in the lexicon. This would mean that for denominal verbs such as saddle and shelve, the lexical entry would

---

\(^{26}\) There seem to be some restrictions on this since the binder is always higher in the theta-hierarchy. While one might argue that this makes this binding look like it is a lexical process and not a syntactic process, the fact that the binder must also be the syntactic subject must be explained. I am assuming that syntax is the obvious place to find such an explanation.

\(^{27}\) Butt and Ramchand (2005) use the term first-phase syntax. This is similar but not identical to L-syntax. I leave the reader to make a comparison.
include a phrase structure tree. They put it in the following way (Hale & Keyser 1993: 95).

[i]n thinking about this [the idiosyncracies of denominal verb formation - LDT], we have taken the conservative view and assumed that this array of facts compels us to suppose that the lexical entry for *shelve* includes at least the full syntactic structure depicted in [(300)].

In being conservative, they keep the idiosyncratic information within the lexicon. But, given that there are syntactic components within their account, this forces them to put a bit of syntax in the lexicon as well. The lexical entry for the verb *to shelve* would then be as in (300) (Hale and Keyser 1993:74).

(300) = (H&K: 74)

```
  VP
     /\  
    V  VP
       /\ 
      DP V' 
         /\  
        V  PP
           /\  
          P  NP
             /  
            N
         shelf
```

There is no distinction made, however, between the syntax that occurs in the lexicon or the syntax that occurs in the computational component.

The ‘structures’ implicated in that usage [Lexical Relational Structure - LDT] are simply syntactic structures expressing such normal syntactic structural relations as ‘head’, ‘specifier’, and ‘complement’. And they are present in the syntactic representations over which normal syntactic processes and principles are defined. The qualification ‘lexical’ refers to the property that the argument structures of verbs are ‘listed’ in the lexicon, perhaps in the manner suggested by the conservative view of lexical entries.
Here I choose to follow a different approach. Rather than assuming that a bit of syntax has slipped into the lexicon, I would like to explore the possibility that a bit of the lexicon has slipped into the syntax. This approach will be outlined below, but before doing that, I would like to begin with a problem that H&K raise. The problem arises with the ungrammatical sentences in (301). I also give H&K’s characterization of this problem.

(301) a. * The clown laughed the child. (i.e. got the child to laugh)
    b. * The alfalfa sneezed the colt. (i.e. made the colt sneeze)

These sentences represent an extremely large and coherent class of impossible structures in English. In particular, unergative VPs cannot appear as complements of V in LRS representations - that is, an unergative may not appear in the lexical syntactic “causative” construction. (H&K 1993: 74-5)

In other words, if a zero causative morpheme can be added to (unaccusative) intransitives such as melt and break to form causative counterparts, why can’t this be done with unergative verbs like laugh and sneeze? Since both of these verbs may undergo productive causativization as the intended meanings show, this question can be reworded as: what is the boundary of L-syntax?

H&K’s answer to this question involves their conception of external argument. They assume that external arguments are truly external and can only be added in the syntax (either through predication or the introduction of functional categories and their requirements). External arguments, then, are not generated in the Spec, VP position through the argument requirements of the verb.

This solves the problem raised by (301). Since the external arguments of laugh and sneeze are added in the s-syntax, they cannot be made into internal arguments by additional L-syntax morphology. In other words, at the point where L-syntax applies,

28 As mentioned earlier, Hale & Keyser (2002) present a slightly different picture. English denominal verbs are derived through conflation rather than syntactic movement, while English deadjectival verbs are derived through head movement, and in the syntax. In both cases, however, there is complex structure in the syntax.
these verbs have no arguments. The addition of the causative, then, cannot create a two argument verb.

For independent reasons (see the discussion of the Malagasy maha- forms in Chapter 7), I assume that external arguments are part of the verb’s lexical entry, so I must look for a different solution. I will argue that my solution solves the problems raised by the data in (301) as well as providing an account of causative morphology in Tagalog and Malagasy productive causatives.

6.4.2 Lexical entries in syntax.

It would be very nice if we could find evidence that L-syntax has to be part of the computational component. Hale and Keyser’s strongest argument was that denominal and deadjectival verb formation appear to be restricted by the Head Movement Constraint, arguably a restriction on syntactic movement. However, if L-syntax is truly syntax, we expect to see other types of syntactic effects. In this section I argue that we have evidence of Spec, Head relations at the level of L-syntax. It is difficult to argue that Spec, Head relations cannot be captured in the lexicon. I will suggest, however, that this is most easily done in the computational component.

The data relevant to this argument involve the pag- drop that we have already seen in the discussion of Tagalog causatives. We have seen two different cases of this in the productive causative. First we saw that if the productive causative was added to a lexical causative, the lexical causative morpheme pag- dropped. This is shown again in (302) and (303) below.29 (302) shows this schematically while (303) gives a relevant example.

(302) Productive Causative (Causer = Topic)
\[
\begin{align*}
&m + \text{ pag } + \text{ pa } + \text{ ?? } + \sqrt{\text{walis}} & \text{‘to cause to sweep’} \\
&M + \text{ PC } + \text{ E } + \text{ LC } + \text{ ROOT }
\end{align*}
\]

---

29 I have added glosses to the Schachter and Otanes examples.
(303)a.  *magsama* ‘combine’

Magsama ka ng mansanas sa dalandan

PRES.PAG.be.with 2P ACC apple SA orange

‘(You) combine the apples with the oranges.’

b.  *magpasama* ‘permit/cause to combine

Magpasama ka ng mansanas sa dalandan kay Maria

PRES.PAG.PA.be.with 2P ACC apple SA orange KAY Maria

‘(You) have Maria combine the apples with the oranges.’

c.  pag - pa - 0pag - √

If, however, the Theme Topic form of the productive causative is taken (meaning that the causee is the Subject/Topic), then the lexical causative morpheme reappears, but the productive causative pag- is dropped.

(304)Productive Causative (Causee = Topic)

papagsamahin ‘Y is made to combine X by W’

?? + pa + pag + sama + in

PC + E + LC + root + ThemeTopic

(305)a.  Pinapagsama ko si Maria ng mansanas sa dalandan.

