Bare Adjunction as “Two-Peaked” Structure

Yohei Oseki  
*New York University*  
yohei.oseki@nyu.edu

1. Introduction
1.1. Background

(1) One of the fundamental distinctions in natural language syntax:

a. *argument*: obligatory, selected, fixed, core, saturation…  
b. *adjunct*: optional, unselected, flexible, peripheral, modification…

(2) X-bar theory in GB/early MP-era (May 1985; Chomsky 1986, 1995):

a. ‘substitution’/category \((\alpha, \beta) = \{H(\beta), \{\alpha, \beta\}\}\)

\[
\begin{array}{c}
\text{XP} \\
\text{X} \\
\text{YP}
\end{array}
\]

b. ‘adjunction’/segment \((\alpha, \beta) = \{<H(\beta), H(\beta)>, \{\alpha, \beta\}\}\)

\[
\begin{array}{c}
\text{XP} \\
\text{XP} \\
\text{YP}
\end{array}
\]

+ *Late-Merge* (Lebeaux 1988; Stepanov 2001)

(3) Bare phrase structure theory in modern MP-era (Chomsky 2000, 2001, 2004 *et seq.)*:

a. *Set-Merge* \((\alpha, \beta) = \{\alpha, \beta\}\)

“For structure building, we have so far assumed only the free symmetrical operation Merge, yielding syntactic objects that are sets, all binary: call them simple. The relations that come “free” (contain, c-command, etc.) are defined on simple structures.”

(Chomsky 2004: 117)

b. *Pair-Merge* \((\alpha, \beta) = <\alpha, \beta>\)

“But it is an empirical fact that there is also an asymmetric operation of adjunction, which takes two objects \(\beta\) and \(\alpha\) and forms the ordered pair \(<\alpha, \beta>, \alpha\) adjoined to \(\beta\),”

(Chomsky 2004: 117-118)

---

* The main idea of this paper originates from the manuscript which I wrote at University of Massachusetts Amherst in 2012-2013. I am grateful to Kyle Johnson, Margaret Speas, and especially Tom Roeper for crucial input in developing this work. An earlier version of this material was presented at WCCFL 32 (University of Southern California), Yale University, and University of Connecticut, where I got insightful suggestions from the audiences. I especially thank Mark Baltin, Chris Barker, Željko Bošković, Noam Chomsky, Guglielmo Cinque, Chris Collins, Marcel den Dikken, Samuel David Epstein, Stephanie Harves, Yusuke Imanishi, Richard Kayne, Hisatsugu Kitahara, Heejeong Ko, Alec Marantz, Hiroki Narita, Liina Pylkkänen, Joachim Sabel, Daniel Seely, Anna Szabolcsi, Coppe van Urk, Patricia Schneider-Zioga, and my fellow students Itamar Kastner, Neil Myler, Vera Zu for thought-provoking comments and/or encouragement. The usual disclaimers apply.
Metaphorically speaking, Set-Merged objects are on the “primary plane” where c-command is defined, whereas Pair-Merged ones are on a “separate plane” where it is not.

“Given the basic properties of adjunction, we might intuitively think of \( \alpha \) as attached to \( \beta \) on a separate plane, with \( \beta \) retaining all its properties on the “primary plane,” the simple structure.” (Chomsky 2004: 117-118)

In the history of the argument/adjunct distinction, special construction-specific operations have been proposed for each object!

4 However, in concert with Chomsky’s (2013) recent ‘simplest Merge’ conjecture that there is Merge \( (\alpha, \beta) = \{\alpha, \beta\} \) only, it is safe to say that adjunction is still very much open to debate:

“Nobody seems to know exactly what to do with adverbs.” (Ernst 2001: 1)

“There has never, to my knowledge, been a really satisfactory theory of adjunction, and to construct one is no slight task” (Chomsky 2004: 117)

“It is fair to say that what adjuncts are and how they function grammatically is not well understood” (Hornstein & Nunes 2008: 57)

(5) **Issues to be addressed:**
   a. What kind of structure does adjunction have?
   b. What kind of operation is necessary and sufficient to explain adjunction?

