Why there is an EPP*

Norvin RICHARDS
(MIT)

In this paper I develop a general theory of why certain heads require specifiers, unifying the classic EPP cases with the conditions on ellipsis identified by Lobeck (1995) and Saito and Murasugi (1990), and accounting for apparent counterexamples from VSO languages like Tagalog and Irish. The account makes use of Kayne's (1994) Antisymmetry and Chomsky's (2000, 2001) notion of phase; offending structures are shown to be unlinearizable on straightforward assumptions about linearization.

Key words: EPP, ellipsis, antisymmetry, phases, Tagalog

A number of languages, including English, are subject to the peculiar constraint illustrated in (1-2):

(1) a. It rained
   b. *Rained

(2) a. There seems to be a wombat in the room
   b. *Seems to be a wombat in the room

* Many thanks to Genara Banzon, Roberto Figueroa, Elena Francisco, Louie Locsin, Valerie Mayo, Raphael Mercado, Ryan Morales, and Reine Protacio for their invaluable help with the Tagalog facts; all Tagalog facts are from conversations with them, unless otherwise indicated. Thanks, too, to Ken Hiraiwa, Shigeru Miyagawa, and Shoichi Takahashi for some discussion of Japanese ellipsis facts. Finally, thanks to Danny Fox, Shigeru Miyagawa, John Whitman, an anonymous reviewer for Gengo Kenkyu, and audiences at AFLA 9 and at my Spring 2002 Syntax seminar for helpful comments and suggestions. None of these people should be held responsible for any errors of fact or analysis, which are my responsibility.
Chomsky (1981, 1982) coined the name “EPP” for this problem: apparently, in some languages, Spec TP must be filled, for reasons which have never been very well understood. The problem has become more salient in the framework developed in Chomsky (2000, 2001), where it is unclear why checking relations should ever involve movement at all, at least in cases where no phase boundary intervenes between the Probe and the Goal. The relevant operation for feature-checking, in this approach, is Agree, an operation which is taken to be independent of Move. Chomsky (2000, 2001) therefore generalizes the notion of EPP to include all cases in which a specifier is occupied for reasons having nothing to do with selection.

In this paper I will try to address the question of why specifiers must sometimes be occupied. I should make it clear at the outset that what I am interested in is not so much the mechanism we should use to force specifiers to be occupied; we might decide, in the end, to make use of a feature labeled “EPP”, as in Chomsky (2000, 2001), or perhaps a feature labeled “Case”, as in Castillo, Drury, and Grohmann (1999) and much other work. I am more interested in a question of “language design”, in the sense of Chomsky (2000, 2001): why is the syntax put together in such a way that Spec TP is always filled, at least in some languages? Of course, it is possible that a full understanding of the nature of the feature that drives movement to this position would give us an answer to that question. For reasons that will hopefully be apparent soon, I will take a different tack here. I will concentrate, for the most part, on the “classic” EPP which requires Spec TP to be filled in certain languages, though the hope is that this approach may be generalizable to the more universal EPP of Chomsky (2000, 2001).

Another domain in which specifiers are apparently required to appear is that of ellipsis. Saito and Murasugi (1990) and Lobeck (1995) posit a syntactic condition on ellipsis sites with more or less the following character, which I will refer to here as the Ellipsis EPP (EEPP):
An ellipsis site must be the complement of a head with a specifier.

1) In fact, their claim is that an ellipsis site must be the complement of a head which agrees with its specifier. I think there is no good reason to prefer this formulation over the one in (3). Martin (1992, 2001) and Bošković (1997) make use of the agreement requirement in their account of contrasts like the one in (i - ii):

(i) I don't know if I can understand this, but I'll try to [ ___ ]
(ii) *I don't know if John understands this, but I believe him to [ ___ ]

On this account, (i) is well-formed because to must agree with PRO in order to assign/check Null Case, while (ii) is ill-formed because to is not in an agreement relation with its specifier. As Martin and Bošković acknowledge, the well-formedness of examples like (iii) is a potential problem for their approach:

(iii) I don't know if John (really) understands this, but he seems to [ ___ ]

Following Lasnik and Saito (1992), Martin and Bošković suggest that apparent raising infinitives like the one in (iii) are actually structurally ambiguous, having the structure either of raising infinitives or of control infinitives. The example in (iii), on this approach, is well-formed because of the availability of a control parse. Crucially, ECM infinitives will have to not be ambiguous in this way, given the ill-formedness of (ii). As evidence for their approach, Martin and Bošković offer the ill-formedness of examples like (iv), where the control parse is presumably unavailable (Bošković 1997, 179):

(iv) *John doesn't believe there is likely to be any Asian team in the final game, but I believe there is likely to [ ___ ]

It is not clear, however, that the ill-formedness of (iv) has anything to do with the unavailability of a control parse. VP-ellipsis of be seems to be ill-formed regardless of the kind of infinitive involved (see also Warner 1986, Hagstrom 1994, Lasnik 1999, and references cited there for interesting discussion of related ellipsis paradigms):

(v) *John is trying to be happy, but Mary isn't trying to [ ___ ]
(vi) *I believe John to be happy, but I don't believe Mary to [ ___ ]
(vii) *John seems to be happy, but Mary doesn't seem to [ ___ ]

Moreover, if we use idioms to eliminate a control parse for putative raising infinitives, VP-ellipsis still seems to be well-formed (as long as we avoid be):

(viii) I don't know if the shit really does hit the fan every Monday around here, but it sure seems to [ ___ ]

Descriptively, then, the generalization seems to be that (i) VP-ellipsis in ECM infinitives is ruled out, and (ii) VP-ellipsis of be is ruled out. Space considerations prevent me from developing full accounts of these facts, though it seems reasonable to relate the second to the facts in (ix-x), perhaps by claiming that be, unlike other verbs, raises out of the domain affected by VP-ellipsis:
The EEPP in (3) is meant to deal with data like those in (4-5):

(4) a. I wanted to read a book, so I stole [DP John’s [ ]]
b. *I wanted to read a book, so I stole [DP a [ ]]
(5) a. Although she doesn’t know [CP how [ ]], Sue thinks John made it to work on time
b. *Although she doesn’t know [CP if [ ]], Sue thinks John made it to work on time

On standard assumptions about the syntax of DPs and of wh-movement, the EEPP is satisfied in the (a) sentences above but not in the (b) sentences.