PST.PA.PAG.be.with GEN.1S SI Maria ACC apple SA orange

‘I had Maria combine the apples with the oranges.’

b.  0pag - pa - pag - √

As we have seen, a similar phenomenon occurs with the lexical causative alone. In the Actor Topic form of the verb, the lexical causative is overt as in (306a). In the Theme Topic form, however, the lexical causative morpheme drops.
LEXICAL CAUSATIVE

a. ACTOR TOPIC
   \textit{pagsama} ‘X brings along Y’
   pag + sama
   LC + ROOT

b. THEME TOPIC
   \textit{samahin} ‘Y is brought along by X’
   ?? + sama + in
   LC + ROOT + Theme Topic

(307) summarizes these facts and correlates the pag- drop with the overt realization of arguments. \textit{Pag1} is the lexical causative while \textit{pag2} is the productive causative. Agt1 is the Agent of the lower (or sole) verb, Agent2 is the Agent of the productive causative. Note that when Agt1 is external, \textit{pag1} is overtly realized. When Agt2 is external, \textit{pag2} is realized. This has been captured by Ross (1993) by saying that the Agent that is promoted to the subject position must be related to an overt cause morpheme.

(307) a. AT: lexical \pag{1} - \checkmark \quad \text{(Agt1 external)} \quad \text{(Th in place)}
   b. TT: lexical \textit{0pag}1 - \checkmark \quad \text{(Th external)} \quad \text{(Agt1 in place)}
   c. AT: productive \textbf{pag}2 - pa - \textit{0pag}1 - \checkmark \quad \text{(Agt2 external)} \quad \text{(Agt1 in place)}
   d. TT: productive \textit{0pag}2 - pa - \textbf{pag}1 - \checkmark \quad \text{(Agt1 external)} \quad \text{(Agt2 in place)}

(308a) is the most telling example. With this form of the verb, the Theme of the lower predicate moves to the matrix subject position. Since neither Agent has become the subject, neither \textit{pag}- is realized.
(308)a. Pinabuksan ko kay Pedro ang kahon  (S&O: 328)  
PST.PA.open GEN.1S KAY Pedro NOM box  
‘I had Pedro open the box.’

b. 0pag - pa - 0pag - √

The generalization, then, that we want to be able to capture is the relation of syntactic movement of an argument to the subject position in Tagalog and the appearance of the related pag- morpheme.

Ross’s observation is that when the relevant Agent moves, then the related pag- morpheme is overt (309a). Looking at it a bit differently, when the relevant Agent has not moved, the morpheme must be covert (309b). To put this in terms of a filter, we could formulate the generalization as (309c).30

(309)a.  [ tAGENT [ pag- ]]
   b.  [ Agent [ 0pag- ]]
   c.* [ Agent [ pag- ]]

This, in fact, is reminiscent of the Doubly Filled Comp filter in English, which rules out a relative pronoun from appearing with the complementizer that accounting for the following pattern.

(310) a. * the children [ who [ that [ I know t ]]]
   b.  the children [ who [ e [ I know t ]]]
   c.  the children [ e [ that [ I know t ]]]
   d.  the children [ e [ e [ I know t ]]]

Sportiche (1990, 1998) generalizes this restriction to other cases of Spec, Head realization and proposes a Doubly filled Voice Filter.

30In fact we would also have to rule out the possibility of having both the head and the SPEC empty. I assume that this is due to a problem of recoverability of information.
In his paper, this filter is intended to account for language variation in clitic doubling. If a language does not allow clitic doubling, then in that language both the Spec and the head (clitic) encode some relevant property, perhaps Case. How exactly this prohibition on double realization is achieved, it seems that the overt realization of *pag-* above should be part of the same phenomenon. The overt realization of *pag-* then, is sensitive to what is in its Spec position. If the Doubly Filled Voice filter is part of syntax, it seems that *pag-* drop must also be part of syntax. Further, as I assume that externalization of arguments in Tagalog is a syntactic rule and not a lexical rule (vs. e.g. Travis and Williams (1982)), one could say that *pag-* drop is sensitive to a syntactic rule.

Of course there is always an alternative. One could always say that *pag-* is optionally added in the lexicon creating all of the possible combinations. Once the form was introduced in syntax, however, and the features that *pag-* was generated with were checked in the relevant heads, then something similar to (311) could be applied checking the contents of Specs and the feature content of heads, all with the same effect. In other words, what I am claiming to happen in syntax could, with some technology, be applied in the lexicon. Further, Anderson (1974) discusses similar cases, which he calls ‘disagreement’ where agreement morphemes in Abkhaz are sensitive to the position of the relevant arguments. If the argument is adjacent to the verb, the agreement is deleted. If the argument is not adjacent (for example, it has scrambled or there in an intervening adverbial), then the agreement form must be realized. Presumably this too can be captured in the morphological component. I stand by the claim, however, that this Spec,

---

31 A concern I have is that this sort of doubling or lack thereof occurs with functional categories and I have been arguing that *pag-* and what it stands for is a lexical category. Others, however, such as Bowers 1993, Chomsky 1995, Harley 1995, and Kratzer 1996 would base-generate the subject in the Spec of a functional category. These facts from Tagalog could be used as an argument in favor of their view of phrase structure and against mine. I nevertheless maintain that V₁ is a lexical category as discussed in Chapter 1xx.
Head effect looks syntactic enough to at least be suspicious that L-syntax is part of the computational component. We will see other examples later of lexicon-like behavior of syntax, but now I discuss an alternative syntactic account for the appearance of pag-.