1.2. **The gist of the talk**

6 Two descriptive generalizations about adjunction are to be established and explained:

   a. Adjuncts are generally *invisible* to syntax.
   b. Adjuncts entering Feature-Sharing become *visible* to syntax.

7 Late-Merge and Pair-Merge are untenable both theoretically and empirically (See Rudin (2003), Irurtzun & Gallego (2007), and Richards (2009) for elaboration of Pair-Merge):

   a. **Theoretical point:**
      Late-Merge and Pair-Merge have been introduced into the system only to capture invisibility of adjuncts.
   b. **Empirical point:**
      Late-Merge and Pair-Merge are too restrictive to capture the generalization (6b).

(8) **Answers to (5):**

   a. Adjunction has “two-peaked” structure.
   b. Simplest Merge \( (\alpha, \beta) = \{\alpha, \beta\} \) is necessary and sufficient to explain adjunction.

Invention of “two-peaked” structure is due to Epstein, Kitahara, & Seely’s (2012) analysis of “countercyclic” subject raising and deduction of subject island and Cyclic Transfer.

Proposal to be developed:
Adjunction has “two-peaked” structure.

1.3. Roadmap
- §2: Invisibility of Adjuncts in Syntax
- §3: \{XP, YP\} = “Two-Peaked” Structure
- §4: Visibility of Adjuncts in Syntax
- §5: \{XP, YP\} + Feature-Sharing = “One-Peaked” Structure
- §6: Lebeaux vs. Lasnik on Anti-Reconstruction
- §7: Conclusion

2. Invisibility of Adjuncts in Syntax
In this section, we attempt to establish the first descriptive generalization that adjuncts are generally invisible to syntax. Core data come from island (2.1.), intervention (2.2.), reconstruction (2.3.), scrambling (2.4.), and ellipsis (2.5.).

2.1. Island
Adjuncts are generally invisible to extraction.

a. *Who did Mary cry [\textit{ADJ} after John hit \textit{<who>}]?
   (Huang 1982)
b. *Who did an advocate speak to Betsy [\textit{ADJ} before a discussion of \textit{<who>}]?
   (Johnson 2003)

2.2. Intervention
Adjuncts are generally invisible to intervention.

a. √John seems [\textit{ADJ} to Mary] \textit{<John> to be smart.}
b. *They seem [\textit{ADJ} to him\textsubscript{i}] to like John\textsubscript{i}.
   (Stepanov 2001)
c. √John T\textsubscript{[present]} [\textit{ADJ quickly}] dance.  (cf. John quickly dances.)
d. √John T\textsubscript{[past]} [\textit{ADJ already}] leave.  (cf. John already left.)
   (Ochi 1999)

2.3. Reconstruction
Adjuncts are generally invisible to binding after fronting.

a. √[\textit{ADJ} In Ben\textsubscript{i}’s office], he\textsubscript{i} is an absolute director.
b. √[\textit{ADJ} With John\textsubscript{i}’s novel finished], he\textsubscript{i} began to write a book of poetry.
c. √[\textit{ADJ} To Ben\textsubscript{i}’s surprise], he\textsubscript{i} noticed that the others had left.
d. √[\textit{ADJ} For Mary\textsubscript{i}’s valor], I heared she\textsubscript{i} was given a medal.
   (Speas 1991)
The same observation can be found in extraposition (Fox & Nissenbaum 1999) and ACD (Fox 2002). See Chomsky (2004) for “after though” view on these constructions.

2.4. Scrambling

(15) Adjuncts are generally invisible to long-distance scrambling.

\[ [\text{[ADJ Riyuu-mo naku]}, \text{Mary-ga [John-ga t\_1 sono setu-o sinziteiru to] omotteiru.}] \]

reason-any without Mary-Nom John-Nom that theory-Acc believes that thinks
‘Mary thinks that John believes in that theory without any reason.’ (Saito 1985)

2.5. Ellipsis

(16) Adjuncts are generally invisible to ellipsis.

a. John visited Sally after the party. \( <\text{VP}> [\text{ADJ after the lecture}]. \) (Johnson 2004)

John-Top car-Acc carefully wash-Past
‘John washed a car carefully.’

\[ \text{Bill-wa <DP> arawa-nakat-ta.} \]
Bill-Top wash-Neg-Past
‘Bill didn’t wash a car (*carefully).’
(Oku 1998)

2.6. Summary

(17) Descriptive Generalization
Adjuncts are generally invisible to syntax.