In a framework like that of Chomsky (2000, 2001), the EEPP and the EPP bear a certain resemblance to each other, in that both involve a requirement that heads with complements of a certain kind also have specifiers. For the EEPP, as we have seen, this requirement holds for heads with elided complements; for the EPP, the requirement holds of a head which is taken to have a phase as its complement:

(6) a. XP
   \hspace{1cm} \text{EEPP position} \hspace{1cm} X' \hspace{1cm} \text{ellipsis} \hspace{1cm} YP
   \hspace{1cm} X \hspace{1cm} \text{EPP position} \hspace{1cm} T' \hspace{1cm} \text{phase} \hspace{1cm} vP

(ix) John seems to be happy, but Mary doesn’t seem to be [ ___ ]
(x) *John seems to put books on tables a lot, but Mary doesn’t seem to put [ ___ ]

In any event, it seems clear that agreement is not the relevant notion; what we need is something distinguishing between ECM infinitives on the one hand and control and raising infinitives on the other.

2) Chomsky (2001) assumes that only transitive vPs are phases; I will have to crucially assume here, following Legate (2002), that all vPs are phases.
If we believed that ellipsis sites and phases had something in common, then, we might be in a position to collapse the EPP and the EEPP. Suppose we assume that this is true; in particular, let us assume that ellipsis, like Spell-out of a phase, involves taking a part of the structure and rendering it atomic as far as the computational system is concerned, with no internal syntactic structure\(^3\).

With this assumption, we are in a position to offer the generalized EPP in (7):

\[(7) \text{ A phrase with no internal syntactic structure must be the complement of a head with a specifier.} \]

The EPP in (7) is still fairly mysterious, of course; we would like to know why it should be true. In the next section, I will investigate a class of cases in which (7) appears to simply be false, and will suggest an amendment to (7) which covers them. Next I will develop an approach to the amended version of (7) which makes it follow from general conditions on the interface between the syntax and the phonology of the type pioneered by Kayne (1994); we will see that structures which violate (7) cannot be linearized, on plausible assumptions about how linearization works.

1. **Ellipsis in VSO languages: Irish and Tagalog**

Several VSO languages appear to have a form of ellipsis that violates the condition in (7):

---

\(^3\) One version of this assumption might state that ellipsis sites invariably *are* phases (cf. Holmberg 1999). This approach would entail claiming that the complement of D in (4a), for example, is a phase, while the one in (4b) is not. The resulting approach to phases is reminiscent of Chomsky’s (2001) claim that only transitive vPs are phases; having a specifier would have to have consequences for phasehood. I will not develop this alternative further here.
Q bought they house  
‘Did they buy a house?’  
b. Creidim gur cheannaigh [siad-teach]  
I-believe that bought [they-house]  
‘I believe they did’

(9) a. Nasaan si Juan? [Tagalog]  
where T Juan  
‘Where’s Juan?’  
b. Umalis na [si-Juan]  
left now [T-Juan]  
‘He’s left’

McCloskey (1991) argues that (8b) involves vP-ellipsis, with the verb having raised out of the elided category. In what follows I will try to argue that this is the correct analysis for the Tagalog example in (9b) as well. I will be trying to show, in other words, that in (9b) a verbal projection containing the DP si Juan has been elided, after the verb has moved out. An alternative approach, of course, would claim that (9b) is an instance of DP-ellipsis, or pro-drop; I will try to show that this is not the correct analysis for Tagalog.

If Tagalog and Irish do have vP-ellipsis, they present apparent problems for the generalization in (7), since the heads preceding the ellipsis sites in the (b) examples above appear to lack specifiers. In the following sections I will present evidence that the problem is real, after which I will attempt to solve it.

4) A note is in order about the Tagalog abbreviations used in the glosses. Most Tagalog clauses contain one nominal which has had its case morphology replaced with a special morpheme; this nominal has been referred to in the Austronesian literature as the “topic”, the “subject”, or the “trigger”, among other names. I will refer to it here as the “topic”, and gloss the topic-markers ang and si as T. For discussion of the syntactic properties of the topic, see Schachter 1976, 1996, Kroeger 1993, Richards 1999, Rackowski 2002, and
1.1 vP-ellipsis in a VSO language, part 1: strict and sloppy identity

Like Tagalog, Chinese and Japanese exhibit a type of ellipsis which might be analyzed as either DP- or vP-ellipsis:

(10) a. John-wa tegami-o suteta
   John TOP letter ACC discarded
   ‘John threw away a letter’

b. Mary-mo [tegami-o] suteta
   Mary also letter ACC discarded
   ‘Mary did too’

Ellipsis in (10b) could in principle either be DP-ellipsis of the DP tegami-o ‘letter-ACC’, or of some larger verbal projection which the verb has moved out of. Huang (1987) and Otani and Whitman (1991) offer arguments for a vP-ellipsis analysis, on the basis of the behavior of strict and sloppy readings. They take as their starting point a theory of strict and sloppy readings developed by Sag (1976) and Williams (1977), which is meant to account, among other things, for the fact that examples like (11) have both strict and sloppy readings while examples like (12) have only the strict reading:

(11) Juan gave flowers to his wife, and Bill did too
   a. Strict reading: Bill gave flowers to Juan’s wife
   b. Sloppy reading: Bill gave flowers to Bill’s wife

(12) Juan heard that I gave flowers to his wife, and Bill heard that

---

references cited there. The dative markers sa and kay will be glossed DAT. Nominals which are neither topic nor dative are marked with ng or ni, which I will gloss as Unm (for Unmarked). Verbal morphology indicating, roughly, the thematic role of the topic will be glossed AV (“agent-voice”), LV (“locative-voice”), or TV (“theme-voice”). Finally, the Tagalog “linker”, a morpheme which often serves as a declarative complementizer, will be glossed LI.
you did

a. **Strict** reading: ... that you gave flowers to Juan’s wife
b. *Sloppy* reading: ... that you gave flowers to Bill’s wife

Sag and Williams develop theories in which the distribution of strict and sloppy readings follows from the nature of the predicates which can be taken to have been elided. In (11), the elided predicates could be either of those in (13):

(13) a. \([\forall x \ (x \text{ gave flowers to Juan’s wife})]\) (Strict)
b. \([\forall x \ (x \text{ gave flowers to } x’s \text{ wife})]\) (Sloppy)

Applied to *Bill*, the predicate in (13a) (the property of having given flowers to Juan’s wife) yields the strict reading; Bill has given flowers to Juan’s wife. The predicate in (13b) yields the sloppy reading; Bill shares with Juan the property of having given flowers to one’s own wife.