### 6.4.3 Pag- as an anti-EPP morpheme

Rackowski (2002) presents a very different view of the function of pag- that is quite difficult to distinguish in effect from the one presented here. As noted above, the distribution of pag- can indicate what acts as the highest syntactic argument (the highest argument of the event introduced by pag-) or what is not acting as the highest syntactic argument (any other, necessarily lower, argument). I have chosen to follow the first direction, Rackowski follows the second direction. In Rackowski’s account, pag- is in the head Voice, which is just above v in her structure. It alternates with a zero morpheme that has an EPP feature which forces movement of the closest DP that it c-commands. This means that when pag- is absent (i.e. the morpheme is zero), a DP other than the highest semantic argument will have been moved above this highest argument. It is this other argument (not the Agent), that will behave as the highest syntactic argument. One of the behaviors of this highest syntactic argument, for Rackowski, is that the verbal morphology will agree with its function, accounting for the voice morphology on the verb. The tree below gives the flavor of her account.

(312)

```
TP
  /\  
 T  VoiceP
   /\    
 agreement morph  DP_i VoiceP
     /\      
   Agent Voice'
      /\     
     Voice vP
        /\  
       0 +EPP ... ti ...
```

As mentioned earlier, it is very difficult to find empirical differences since one account focuses on what has moved, and the other focuses on what has not moved. I
provide one set of data that may be used to support my account of pag- deletion. I however concede that the EPP account it an interesting alternative.

In Chapter 2, section 2.5.3xx, we have seen cases of nominative-3rd (N3) languages, such as Kalagan and in Chapter 3, section 3.4xx, I suggested an account using partial A-movement. Rackowski’s account of Tagalog depends on the highest syntactic argument moving to a position above the highest semantic argument when pag- is not present. The actual position of movement is difficult to determine for Tagalog, which has fairly free word among the elements that appear post-verbally. Recall that N3 languages have the following word orders (e.g. Pangasinan adapted from Mulder and Schwartz 1981:244) where the DP in bold is the subject.

(313) Pangasinan: V – (Actor) – Subject
   a. V Act Pat Rec Ben Instru Loc
   b. V Act Pat — Rec Ben Instru Loc
   c. V Act Rec Pat — Ben Instru Loc
   d. V Act Ben Pat Rec — Instru Loc
   e. V Act Instru Pat Rec Ben — Loc
   f. V Act Loc Pat Rec Ben Instru —

There are two ways in which it would be difficult to extend Rackowski’s account to N3 languages: (i) the verb will not be agreeing with the closest DP and (ii) the DP will not have moved across the external argument suggesting that it cannot be an EPP feature that is at work. Note that the verbal morphology in this language agrees with the DP in bold. I refrain from extending Rackowski’s account to these data, but some non-trivial changes would have to be made to account for the similarity both in the distribution of pag- and the voice morphology on the verb.

Now I turn to some questions concerning L-syntax and more generally the role of the lexicon.

6.5 L-SYNTAX AND THE LEXICON

The phrase structure that I have been arguing for is given in (314).
In this section of the chapter I have been arguing that what happens below E is similar but different from what happens above E. The same morpheme may be added both below E and above E. Parts of the effects of this morpheme addition are the same. The $V_1$ morpheme may add an extra Case and an extra external argument. Parts of the effects of this morpheme addition are different. The productivity, the phonology, and the meaning may not be predictable below E but are expected to be predictable above E. The similarities I want to capture by saying (a) that it is the same morpheme, and (b) these morphemes are added in the syntax. The differences I want to capture by saying that below E we find a syntax that is very lexical in nature — L-syntax. Now I want to look more closely on why there should be a difference, and what the extent of L-syntax can be.

We have seen above that productive causatives are constructed in S-syntax while lexical causatives are constructed in L-syntax. Further, it has been shown in the research on causatives that productive causatives often encode two events while lexical causatives encode only one event (see e.g. Fodor 1970, and Shibatani 1972, 1976). Further still, we can see in many languages that, at least on the surface, productive causatives are always morphologically complex while lexical causatives can be monomorphemic. Putting all of
this together, I propose that the limit of L-syntax is the same as the limit for a lexical entry, which is the same as the limit for one event. Carter (1976) investigates what the limit on a ‘word’ should be. By worrying about what information a word can contain, we also worry about how ‘big’ a word can be. For instance, one of the restrictions that Carter proposes is given in (315a) below (Carter 1976: 31 (16)) and one of his observations is given in (315b) (Carter 1976: 39 (k)).

(315)  a. there exists a number n such that there is no verb in the lexicon to which we are led to assign a SR [semantic representation: LDT] with more than n occurrences of “CAUSE”

b. there is no verb paraphraseable as ‘to verb1 to verbk ...’ except where verb1 is ‘cause’

I will claim that the number n is 1 and therefore the largest number of verbs in a lexical representation of a verb is 2.\[^{32}\] These two verbs will correspond to V₁ and V₂ in the trees that I have been presenting.\[^{33}\]

In English, the lexical causative clearly consists of one word and the productive causative consists of two words. In Malagasy and Tagalog, the demarcation between lexical and productive causative is not so clear since both types of causative morphemes are affixal. There is something, however, that distinguishes the lexical causative and the productive causative and that is the head E. I claim, then, that the position of E demarcates the edge of an event and therefore the edge of a word in Carter’s terms (in some sense to be determined later). E binds the event variable in V₁P, but this only makes sense if we understand what V₁ represents. For those who share the assumption that phrase structure and event structure are related, V₁ often introduces some causal element. For those who believe that subjects are internal to the VP, the Spec, V₁P introduces the Agent argument. Work that studies lexical entries in terms of lexical

\[^{32}\]Carter allows for two causes and therefore three verbs. He needs to do this to allow for four argument verbs such as trade (W trades X to Y for Z), though he acknowledges that this sort of verb is quite restricted (Carter 1976:34). I do not have a proposal for how to handle these predicates but still want to stay with a more restrictive system.

\[^{33}\]With two Vs, I also assume that the limit on the number of arguments is three. See Koizumi (1995) for a similar conclusion.
decomposition (e.g. Carter 1976, Dowty 1979) recognize *cause* as the highest possible predicate\(^{34}\). Work that studies lexical entries in terms of theta-grids recognizes Agent as the highest possible theta-role in any theta-role hierarchy (Baker 1988, Grimshaw 1990, Larson 1988). Further, not only are *cause* and Agent the highest predicate or theta-role, respectively, in a lexical entry, they are unique in any lexical entry. In other words, no lexical entry can have two *causes* nor can a single theta-grid contain more than one Agent. This has the result that once a *cause* predicate has been introduced in a lexical entry, or an Agent theta-role added (if we think of constructing a lexical entry from the bottom up), the lexical entry must be complete. In terms of the tree being discussed, once \(V_1\) has been added, no more lexical categories may be added (since no more predicates can be introduced). Therefore, \(E\), by virtue of its position as the binder of the event variable in this top \(V\), marks the edge of a lexical entry, i.e. the edge of the domain of the lexicon. After this, going further up the tree, any more lexical categories must contain an independent lexical entry. As such, \(E\) also marks the boundary between \(t\)-syntax and \(s\)-syntax.