Crucially, in (12), the elided predicate cannot be (13b); this is not the predicate that has been applied to *I* in the first conjunct. The elided predicate can therefore only be (13a), and only the strict reading is available. More generally, Sag’s and Williams’ theories lead us to expect that sloppy readings will only be available when the binder for the variable is the subject of the elided predicate, and this seems to fit the facts in (11-12).

Otani and Whitman (1991) note that the facts are similar in Japanese. Strict and sloppy readings are both available in examples like (14):

    John-TOP self-GEN letter-ACC discarded
    ‘John threw away self’s letters’
b. Mary-mo [____ ] suteta
    Mary-also discarded
On the other hand, in examples like (15), where John is not the subject of the clause containing zibun, only the strict reading is available:

(15) a. John-wa [NY Times-ga zibun-no kizi-o
   John-TOP NY Times-NOM self-GEN article-ACC
   inyoositeiru to] kiita
   is-quoting that heard
   ‘John heard that the NY Times is quoting self’s article’

b. Bill-mo [NY Times-ga [ ____ ] inyoositeiru to] kiita
   Bill-also NY Times-NOM is-quoting that heard
   ‘Bill also heard [that the NY Times is quoting ____ ]’

If we assume that the relevant ellipsis in Japanese is vP-ellipsis, and assume Sag’s and Williams’ approaches to the semantics of elided vPs, the facts follow straightforwardly.

The facts discussed by Otani and Whitman for Japanese hold for Tagalog as well. Examples like (16) have both a strict and a sloppy reading, as expected:

(16) Nagbigay si Juan ng bulaklak sa kanyang asawa,
   AV-gave T Juan Unm flower DAT his spouse
   ‘Juan gave flowers to his wife ...’

a. ... at nagbigay din si Bill [ ____ ]
   and gave also T Bill
   ‘... and Bill did too’

b. ... at nagbigay naman si Bill ng tsokolate [ ____ ]
   and AV-gave NAMAN T Bill Unm chocolate
   ‘... and Bill, on the other hand, gave (her) chocolate’

Examples like (17), on the other hand, where a clause boundary
separates *Juan* and the ellipsis site, have only a strict reading\(^5\): 

(17) Narinig ni Juan na nagbigay ako ng bulaklak sa TV-heard Unm Juan LI AV-gave T-I Unm flower DAT ‘Juan heard that I gave flowers to his wife and Bill, ...’ kanyang asawa, ... at narinig naman ni Bill na his spouse and TV-heard NAMAN Unm Bill LI nagbigay ka [____]. 
AV-gave T-you ‘... on the other hand, heard that you did’ 

In this regard, then, Tagalog ellipsis behaves like Japanese and English ellipsis, and to the extent that Otani and Whitman’s arguments succeed in establishing that Japanese ellipsis is vP-ellipsis, they carry over to Tagalog as well.

On the other hand, a number of arguments have been offered against Otani and Whitman’s approach in the literature, notably by Hoji (1998) and Kim (1999). I will discuss one of these arguments in the following section.

### 1.2 vP-ellipsis in a VSO language, part 2: against DP-ellipsis for Tagalog

A number of people have pointed out counterexamples to the theories of ellipsis developed by Sag and Williams, on which Otani and Whitman’s arguments are based. Fiengo and May (1994), for instance, 

\(^5\) This is not because Tagalog disallows vehicle change. Examples like (i) are acceptable with a sloppy reading (or with a strict one):

(i) Nagbigay si Juan ng bulaklak sa kanyang asawa, AV-gave T Juan Unm flower DAT his spouse at nagbigay ka rin pala [____]
and AV-gave T-you also I-discover
‘Juan gave flowers to his wife, and (I see that) you did too’
discuss examples like the following:

(18) I didn’t know Bill was a bigamist. Mary just said he’s married to her, and Sue said he is [___], too.

Surprisingly for Sag and Williams, (18) seems to have both a strict and a sloppy reading, unlike the structurally similar examples which we were considering earlier; it is possible for Sue to be claiming that Bill is married to Sue. Fiengo and May take this as evidence that we need a theory that does not make the predictions discussed in the previous section, and construct one, in which the distinction between strict and sloppy readings does not rest in properties of the elided predicate as a whole, but rather in properties of the relation between the variable and its binder. If their approach is right, none of the data in the previous section can be used to distinguish between DP-ellipsis and vP-ellipsis.\(^6\)

Kim (1999), working within Fiengo and May’s (1994) approach to ellipsis, notes that Korean ellipsis can have very different properties from English VP-ellipsis. He notes, for example, the availability of at least three readings for examples like (19).\(^7\):

(19) a. Mike-ka [caki-uy ai]-lul ttayli-ess-ta  
   Mike NOM self GEN child ACC hit PAST IND  
   ‘Mike hit his/her child’

b. Kuleca Jeanne-to ttohan [___] ttayli-ess-ta  
   then Jeanne-also too hit PAST IND  
   ‘And then, Jeanne hit...’
   (i) ‘... her (Jeanne’s) child, too’
   (ii) ‘... his (Mike’s) child, too’

---

6) In fact, if their approach is right, then the reported contrast in (11-12) is surprising. They claim that such apparent contrasts are illusory.
7) Ken Hiraiwa, Shigeru Miyagawa, and Shoichi Takahashi (p.c.) tell me that Japanese ellipsis has similar properties, as long as the context is right.
The readings in (i) and (ii) are no surprise; these are the sloppy and strict readings, respectively. The surprise is the reading in (iii), which is definitely unavailable for English vP-ellipsis:

(20) Mike hit his child, and then Jeanne did [ ___ ].
   (i) [hit her (Jeanne’s) child]
   (ii) [hit his (Mike’s) child]
   (iii) *[hit Mike]

As an extreme example of the reading in (iii) above, Kim offers the example in (21):

    Mike NOM hit when John also hit PAST IND
    ‘When Mike hit (John), John also hit (Mike)’

These facts are surprising if we think that the only parse of an example like (19b) involves vP-ellipsis; Kim argues that we must allow DP-ellipsis as an option for Korean (though this argument does not rule out the possibility that vP-ellipsis also exists).

Interestingly, the readings that Kim uses to motivate a DP-ellipsis account of the Korean facts are absent in Tagalog:

(22) a. Sinuntok ni Mike ang anak niya
    TV-hit Unm Mike T child his
    ‘Mike hit his child’

b. At sinuntok din ni Jeanne [ ___ ]
   and TV-hit also Unm Jeanne
   ‘And then Jeanne also hit ...’
   (i) ‘... her (Jeanne’s) child’
(ii) ‘... his (Mike’s) child’
(iii) *‘... Mike’

(23) *Noong sinuntok ni Mike [___], sinuntok din
    when-PAST TV-hit Unm Mike TV-hit also
    ni John [___]
    Unm John
    ‘When Mike hit (John), John also hit (Mike)’

Like English (20) and unlike Korean (19), the Tagalog example in (22) is constrained to the strict and sloppy reading possibilities. The Korean example in (21), translated literally into Tagalog as (23), makes no sense. Whether or not Kim is right about Korean, then, it seems clear that Tagalog is behaving differently. Ellipsis in Tagalog is more restricted in its possible readings than Korean (and Japanese) ellipsis seems to be; in particular, Tagalog ellipsis lacks the readings that motivate Kim’s DP-ellipsis analysis of Korean, suggesting that Tagalog ellipsis may indeed be vP-ellipsis.

1.3 Ellipsis and clitics

Another argument for regarding Tagalog ellipsis as ellipsis of some verbal category, rather than DP-ellipsis, has to do with the interaction of ellipsis with Tagalog clitics. Tagalog has a number of second-position clitics. Some of these are pronouns, like the ones in (24):

(24) a. Masaya siya ngayon
    happy T-he today
    ‘He is happy today’

b. Hindi siya masaya ngayon
    not T-he happy today
    ‘He is not happy today’
c. Bakit siya hindi masaya ngayon?
   why T-he not happy today
   ‘Why isn’t he happy today?’

The pronominal clitic siya “he/she (Topic)” is placed in second position in all of the above examples. This is not ordinary behavior for nominal arguments; non-clitic arguments generally follow the predicate head:

(25) a. Masaya si Juan ngayon
   happy T Juan today
   ‘Juan is happy today’

b. Hindi masaya si Juan ngayon
   not happy T Juan today
   ‘Juan is not happy today’

c. Bakit hindi masaya si Juan ngayon?
   why not happy T Juan today
   ‘Why isn’t Juan happy today?’

In addition to these pronominal clitics, Tagalog has another class of clitics which I will refer to here as adverbial clitics; they have meanings having to do with things like aspect, veridicality, and addressee honorification:

(26) a. Masaya ho raw yata si Pete ngayon
   happy HON they-say apparently T Pete today
   ‘They say that Pete’s happy today, apparently’

b. Hindi ho raw yata masaya si Pete ngayon
   not HON they-say apparently happy T Pete too
   ‘They say that Pete’s not happy today, apparently’

Pronominal clitics and adverbial clitics behave differently under different types of ellipsis, as we will see.
In what I have been calling vP-ellipsis, a pronominal clitic in the model must correspond to a pronominal clitic in the elided version:

(27) a. Sinabi kong magbibigay ako ng pera sa
   TV-said Unm-I-LI AV-will-give T-I Unm money DAT
   simbahan, at nagbigay nga *(ako) [ ]
   church and AV-gave indeed T-I
   ‘I said I would give money to the church, and I did’

b. Sinabi ni Juan na magbibigay siya ng pera
   TV-said Unm Juan that AV-will-give T-he Unm money
   sa simbahan, at nagbigay nga *(siya) [ ]
   DAT church and AV-gave indeed T-he
   ‘Juan said he would give money to the church, and he did’

c. Sinabi kong magbibigay si Juan ng pera sa
   TV-said Unm-I AV-will-give T Juan Unm money DAT
   simbahan, at nagbigay nga (siya) [ ]
   church and AV-gave indeed T-he
   ‘I said Juan would give money to the church, and he did’

In all of the examples in (27), just when the subject of the first conjunct is a clitic pronoun, the subject of the second conjunct cannot be elided by vP-ellipsis. The generalization seems to be that vP-ellipsis cannot elide pronominal clitics (and perhaps cannot elide clitics at all; it is difficult to find a context in which an adverbial clitic would be expected to be obligatory, which makes it hard to test whether adverbial clitics can be affected by vP-ellipsis).

8) For this to be the correct generalization, the version of (27c) with an elided subject must involve a non-pronominal subject, perhaps si Juan ‘Juan’. If the name is actually repeated in the second conjunct, the result is just as awkward-sounding as it would be in English. I have to assume that this awkwardness has to do with actually repeating the name twice in the PF representation, not with simply having it syntactically present twice.
We might be tempted to attribute this difficulty with eliding clitics to the ellipsis process itself. Because clitics are attaching to a head which is not itself being elided, the reasoning might go, they obligatorily escape the ellipsis process. While this approach is perfectly reasonable, further investigation seems to reveal that it is on the wrong track; clitics behave differently under other kinds of ellipsis in Tagalog, even when the heads to which they would cliticize are not elided.