We now have an explanation for why the examples in (301) are ungrammatical. The Agents of *laugh* and *sneeze* must have been introduced by \(V_1\). If an additional argument is to be introduced, then, it must be done via an additional lexical item added above \(E\). In English, such a lexical item would be the productive causative *make*. While we have seen that lexical items that appear monomorphemic (such as *melt\textsubscript{trans}*) are in fact morphologically complex with zero morphology, no productive causative morpheme is consistently represented by a zero morpheme.\(^{35}\) This would explain Carter’s claim concerning the limitations on what can be encoded in one ‘word’.

Now we can see how the lexical causatives and the productive causatives are distinguished in syntax. This is very clear in Tagalog and Malagasy where the morphology is much more transparent. While only one causative morpheme exists in each of these languages, it can serve as both the lexical causative or the syntactic causative depending on where it occurs in the phrase structure. If it occurs below \(E\), it is

---

\(^{34}\) Or *do* if an Agent may be introduced by a *do* predicate. It may be that *do* when selecting a become predicate is *cause*.

\(^{35}\) We have seen, however, that Tagalog productive causative *pag-* is, in certain configurations, realize as zero, but it does have an overt form.
part of L-syntax and is the lexical causative. This is because it is part of a lexical entry and as such shows the idiosyncracies of lexical entries. If it is above E, then it must be attached to the stem via S-syntax and it represents a lexical item on its own. This explains its productivity and predictability. The position of the morpheme is easy to determine in these languages due to the fact that E is lexically realized. A causative morpheme appearing closer to the root than the E morpheme will be a lexical causative and a causative morpheme appearing further from the root than the E morpheme will be the productive causative.

(316) \[ V \quad - \quad E \quad - \quad V \quad - \quad \checkmark \]
Malagasy: \[ an \quad - \quad f \quad - \quad an \quad - \quad \checkmark \]
Tagalog: \[ pag \quad - \quad pa \quad - \quad pag \quad - \quad \checkmark \]
\[ \text{PC} \quad \text{LC} \]
6.6 SUMMARY

The goals of this section were two-fold. I want to contribute to the discussion of L-syntax as a definable sub-module of the grammar, more particularly a sub-module of the syntactic component. I also want to show the importance of looking at a wide variety of languages when investigating these questions. In English, where many morphemes are arguably zero, it is often hard to find evidence for abstract heads (like E) or operations (like lexical causativization). Other languages often provide the needed evidence for these heads or processes. By looking carefully at causativization processes in Tagalog and Malagasy, I have proposed that the differences between lexical causatives and productive causatives are determined not by choice of morphemes since the same morpheme is used for both processes in each of these languages. The difference stems from the position of these morphemes in the tree. Further, I argue that the difference in these positions is easily determined in these two languages because of an intervening morpheme which I assume indicates the existence of an event related head marking the boundary between lexical processes and syntactic processes. I present a picture of syntax in general and the interaction of L-syntax and s-syntax as shown in (317).

(317)

```
(317) VP3DP V'S-syntax3###V EP    Lexicon===========3E VPL-syntax 3DP V'3VASPP  3ASPVP  3 DP V' 3 V PPSyntax
```

draft: 8/10/06
I claim that the lexicon and the computational component are allowed to overlap up to a point. In terms of phrase structure, that boundary is the event related head which I have labelled E. In terms of semantics, the lexicon can encompass, at most, one event. In other words, a lexical entry may refer to any of the lexical head positions that occur below E since those head positions encode sub-parts of a single event. It is harder to argue that this must be a case of the lexicon exerting an influence on the syntax rather than the syntax exerting an influence on the lexicon. One of my main reasons for taking the direction that I do has to do with the productivity and predictability of the event related categories that appear within the domain of L-syntax. Hale & Keyser (1993) make it very clear that functional categories cannot be part of L-syntax. They write “no functional categories are involved in the verb formation processes at issue here … no functional projections are present at points internal to the domains defined by lexical entries” (H&K 1993: 98). But here we encounter a problem with forms such as the Navajo forms we saw in Chapter 3 (section 3.2.2xx). In the template given for the order of morphemes in Navajo, functional material is interspersed with the lexical material. Further, the lexical material has the idiosyncratic earmark of L-syntax. To repeat an example from Chapter 2, below I give the morphological make-up of the verb meaning ‘to pray’. It consists of three parts — two prefixes, which, according to Speas (1990: 208), are not productive, and a stem which cannot occur on its own.

\[(318) \quad so \ldots di \ldots zin \quad \text{‘to pray’} \quad 1 \ldots 6 \ldots \text{stem}\]

These subparts of the lexical entry occur in particular places in the template as indicated by the numbers given to the right of the entry above. In between these parts of the lexical entry appear such inflectional type elements such as aspectual markers, tense, and object and subject agreement. If the inflectional like material cannot appear in the lexicon, the solution is to allow the lexical formation of the lexical entry to be formed in the syntax.

This view of syntax has drawbacks mainly having to do with semantic and lexical idiosyncrasies. For example, the semantics within this component is often not
compositional. The whole does not entail the subparts, so, for example, the (a) utterances below do not entail the (b) utterances in English and Tagalog.³⁶

(319) a. Nagsabog ng bato ang magsasaka
  PERF-PAG-sabog ACC stone NOM farmer
  ‘The farmer scattered the stones.’

  b. Sumabog ang bato
  AT-PERF-sabog NOM stone
  ‘The stone exploded.’