Tagalog negation is preverbal, and ellipsis of the complement of negation is possible:

9) There is an exception to this generalization. Pronominal clitics may be elided by vP-ellipsis when the ellipsis site is in the main clause, and there are no adverbs or negation in the sentence. Thus, successful ellipsis of the pronominal clitic in (ic) contrasts with failure of ellipsis in all the other examples in (i):

(i) a. Sinabi kong magbibigay ako ng pera sa
   TV-said Unm-I-LI AV-will-give T-I Unm money DAT
   simbahan, at nagbigay nga *(ako)
   church and AV-gave indeed T-I
   ‘I said I would give money to the church, and I did’
b. Nagbigay ka ba ng pera sa simbahan?
   AT-gave T-you Q Unm money DAT church
   ‘Did you give money to the church?’
c. Oo, nagbigay *(ako)
   yes AV-gave T-I
   ‘Yes, I did’
d. Oo, nagbigay *(ako) kahapon
   yes AT-gave T-I yesterday
   ‘Yes, I did yesterday’
e. Hindi *(ako) nagbigay
   not T-I AV-gave
   ‘No, I didn’t’

Holmberg (1999) discusses a condition on Finnish subject-ellipsis which is remarkably similar, and it is possible that his account will generalize to the Tagalog facts (he claims that the relevant kind of ellipsis in Finnish involves ellipsis of a larger structure than vP).
(28) Hindi ko alam kung nagbigay si Juan ng pera not Unm-I know if AV-gave T Juan Unm money sa simbahan, pero sinabi ni Maria na hindi [ ____ ] DAT church but TV-said Unm Maria LI not ‘ I don’t know if Juan gave money to the church, but Maria said he didn’t ’

In this kind of ellipsis, pronominal clitics are required to be absent, even though the head to which they would cliticize, hindi ‘not’, is not elided. Adverbial clitics, by contrast, are allowed to appear:

(29) Hindi ko alam kung nagbigay ako ng pera sa not Unm-I know if AV-gave T-I Unm money DAT simbahan ... church ‘ I don’t know if I gave money to the church ... ’

a. ... pero sinabi ni Maria na hindi (*ako) but TV-said Unm Maria LI not T-I ‘ ... but Maria said I didn’t ’

b. ... pero sinabi ni Maria na hindi pa / raw / yata but TV-said Unm Maria LI not yet they-say apparently ‘ ... but Maria said I didn’t yet/reportedly/apparently ’

Tagalog also exhibits Sluicing (and Sprouting):

(30) a. May dumating, pero hindi ko alam kung sino EXIST AV-came but not Unm-I know +WH who ‘ Somebody came, but I don’t know who ’

b. Darating si Juan, pero hindi ko alam kung AV-will-come T Juan but not Unm-I know +WH kailan when
‘Juan will come, but I don’t know when’

In Sluicing examples, no clitics at all may appear\(^\text{10)}\):

(31) Gusto kong bumalik sa Pilipinas,
want Unm-I-LI AV-return DAT Philippines
‘I want to go back to the Philippines ...’

\begin{itemize}
  \item a. ... pero hindi ko alam kung kailan (*ako)
      but not Unm-I know +WH when T-I
      ‘... but I don’t know when’
  \item b. *... pero hindi ko alam kung kailan
      but not Unm-I know +WH when nga / kaya / naman
      indeed I-wonder NAMAN
      ‘... but I don’t know when indeed/I-wonder/by-contrast’
  \item c. ... pero hindi ko alam kung kailan
      but not Unm-I know +WH when nga / kaya / naman *(ako) babalik
      indeed I-wonder NAMAN T-I AV-will-return
\end{itemize}

As the contrast between (31b) and (31c) shows, these particular adverbial clitics are semantically felicitous with this kind of question; they simply cannot appear in Sluicing contexts.

It is worth noting, first of all, that these facts are problematic for an

\(^\text{10)}\) This is not quite true. The adverbial clitic \textit{pa}, in addition to its aspectual meaning (‘still’) has a related meaning something like English adnominal \textit{else}:

\begin{itemize}
  \item (i) Alam kong darating si Juan, pero hindi ko alam
      know Unm-I-LI AV-will-come T Juan but not A-I know
      kung sino \textit{pa}
      +WH who else
      ‘I know Juan is coming, but I don’t know who else’
\end{itemize}

I assume that the clitic \textit{pa} here is modifying the DP \textit{sino} ‘who’, rather than the clause containing it.
account that attributes to Tagalog the possibility of eliding a DP without eliding any other material (DP-ellipsis, or pro-drop, or any equivalent mechanism). It is unclear why, if Tagalog had DP-ellipsis, DP-ellipsis just of pronouns should be impossible, or why it should become possible again, and in fact obligatory, when the verb is elided.

The facts discussed above are summarized in the tree in (32):

The concentric ovals represent the different types of ellipsis we have just discussed. As we saw above, vP-ellipsis (represented by the smallest oval) is unable to affect clitics at all, while ellipsis of the complement of negation (the next larger oval) obligatorily eliminates pronominal clitics, and Sluicing (the largest oval) is incompatible with all clitics.

This subset relation between the different clitics could be captured straightforwardly if ellipsis operated on a structure in which adverbal clitics were placed between C and Neg, and pronominal clitics between
Neg and the raised position of the verb:

(33)

The facts about the interactions between clitics and ellipsis follow straightforwardly from a tree like the one in (33). vP-ellipsis, according to this tree, cannot affect clitics because they are not inside the constituent being elided. Ellipsis of the complement of negation obligatorily elides pronominal clitics, which are inside the elided constituent, but not adverbial clitics, which are outside it. Finally, sluicing elides all clitics, along with everything else inside the complement of C.

Of course, the tree in (33) does not accurately predict the word order of the clitics. All other things being equal, this tree leads us to expect that a sentence containing an adverbial clitic, a pronominal clitic, and a verb should put them in the order in (34a); the correct order is in fact (34b):

(34) a. *Ba siya nagngangawa?*

    Q  T-he AV-is-babbling
I am forced to conclude, then, that these second-position clitics are placed in second position by a process which follows ellipsis in the derivation, perhaps a post-syntactic process. I will give two arguments for such a conclusion in the following section, trying to show that clitics are indeed ordered by a PF operation.

### 1.3.1 Post-syntactic clitic ordering, part 1: sorting by syllables

Tagalog clitics appear in an order which is more or less fixed, as shown in (35) (in the list in (35), the highest clitics are the once which come first in a sequence):

(35)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>ko ' I (Unm)'</td>
</tr>
<tr>
<td>b</td>
<td>na ' as of now'</td>
</tr>
<tr>
<td>c</td>
<td>man ' even'</td>
</tr>
<tr>
<td>d</td>
<td>nga ' indeed'</td>
</tr>
<tr>
<td>e</td>
<td>din ' also'</td>
</tr>
<tr>
<td>f</td>
<td>lang ' just, merely'</td>
</tr>
<tr>
<td>g</td>
<td>daw ' they say'</td>
</tr>
<tr>
<td>h</td>
<td>po, ho ' honorific'</td>
</tr>
<tr>
<td>i</td>
<td>ba ' question'</td>
</tr>
<tr>
<td>j</td>
<td>muna ' first, for a while'</td>
</tr>
<tr>
<td>k</td>
<td>naman ' on the other hand'</td>
</tr>
<tr>
<td>l</td>
<td>kasi ' because'</td>
</tr>
<tr>
<td>m</td>
<td>kaya ' maybe'</td>
</tr>
<tr>
<td>n</td>
<td>sana ' counterfactual'</td>
</tr>
</tbody>
</table>

[galá ' I have just discovered', yata ' apparently']
As the list in (35) shows, monosyllabic clitics generally precede disyllabic clitics in Tagalog; (a-i) are monosyllabic, and (j-p) are disyllabic. This by itself suggests that there are non-syntactic principles at work in the ordering of the clitics; the syntax should not be responsible for sorting the clitics by number of syllables.

Within the disyllabic clitics, the ordering is what we would expect given the tree in (33); adverbial clitics precede pronominal clitics. Within the monosyllabic clitics, the ordering is the reverse of what we expect. The result is that the pronouns are on the periphery of the clitic cluster, with the monosyllabic pronouns coming first and the disyllabic pronouns coming last. If the claim developed here about the syntactic positioning of clitics is on the right track, then the post-syntactic process responsible for sorting the clitics by number of syllables apparently reverses the order given by the syntax to the monosyllabic clitics, while leaving the disyllabic clitics as they are.

In a perfect world, the easiest way to argue for this conclusion would be to look at the ordering of the pronominal clitics. For disyllabic clitics the principle is straightforward: clitics in the "unmarked" case precede Topic clitics\(^1\). The theory then predicts that for monosyllabic clitics the order should be reverse. Unfortunately, monosyllabic clitics never cooccur, making the prediction untestable\(^2\).

The ordering of monosyllabic adverbial clitics, however, seems to

\(^1\) For reasons having to do with the properties of Tagalog Topics (see the literature cited in footnote 4 for details), this means that among the disyllabic pronominal clitics, thematic subjects precede thematic objects.

\(^2\) The one case where one might expect monosyllabic pronominal clitics to cooccur is that of an unmarked 1st person singular clitic *ko* and a Topic 2nd person singular clitic *ka*. As it happens, however, this sequence is always replaced with a portmanteau clitic *kita* (historically a 1st person dual inclusive pronoun).
be consistent with the idea that structurally higher clitics are ordered later within this field. For example, the adverbial clitics *na* "as of now" and *pa* "still", which modify the aspect of the clause, are the earliest monosyllabic adverbial clitics, while the latest is *ba*, which marks the sentence as a question. On the common assumption that the syntactic representation of aspect is structurally lower than the part of the structure responsible for making the clause a statement or a question, this ordering suggests that hierarchically higher clitics are following hierarchically lower clitics, within the monosyllabic clitic domain.

In this section I have tried to show that clitic placement is at least partly achieved by post-syntactic operations which make reference to factors like number of syllables. I have also suggested that the facts of clitic ordering offer indirect support for the conclusion reached in the previous section on the basis of the ellipsis facts, that adverbial clitics generally are structurally above pronominal clitics.

### 1.3.2 Post-syntactic clitic ordering, part 2: order and scope

Facts about the scope interactions of clitics with non-clitics also suggest that clitics sometimes undergo post-syntactic lowering into second position. Tagalog has two aspectual clitics, *na* and *pa*, exemplified below:

\[(36) \begin{align*}
(a) & \text{ Matanda } na & \text{ siya} \\
& \text{ old as-of-now T-he/she} \\
& \text{ 'He/she is old now (and wasn't before)'} \\
(b) & \text{ Bata } pa & \text{ siya} \\
& \text{ young still T-he/she} \\
& \text{ 'He/she is still young'}
\end{align*} \]

When negation is added to these sentences, the clitics obligatorily take scope over negation:
(37) a. Hindi *na* siya bata
   not as-of-now T-he/she young
   'He's not young any more'
   as-of-now > NEG

b. Hindi *pa* siya matanda
   not still T-he/she old
   'He's not old yet'
   still > NEG

This is what the tree in (33) would lead us to expect. The adverbial clitics *na* and *pa* begin the derivation above negation and are interpreted there; they only follow negation in (37) because they are post-syntactically lowered into second position.

1.3.3 Ellipsis and clitics: conclusions

We have seen some evidence in this section that Tagalog sentences like (38) should be analyzed as involving vP-ellipsis, rather than DP-ellipsis or pro-drop:

(38) a. Nasaan si Juan?
   where T Juan
   'Where's Juan?'

b. Umalis na [si-Juan]
   left now [T-Juan]
   'He's left'

The vP-ellipsis account, unlike the DP-ellipsis account, has the potential to explain why ellipsis of postverbal material is unable to delete pronominal clitics. The account developed here of that fact claimed that pronominal clitics are in fact above the verb in the syntax, and are lowered into second position post-syntactically, after ellipsis has applied. I offered some independent evidence that Tagalog clitics are moved post-syntactically, including evidence that they are postsyntactically sorted so that monosyllabic clitics precede disyllabic clitics, and
evidence from scope interactions that they are sometimes pronounced in a lower position than they are interpreted\(^\text{13}\).

2. EPP, ellipsis, and linearization

At the outset of this paper, building on work by Saito and Murasugi (1990) and Lobeck (1995), I proposed a generalization which required heads with phases or ellipsis sites as their complements to have specifiers:

\[(39)\] A phrase with no internal syntactic structure must be the complement of a head with a specifier.

I then noted an apparent problem for (39); there are several VSO languages, including Irish and Tagalog, in which ellipsis of post-verbal material is possible, even though there is no apparent specifier for the projection hosting the verb:

   ‘Did they buy a house?’