(320) a. The make-up artist reddened the movie star’s cheeks.

  b. The movie star’s cheeks reddened.

In order to account for those idiosyncracies, I am proposing that syntax has recourse to the lexicon once the structure of an event has been complete. While head movement may continue to form longer words beyond this domain, they are not the ‘words’ that Carter investigated the limits on. To distinguish between the two types of words, we can label one set E-words (event words) and the other M-words (morphological words). Clearly the two are not necessarily represented through a one to one mapping. We can see summed up in the table below that boundaries of M-words and E-word may vary.

³⁶ The view of modularity internal to syntax that I have sketched has much in common with Marantz’s work (e.g. Marantz 1997, 2001). Marantz also develops a view of syntax that includes an idiomatic/idiosyncratic component. I leave it to the reader to compare the two views.
(321) M-words vs. E-words

<table>
<thead>
<tr>
<th>1 M-word</th>
<th>2 M-words</th>
</tr>
</thead>
</table>
| English: wash | Edo: naki kiri  
|          | ‘kill’  
|          | Fongbe: kú drç  
|          | ‘dream’  
| Malagasy: m-an-f-an-sasa  
|          | ‘make wash’  
| Tagalog: m-pag-pa-0-sama  
|          | ‘make combine’  
| English: make wash |

The category E limits the domain of an E-word and one event. We can see why English is not the best language to study when investigating these boundaries since English generally has a one-to-one mapping. English does not have M-words that go beyond the E-domain while Malagasy and Tagalog do. Further, English does not have multiple M-words within one E-word while Edo and Fongbe do as we will see below.

Chomsky (1998:12-13) dismisses the possibility that the lexicon be accessed more than once as being uneconomical. He uses a metaphor to clarify the notion of operational complexity.

Suppose automobiles lacked fuel storage, so that each one had to carry along a petroleum processing plant. That would add only bounded ‘complexity’, but would be considered rather poor design. Something similar might well be true for language.

In applying this to the question of access to the lexicon, he writes,

The obvious proposal is that derivations make a one-time selection of a lexical array LA from LEX[icon], then map LA to expressions, dispensing with further access to LEX. … If the derivation accesses the lexicon at every point, it must carry along this huge beast, rather like cars that have to replenish fuel supply constantly.

37 A different question arises with words that represent features on a root like tense. This sort of syncretism would be dealt with differently.
Of course, Chomsky also points out in the discussion that ultimately “the issues are empirical and can be settled only by investigating consequences of alternative conceptions.” Given the idiosyncratic nature of L-syntax, it is clear that the lexicon has to be involved with the output of the lexical item which encodes the final argument (e.g. Agent) or the final predicate (e.g. DO, or CAUSE). If the lexicon can be accessed only once, then the process of L-syntax in its entirety has to occur within the lexicon. We lose on several counts, however, if this is the case. We lose on the generalizations that would now appear in two different grammatical components — in the computational system and in the lexicon. The processes that we have seen are restrictions on head movement, subject sensitivity of anaphors, and restrictions on Spec and head positions that are filled simultaneously. We also lose generalizations provided by a theory of multifunctionality that allows lexical items (such as an- in Malagasy and pag- in Tagalog) to appear at different positions on a syntactic tree, with differences in function being derived from differences in position. Finally, we lose the flexibility of allowing inflectional type material from appearing internal to lexical entries as we have seen is the case in Tagalog (aspectual reduplication) and Navajo. As is often the case, we have found the computational system to be less than optimal. However, we can contain the domain within which this less than perfect design must function.

6.7 DISCONTINUOUS LEXICAL ITEMS

The cases we have seen for L-syntax above all involve predicates that are encoded in one M-word, created through head movement. It is fairly easy to argue that M-words can contain bits of syntax as implied by Baker (1985) and argued for explicitly in Baker (1988). The clearest cases of this are words that are created in the s-syntax as they are morphologically productive and semantically compositional. Let us take the example of future tense in English versus French. In English, the string will eat is represented by two syntactic heads. In French, it is less clear that the morphological word mangerai ‘will eat (1SG)’ represents two (or more) separate syntactic heads, but one can make the conceptual argument that, languages being more similar than different, the French affixes act similarly to the separate words in English. The French morphological word
mangerai, then, is also represented by (at least) two heads, T and V.\textsuperscript{38} We have used similar argumentation to claim that the Malagasy morphological word \textit{mpanasa} ‘make wash’ should have the same syntactic representation as its English translation, which is represented by (at least) two syntactic heads. In both the French future and the Malagasy productive causative, the morphology is productive and the meaning is compositional. This combined with the language variation makes a syntactic analysis appealing.

It is harder to argue for a syntactic account of phenomena which are less productive and where the meaning is not compositional (such as \textit{magsabog} ‘to scatter’ in Tagalog). It becomes harder still to argue for syntactic complexity for a form that appears to be morphologically simple (such as \textit{kill} in English). However, as in the cases of \textit{s}-syntax, we can find cross-linguistic variation that provides support for a syntactic analysis. For this reason I turn to other languages where one \textit{e}-word can be represented by two \textit{m}-words, in other words, languages that appear in the top righthand corner of the table in (321).

I have been arguing for an articulated VP structure. While one argument for this structure comes from the position of derived elements within the VP, many of the other arguments come from verbal morphology such as reduplication in Tagalog and morpheme order in Navajo. In this view of phrase structure, the VP contains more than one lexical head and we have seen these heads filled with different morphemes in different languages. In Tagalog, \( V_1 \) was filled with \textit{pag-} and \( V_2 \) with the verb root while a reduplicative morpheme could be attached between the two. In Navajo, a verb like \textit{so…di…zin} ‘to pray’ had parts that could be separated by aspectual type material. If all of these heads do exist, however, we might expect to see them filled with free-standing words as well. In this section, I present some plausible examples of languages that do fill these heads with separate words. As pointed out by Hale and Keyser (1993), given various views of articulated VPs and especially their own view of the projection of arguments, such languages are expected to exist (Hale and Keyser 1993:96).

In reality, all verbs are to some extent phrasal idioms, that is, syntactic structures that must be learned as the conventional “names” for various

\textsuperscript{38} I am representing the minimum number of heads. In my phrase structure, \textit{mangerai} contains many more heads..
dynamic events. That is our view of the matter, in any event, and it seems to be forced on us by the very framework we are considering. Moreover, it is not without empirical support, at least at the observational level. In many languages a large percentage of verbal lexical items are overtly phrasal (e.g. Igbo, Nwachukwu 1987); in others a healthy inventory of “light verb” constructions represent the class of overtly phrasal lexical entries (e.g. Japanese, Grimshaw and Mester 1988; English, Kearns 1988); and in still others (e.g. the Tanoan languages, including Jemes, Tewa, and the Tiwa languages), the verbal lexicon contains an extraordinary number of entries whose morphological make-up is overtly the result of incorporation. To be sure, many languages boast a large inventory of simple monomorphemic verbs. But our guess is that most, probably all, superficially monomorphemic verbs are lexically phrasal, possessing a structure that is syntactic…”

In the next two sections, I present two language types with phrasal lexical items. First I discuss serial verbs, then inherent complement verbs.

### 6.7.1 Serial Verb Constructions

In Navajo we have seen an extreme case where a lexical item seems to be split up over several non-adjacent morphemes in a morphologically complex structure. Serial verb constructions are cases where a lexical item can be seen to be split up over several non-adjacent words in a syntactically complex structure. Here I present the possibility that serial verb constructions exemplify a case where we can see the articulated VP in the syntax. In particular, I suggest that serial verb constructions are the most obvious place to find V2Ps being realized independently. This view of serial verb constructions follows up on the analyses proposed by Baker (1989, 1991) and Larson (1991). 39

To start with the theory rather than with the data, we can ask what properties a V2P standing alone would have. As we have seen in Chapter 4, a single VP generally characterizes and (end)state. Therefore, one property that V2P would have would be stativity. Further, given that no higher VP will be projected, the external theta-role in the
sense of Williams (1981), generally the Agent, will have to be satisfied through an alternative to regular theta-role assignment. Finally, since the traditional external theta-role cannot be assigned, some other theta-role will appear to be the highest. These properties are summarized below.

(322) Properties of V₂P

(a) Stative
(b) External theta-role (Agent) satisfied by different means
(c) Internal theta-role (Theme) appears external

Turning to the relevant data, we can see that all of these properties can be found in a certain set of SVCs. Further, by assuming that the second (and in these cases) last V in a SVC is, in fact, an instance of V₂, we are able to use an analysis of SVCs proposed by Larson (1991) while avoiding the problems that this analysis raises as pointed out by Baker (1989).

Larson (1991) suggests that SVCs of the type given in (323a, b) below are like resultative predicates in English (323c).

(323a. (English based Creole of Surinam)

Kofi naki Amba kiri⁴⁰ (Larson 1991: 10)
Kofi hit Amba kill
‘Kofi struck Amba dead.’

---

⁴⁰ The use of the verb ‘kill’ here rather than ‘die’ is an interesting cross-linguistic distinction which I put aside here.
b. Fon (West African Kwa language)

\[
\text{Koku sO asO do tavo-ji} \quad \text{(Larson 1991: 7)}
\]

Koku take crab put table-on
‘Koku put the crab on the table.’

c. Black Flag kills bugs dead. \quad \text{(Larson 1991: 20b)}

The analysis that Baker (1989) proposes for a SVC as in (323a) is given in (324) below.\textsuperscript{41}

\[
\begin{array}{c}
\text{VP} \\
\text{V} \\
\text{DP} \\
\text{naki} \\
\text{Amba} \\
\text{kiri} \\
\text{V}'
\end{array}
\]

Baker tries to formally capture the following characteristics of SVCs. If the first V has an object, this object is shared by the second V. Further, the second V must assign an internal theta-role to this object. Baker’s phrase structure captures this observation since the DP between the two Vs is structurally internal to the projection of both of the Vs. The Projection Principle, then, would force both Vs to assign a theta-role to the DP. UTAH will ensure that both of these theta-roles are internal theta-roles (under the assumption that external arguments must be external to the VP).

Larson (1991) views this differently. It is clear that Baker must make some innovative assumptions concerning X’-theory to allow structures such as the one in (325) since the VP will contain two head Vs. Larson suggests that these innovations are unnecessary. He, instead, proposes the following structure for the same string, relating it to resultative predicates, which he would give the structure in (325).

\textsuperscript{41} In the discussion of older accounts, I update the terminology using DP instead of NP.
Just as the internal NP *her finger* in the resultative structure is “shared” by both the main V *rub* and the secondary predicate *raw* (i.e. both assign theta-roles to it) in (326), so is the internal NP *Amba* in the SVC shared by both Vs in (325). In both constructions, the main V assigns the primary theta-role, and the resultative predicate, be it a VP or an AP, assigns a further theta-role.

Baker (1989) presents an interesting argument against Larson’s analysis of serial verb constructions. As he points out, secondary predicates in English, for example, are APs and PPs, and while they assign an additional theta-role to the object of the transitive main verb, they assign an external theta-role to this position, not an internal theta-role. Taking our original SVC example in (323a), we see that the second verb *kiri* ‘kill’, the one we are trying to relate to a secondary predicate, assigns its internal theta-role, Theme, to the shared object. Its external theta-role, Agent, is co-assigned with the main verb to the subject position. This is shown schematically below.
Baker points out that a true parallel between a SVC and a secondary predicate would appear in examples such as the following, which are clearly ungrammatical.

(328)  

a. * I locked the house in.  

= I locked the house such that I was in the house.

b. * John tested his son proud.

= John tested his son, and, as a result, John was proud of his son.

In each case, the external argument of the secondary predicate is co-assigned with the main verb to the subject of the sentence, and the internal argument of the secondary predicate is co-assigned with the main verb to the shared object. In other words, in (328a), the shared subject I is the external argument of both lock and in, and the shared object the house is the internal argument of both lock and in, giving the intended reading that I locked the house and I was in the house.