\(^{13}\) This account has a peculiar consequence for the nature of pronominal clitics in Tagalog. The facts about the interaction of ellipsis with clitics led us to the proposal that, as far as the syntactic representation operated on by ellipsis is concerned, pronominal clitics are in a preverbal position, and are lowered post-syntactically to a post-verbal position. Non-pronominal clitics, on the other hand, are apparently always in a post-verbal position. We therefore arrive at the conclusion that pronouns in Tagalog undergo a kind of shift to a pre-verbal position in the syntax which is not undergone by non-pronouns, after which they post-syntactically retreat back to a post-verbal position. There are, of course, other languages in which pronouns undergo an obligatory shift to a high position which cannot be occupied by non-pronouns (this is the behavior, for instance, of object pronouns in Swedish; cf. Bobaljik (1995) and much other work). What makes Tagalog special, on this account, is the subsequent post-syntactic lowering of the shifted pronoun, which conceals the fact that shift has taken place at all.
b. Creidim gur cheannaigh [siad-teach]
   I-believe that bought [they-house]
   ‘I believe they did’

(41) a. Nasaan si Juan?
   where T Juan
   ‘where’s Juan?’

b. Umalis na [si-Juan]
   left now [T-Juan]
   ‘He’s left’

McCloskey (1991) argued for a vP-ellipsis analysis of examples like (40b) in Irish, and I have argued above that this is the correct analysis for Tagalog examples like (41b) as well. The problem would therefore seem to be a real one; Tagalog and Irish are counterexamples to (39).

Tagalog and Irish have some other properties in common as well. McCloskey (1996a) argues that the Irish complementizer lowers to the verb. He notes, for example, that clause-initial adverbs in Irish can precede the complementizer, a fact he attributes to the lowering of the complementizer past the adverb and onto the verb:

(42) Is doiche [faoi cheann cúpla lá go bhfáidfaí imeacht]
   is probable at-the-end-of couple days that could leave
   ‘It is probable that at the end of a couple of days they could leave’

I argued in the previous section that Tagalog clitics also lower to the verb. We might try to amend the generalization in (39) to account for the Irish and Tagalog exceptions, then, as follows:

(43) A phrase which becomes invisible to the computational system must be the complement of a head which either:
   (a) has a specifier, or
Why there is an EPP

(b) is lowered to by higher material.

Since (39) was meant to capture both the conditions on ellipsis and the conditions under which EPP effects are found, the revised version in (43) predicts that Tagalog and Irish will lack EPP effects; since the verb is lowered to by higher material, it will not need to have a specifier even though its complement (by hypothesis) is a phase. McCloskey (1996b) argues that Irish does indeed lack EPP effects, noting the existence of examples like (44) where no DP arguments are present (and arguing against the existence of a null expletive satisfying the EPP):

(44) Chuaigh [de mo neart]
    went of my strength
    'My strength waned'

Tagalog seems to lack examples like (44), as far as I can tell, but there are certainly no examples that would compel one to posit an EPP for Tagalog; examples which would involve expletives in languages like English are apparently subjectless:

(45) a. Umulan
    AT-rained
    'It rained'

b. Mukha-ng masaya si Juan
    seem LI happy T Juan
    'It seems that Juan is happy'

The condition in (43), then, seems to account for the distribution of possible ellipsis sites, as well as the distribution of EPP positions. On the other hand, (43) is rather descriptive. Is there any way of deducing the effects of (43) from general principles?

The account I will develop of the generalization in (43) will be
based on Kayne's (1994) Antisymmetry; I will assume, following Kayne, that there is a mapping between the asymmetric c-command relations in the tree and precedence relations\(^\text{14}\). Consider, for instance, the linearization of a tree like (46):