Obviously, this sort of construction is not possible. Given the proposal that V₂P can act independently, however, we can explain this difference, not by changing how we look at secondary predication, but by changing how we look at SVCs.

I will assume Larson’s structure given in (325) with the additional claim that the VP of the secondary predicate is crucially a V₂P. Why must it be a V₂P? This is explained by the first property of V₂Ps given above. The resultative must be a state. This is clear in the English examples where the resultative is encoded by an AP or a PP,
archetypal states. The requirement that the result is a state, then, forces a verbal projection to be restricted to V₂Ps. Once this structure is set up, the rest follows since the theta-assignment properties of V₂P more closely follow the theta-assignment properties of PP and AP. The “external” argument (i.e. the highest argument) of the V₂ will be an internal argument within its theta-grid. In other words, as we saw for Larson’s tree, the Theme is like the inner subject, external argument of V₂.

How exactly this shared theta-role is assigned is one of debate. There are two possibilities that I see. One is that PPs and APs, when acting as predicates, do not assign their external arguments to their Spec positions, but rather, as a maximal projection with an unsaturated theta-role, they assign this theta-role through predication as in Williams (1980). Another possibility is that there is an empty category in the Spec position. Collins (1994) argued for the existence of an empty category in SVCs using agreement facts to support his claim.

The topic of SVCs is an enormous one and deserves a book of its own. The conclusion to be drawn now for the purposes of the present chapter is that the separate parts of the articulated VP can occur as independent lexical items.

6.7.2 Inherent Complement Verbs.

Another case where the articulation of the VP can be seen in the syntactic configuration of the VP is in VPs that contain inherent complement verbs (ICVs). Some examples of this construction are taken from Fongbe (see Avolonto 1995: 72ff) are presented below. In (330a) we can see where the verb kú, used on its own, means ‘to die’. However, this same verb, when in construction with the N drɔ, ‘dream’, means ‘to dream’. In (331), depending on what the direct object of the verb dó is, the sense of the VP changes

---

42 Note that a change of state predicate cannot be used as the following contrast shows (see Embick 2004).

(i) I cracked the egg open/*opened.

Prepositions, however, seem to have to show motion giving the different between ‘I put the book onto the table’ and ‘*The book is onto the table’; ‘I pushed the table to/*at the wall’ and ‘The table is *to/at the wall’. I have no explanation for this.

43 This raises the question of why stative verbs like ‘know’ are not acceptable in such constructions. Transitive stative verbs arguably have a more complex structure (see, e.g. Noonan 1992).
completely. In (331a), the verb takes a normal DP complement, while in (331b), the verb is merged with an inherent object.\textsuperscript{44}

(330)a. \textit{gbɔ́ kú}

sheep DET die

‘The sheep died.’

b. Kɔ̀kú kú drò

Kokou \textit{ku} dream

‘Kokou dreamt.’

(331)a. Àsibá dó gbàdé

Assiba sow corn

‘Assiba sowed some corn’

b. Àsibá dó wèzùn

Assiba DO race

‘Assiba ran.’

(332)a. Kɔ̀kú xà àkwɔ́é

Kokou count money

‘Kokou counted money.’

b. Kɔ̀kú xà yèwhè

Kokou \textit{xà} prayer

‘Kokou prayed.’

Other examples of ICV constructions are given in (333). Here the same verb can take a variety of the inherent objects with a concomitant change of meaning.

\textsuperscript{44} All examples are taken from Avolonto (1995) but the English glosses and translations are my versions of
In these constructions, while the constituent parts, the verb and in particular the inherent object, lend some meaning to the whole, the specific meaning of the expression is non-compositional. In comparing (333b) and (334a), it appears that dó is causative while dí is stative. This contrast, however, is not found in the comparison of (331b) and (333a). As suggested in the passage below describing the same construction in Igbo, given in Avolonto (1995: 71) and credited to Nwachukwu (1987:22 and 1985:61), these seem to be fixed expressions that must be learned and stored as such.

An inherent-complement verb is one whose citation form is obligatorily followed by a meaning-specifying noun complement ... Because it is lexically specified as part of the verb, the inherent complement is by definition strictly obligatory ... and any dictionary entry which excludes the complement is so ambiguous as to be meaningless 1987: page 40.
Avolonto clearly shows that these nominals do not behave syntactically like objects. They cannot undergo wh-movement, cliticization, relativization, or clefting. Examples from the discussion on wh-movement are given below. The construction in (335a) can only be a question formed from the construction in (331a), and could not be used to ask for the content of a ICV construction using the verb dô, such as those given in (331b), and (334a,b) above. Likewise, the question in (335b) can be used to form a question about (332a) but not (332b).

(335)a. été Àsibá dô  
   what Assiba sow 
   ‘What did Assiba sow?’

b. été Kôkû xà 
   what Kokou count 
   ‘What did Kokou count?’

Cliticization and relativization work similarly, which is not surprising given that the N has no referential content. Clefting, however, has an interesting twist. Avolonto first shows that there is clefting of objects and of verbs, and that when the object is clefted, there are three possible interpretations. In (336) below, we see a case where an object is clefted. The three interpretations are given. In essence, the focus can be on the object alone (i), the verb and perhaps the object (ii), or the whole VP (iii) (Avolonto 1995: 83).

(336)àsâñ wè Zuléma xò 
  crab FOC Zulema buy 
  i. It is crab that Zulema bought (and not bread)  
  ii. It is buy crab that Zulema did (and not prepare crab/rice)  
  iii. It is buy crab that Zulema did (and not prepare rice)

When the verb undergoes predicate clefting, there is only one reading that involves the verb meaning alone. This is shown in (337) below (Avolonto 1995: 83).
It is buy crab that Zulema did (and not prepare crabe/*rice)

Avolonto stresses the difference in meaning between (336) and (337). This is interesting when applied to clefting in ICV constructions. To begin, note that only the “object” and not the “verb” can cleft. This is shown below in (339) where a cleft construction is being formed from the example given in (338).

(338) Kòfì dì sà
   Koffi DI walk
   ‘Koffi went for a walk.’

(339)a. sà wè Kòfì dì
   walk FOC Koffi DI
   (i) It is walk that Koffi did and not work.
       (ii) *It is a walk that Koffi did’

b. * dì wè Kòfì dì sà

This is particularly interesting given that this is the object which resisted movement in WH-constructions, cliticization and relativization. Further, the element that we have been led to believe is the lexical entry is now a discontinuous element. Finally, what appears to be the verb cannot cleft. This shows clearly that the object has to be visible to the syntax as a separate element in spite of its semantic dependency on the verb. Also Avolonto reports that the cleft has is closer to the interpretation of the predicate clefting than it is to the object clefting. One could imagine that the contrast could be set up between dì sà ‘walk’ and dì xèsi ‘fear’ by clefting the inherent object but this does not seem possible.
In sum, inherent objects are syntactically independent as can be seen in the formation of the cleft construction. However, they are have no independent semantics. Like serial verb constructions, bits of lexical entries can appear as independent words.

6.8 THE LEXICON AND LEXICAL CATEGORIES

There are two more questions that I would like to explore having to do with L-syntax and S-syntax. In the end, I will have only suggestions of answers. One question is what information exactly is contained in a lexical entry, and the other question is whether there are limits on M-words.

6.8.1 Lexical entries

Recent work has suggested that the lexicon is quite impoverished (e.g. Marantz 1997, Borer 2005). This trend started with Hale and Keyser (1993) when much of the lexicon was argued to be, in fact, part of syntax. If argument structure can be derived from the syntactic structure, one can ask what is left in the lexicon. In fact, the picture that I am developing here is quite conservative. It also contains a certain amount of redundancy. I assume that roots have categorial signatures. This conclusion comes from work by Demirdache and Matthews on Salish (1995) and work I have done on Malagasy (Travis In press). I also assume that full theta-grids have to be specified for roots. I argue in Chapter 7 that certain theta-roles are assigned by structural configuration and some from the lexical specifications of the root. Clearly, because there is a distinction, not all theta-roles can be created by syntax. We have also seen that the addition of the external argument through lexical causation can change the meaning substantially. In other words Agent + \text{EXPLODE} doesn’t mean $X$ explodes $Y$ in Tagalog. Therefore, there has be some return to the lexicon for the details of the semantics of these forms. While I don’t have the whole view of the lexicon fleshed out, it still houses a fair amount of information some of which is redundant with the computational component, not unlike older views. In fact, since must of the lexical information, in my view, is scattered over syntactic heads, the overlap is substantial.
6.8.2 M-words

In the discussion above, I have suggested that there is a limit on the domain of an E-word. One could ask whether there is a limit on the domain of an M-word. In a way, this is the question that Li (1990) addresses in his research on the constraints on head-movement. Li claims that improper movement includes not only XP movement from an A position to an A' position back to an A position, but also X0 movement from a lexical head to a functional head back to a lexical head. If Li’s generalization is derivable from a principle of grammar, the phrase structure I propose in this book runs into serious problems. In Chapter 3 I argue that there is movement from V (a lexical category) to ASP (a functional category) back to V.

(340) Tagalog Aspect: \( V \Rightarrow ASP \Rightarrow V \)

\[
nagtutumba: \text{n-} m- \text{ pag-} \text{ RED-} \sqrt{\text{tumba}}
\]

\[
  V_1 \quad \text{ASP} \quad V_2 \quad \text{IMPERFECTIVE 'is taking out'}
\]

In this chapter, I propose that causatives in Malagasy involve head movement from V to E to V.

(341) Malagasy causatives: \( V \Rightarrow E \Rightarrow V \)

\[
0 + m + an + f + an + \sqrt{ala} = \text{mpanala} \quad \text{'to make x take y out'}
\]

\[
T \quad E \quad V_1 \quad E \quad V_1 \quad V_2
\]

While on one hand, the presence of an intermediate functional category looks problematic, on the other hand, the presence of such a category may solve other problems. In Baker (1985: 374), an example of a causative in Chamorro shows that agreement may occur between the causative morpheme and the verb stem.\(^{45}\)

\(^{45}\) Transcription and glosses are from Baker (1985).
In (342) we see the plural agreement marker *fan-* occurring between the causative marker *na'* and the lower verb which has been passivized.\(^{46}\) If agreement is to be encoded through Spec-head agreement or an AGREE relation with a functional category as assumed above, then the existence of agreement here indicates the existence of a non-lexical category between the causative verb and the lower verb. While this may appear to be counter-evidence for Li (1990), his otherwise robust general observation that inflectional type material does not occur within causatives needs an explanation. We have seen that the type of phonology in Malagasy changes at E. Further, Li observes that there are no cases of head-movement from V through T and C to another V. In my terms, there is no m-word that crosses C. Following ideas of Morita (in press) and Newell (2004), I suggest that these two observations are due to phases. E would represent the edge of one phase. If the material below E is sent to PF, any further morphological process must be similar to compounding. C would represent the edge of the next phase,\(^ {47}\) and there would be no morphological process to combine elements further. With this much inflectional structure, the construction would have to be periphrastic.

### 6.9 CONCLUSION

While most syntax has the blind productivity that one would want in a computational system, I have argued that one corner of syntax, L-syntax, has to be allowed one pit-stop before continuing in the derivation. An event related category, E, selects \(V_1P\) and represents the point at which the syntax may return to the lexicon. This head E delimits the edge of an event and therefore the edge of an \(E\)-word. By allowing the syntax this possibility, we allow the syntax to keep control over Spec, head relations, binding

\(^{46}\) I assume that in some languages passive morphology may appear in \(V_1\) thereby allowing a different representation of the external argument that is introduced by the V. In Chamorro, then, the passive morpheme would be in \(V_1\) just below E.
possibilities, adjunction structures, and head movement — all phenomena which arguably hold in L-syntax.

47 More exactly, any inflectional material above E would be part of the next phase so that even a TP that is selected by a V would still constitute a phase. This is similar to Li’s observation that any move back to a lexical category after T or C has been added would violate proper head movement.