\[\text{(46)}\]

```
CP
  \|-- C
    \|-- TP
       \|-- DP
           \|-- T
               \|-- John
                   \|-- will
                       \|-- v-V
                           \|-- dance
                       \|-- vP
                   \|-- T
               \|-- VP
       \|-- \n```

Linearizing the tree in (46) is a matter of determining the pairs of (XP or X\(^0\)) nodes \(<\alpha, \beta>\) such that \(\alpha\) asymmetrically c-commands \(\beta\); such pairs are interpreted by the PF interface as instructions that \(\alpha\) should precede \(\beta\). If all the words in the tree can be linearized by this method, then the tree is well-formed. In (46), for instance, the DP dominating \(\text{John}\) asymmetrically c-commands the T dominating \(\text{will}\), which means that \(\text{John}\) precedes \(\text{will}\).

I will assume that the derivation proceeds by phases, as in Chomsky (2000, 2001); in the tree in (46), for example, the vP is a phase which is sent to PF once it is completed, rendering its internal structure invisible to the syntax and effectively making vP a head as far as the syntax is concerned\(^\text{15}\). In order to make this consistent with Kayne

---

14) Kayne (1994), along with much subsequent work, assumes one maximally simple version of such a mapping, claiming that if \(\alpha\) asymmetrically c-commands \(\beta\), \(\alpha\) precedes \(\beta\). While this is conceptually attractive, any kind of mapping between c-command and precedence is in principle consistent with Kayne's (1994) basic insight (e.g., "if \(\alpha\) asymmetrically c-commands \(\beta\), then \(\alpha\) precedes \(\beta\) unless \(\alpha\) is a head, in which case \(\alpha\) follows \(\beta\)", which would be a mapping that would yield a head-final language). Thanks to Danny Fox for raising this point.
(1994), I will assume that Spell-Out of the vP phase occurs after T has already been linearized with the interior of vP. Thus, PF already has instructions for T to precede v, before vP is made into a syntactic atom. This avoids a problem with linearization that might otherwise arise; if vP is effectively a head, then T and the vP head are in a mutual c-command relation, and ordering should be impossible. I will also assume, as before, that ellipsis is like Spell-out in reducing the elided XP to a syntactic atom.

I will also follow Kayne in assuming that the specifier of XP is not dominated by XP. Kayne collapses the standard distinction between specifiers and adjuncts, claiming that like adjuncts, specifiers are only contained in the maximal projection of which they are a specifier, not dominated by it. Thus, the DP John in (46) is not dominated by TP; the lowest node dominating it is CP, and it is therefore in a mutual c-command relation with C. Linearization is still possible in this case, since C can c-command the internal structure of this DP (for instance, the N John).

Finally, I will assume with Chomsky (1995, 2000) that the specifier of XP is just the YP daughter of XP, and that the complement of X is just the YP sister of the head X. These definitions are not exclusive in a phrase structure like the one posited by Chomsky (1995); in the tree in (46), for example, VP is both the complement of v and the specifier of vP. By the assumption in the last paragraph, then, VP, as the specifier of vP, is in a mutual c-command relation with T.

Armed with these assumptions, we can now consider a couple of relevant cases. Consider first a case in which a head W has a phase (or,  

---

15) Throughout, I will talk as though Spell-Out of a phase XP renders XP completely opaque to the computational system, and takes place after the head taking XP as a complement is Merged. Other approaches to phases have been proposed, of course (see Chomsky 2000, 2001, Nissenbaum 2000), and I am using this one here just for ease of exposition. The story developed here can be adapted to other approaches to phases.
equivalently, an ellipsis site) as its complement and has a specifier. The tree is built up to the point where \( W \) is introduced, yielding the tree in (47):

\[
\text{(47)}
\]

\[
\begin{array}{c}
\text{WP} \\
\text{W} \\
\text{XP} \\
\text{YP} \\
\text{X'} \\
\text{X} \\
\text{...}
\end{array}
\]

The phase XP undergoes Spell-Out. At this point, \( W \) is linearized with various nodes inside XP (including \( X \) and \( Y \)). After Spell-Out has taken place, XP is a syntactic atom, with no internal structure:

\[
\text{(48)}
\]

\[
\begin{array}{c}
\text{WP} \\
\text{W}
\end{array}
\]

WP's specifier is constructed and Merged:

\[
\text{(49)}
\]

\[
\begin{array}{c}
\text{WP} \\
\text{MP} \\
\text{W'} \\
\text{W}
\end{array}
\]

Finally, the head \( Q \) which takes WP as a complement is Merged:

\[
\text{(50)}
\]

\[
\begin{array}{c}
\text{QP} \\
\text{Q} \\
\text{WP} \\
\text{MP} \\
\text{W'} \\
\text{W}
\end{array}
\]
The subtree in (50) will be linearizable when it undergoes Spell-Out. Q asymmetrically c-commands W, XP, and various parts of the interior of MP\textsuperscript{16}. MP asymmetrically c-commands W and XP, and W has previously been ordered with material inside the XP, in the phase that was spelled out previously. A total ordering can therefore be constructed, with Q preceding the material inside MP, which precedes W, which precedes XP.

Suppose we now consider an alternative derivation without a specifier for WP. The derivation proceeds as before up to the step in the tree in (48), at which point the head Q is Merged:

\[
\begin{array}{c}
\text{(51)} \\
\text{QP} \\
\text{Q} \\
\text{WP} \\
\text{W} \\
\end{array}
\]

This subtree is not linearizable. W is linearized with the interior of XP prior to Spell-Out, as before. And Q asymmetrically c-commands W, as before. But XP, as the daughter of the maximal projection WP, is the specifier of WP (as well as the complement of W). As a result, Q and XP are in a mutual c-command relation. Because XP has no internal structure at this point in the derivation, linearization of Q with XP is impossible; they are simply two heads in a mutual c-command relation, as far as the syntax is concerned. Linearization will therefore fail.

There are several imaginable ways of avoiding this outcome, however, even if WP lacks a specifier. Since the relation between Q and XP is the problematic one, one way to avoid the problem would be to post-syntactically lower Q onto W. This makes the ordering between

\[\text{16) The account given here assumes that MP always has complex internal structure; this would have to be the case even when MP is a pronoun or an expletive, for example. Alternatively, pronouns and expletives may make use of one of the other ways of licensing these structures (postsyntactic lowering, for example). I will have to leave this issue open for now.}\]
Q and W the problem of the morphology; Q is now in the same position as W, as far as the syntax is concerned, and W’s linearization with the interior of XP has been established on a previous phase\textsuperscript{17).

3. Conclusion

We have now arrived at the result we wanted. In a tree like (51), if W’s complement XP is a phase or an ellipsis site, some way must be found of linearizing XP with the head Q taking WP as a complement. If WP has a specifier, then the relation between Q and XP is one of asymmetric c-command, and linearization succeeds. If Q post-syntactically lowers onto W, then Q and W are in the same position as far as the syntax is concerned, and since W has already been linearized with the interior of XP on a previous phase, linearization succeeds. The descriptive generalization in (43), repeated here as (52), is thus made to follow from general principles:

\begin{center}
\textbf{(52)} A phrase which becomes invisible to the computational system must be the complement of a head which either:
\end{center}

\textsuperscript{17) An anonymous reviewer raises another possible way of avoiding linearization difficulties in trees like (51): what happens if the higher head Q is not Merged at all? The reviewer suggests that we might expect, on the approach developed here, to find cases in which the highest projection of the clause must have its specifier filled in embedded clauses, but need not in matrix clauses. The status of the prediction depends partly on where we think phase boundaries are; if phases are the units that are sent to PF for linearization, then the highest projection of a phase will never be forced to have a specifier by the considerations outlined in this paper, and tensed embedded clauses seem like a good candidate for phases. It will have to be the case, on this approach, that matrix clauses are never simply TP, but always contain a CP projection; otherwise, as the reviewer notes, we might expect not to find EPP effects in matrix clauses. There are a number of phenomena in matrix clauses that one could imagine accounting for in terms of some version of the reviewer’s suggestion, including French optional wh-in-situ (Chang 1997, Bošković 2000) and “diary drop” (Haegeman 1990); I will not try to develop these accounts here.
Why there is an EPP

(a) has a specifier, or
(b) is lowered to by higher material.

It is perhaps worth emphasizing that we have not drawn any conclusions about the specific mechanisms that the grammar uses to implement the EPP. It could be, for instance, that the grammar is capable of determining that linearization will fail unless a specifier is filled, and therefore performs operations that will fill it. Alternatively, we may want to deny the grammar that kind of understanding of its interface with PF, and claim that there is a feature (either a pure “EPP” feature, or some other feature that plays this role) responsible for ensuring that the EPP is satisfied. The account given here has been an explanation of why such a feature, if there is one, should be required.
References


Dordrecht: Foris.