3. \{XP, YP\} = “Two-Peaked” Structure

In this section, we explain the descriptive generalization (17) resorting to Chomsky’s (2013) Labeling Algorithm and simplest Merge. After reviewing Chomsky’s Labeling Algorithm in 3.1., the main proposal will be developed in 3.2. and explain each data of invisible adjuncts in 3.3. Binding transparency of adjuncts is also discussed briefly in 3.4.

3.1. Chomsky’s (2013) Labeling Algorithm

(18) Chomsky (2013): Labeling Algorithm = Minimal-Search/Head-Detection

<table>
<thead>
<tr>
<th>a. {X, YP}</th>
<th>b. {XP, YP}</th>
<th>c. {XP_{[+F]}, YP_{[+F]}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{XP} )</td>
<td>( \text{XP} )</td>
<td>( \text{FP} )</td>
</tr>
<tr>
<td>( \text{X} )</td>
<td>( ? )</td>
<td>( \text{XP} )</td>
</tr>
<tr>
<td>( \text{YP} )</td>
<td>( \text{YP} )</td>
<td>( \text{YP} )</td>
</tr>
<tr>
<td>( \sqrt{\text{Minimal-Search}} )</td>
<td>( *\text{Minimal-Search} )</td>
<td>( *\text{Minimal-Search} )</td>
</tr>
<tr>
<td>#Feature-Sharing</td>
<td>*Feature-Sharing</td>
<td>*Feature-Sharing</td>
</tr>
</tbody>
</table>

a. Minimal-Search can easily locate the unique head X.
b. Minimal-Search cannot find the unique head given X and Y are equidistant.
c. Minimal-Search cannot find the unique head but X and Y share [+F].
3.2. Proposal

(19) **Proposal:**
Adjunction has “two-peaked” structure.

```
  WP
 /  \               
W    ZP
 / \
Z   XP  YP (= ADJ)
```

(20) The proposal is a *theorem* naturally deduced from the axioms of phrase structure:

a. \{XP, YP\}: The defining geometry of adjunction is Merge (XP, YP).
   (May 1985; Chomsky 1986; Chometzky 1994; Kayne 1994; Narita 2011; Ott 2011)

```
XP                      \sqrt{\text{Merge} (XP, YP)}            YP (= ADJ)
```

b. **Labeling Algorithm:** Outputs of Merge are labeled independently.

   Chomsky’s (2013) Labeling Algorithm (LA):
   i. In \{X, YP\}, X is the label.
   ii. In \{XP, YP\}, there is no label.
   iii. In \{XP, YP\}, if XP and YP share [+F], then F is the label.

Notice that Chomsky’s (2013) LA is inapplicable to Merge \{XP, YP\} unless [+F] is shared.

```
XP       YP (= ADJ)          *\text{LA} (\{XP, YP\})               YP (= ADJ)
```  

c. **Label Accessibility Condition:** Unlabeled outputs are inaccessible to Merge.
   (Chomsky 2000; Collins 2002; see also Hornstein’s 2009 Concatenate/Label system)

   Hornstein’s (2009: Ch.3) Label Accessibility Condition:
   Only the label of a syntactic object is accessible to Merge.

```
Z                      *\text{Merge} (Z, \{XP, YP\})            ZP
XP                      \sqrt{\text{Merge} (Z, XP)}            YP (= ADJ)
```

Since unlabeled objects are inaccessible to Merge, one of two daughters of the unlabeled object should be the target of the next application of Merge. Importantly, Merge (Z, XP) is defined via LA.
(21) Epstein, Kitahara, & Seely (2012): Once “two-peaked” structures are generated, one of two peaks has to undergo Transfer immediately for the derivation to converge.

\[ ZP \quad \text{Transfer} \quad \{XP, YP\} \quad ZP \]

\[ Z \quad XP \quad YP (= \text{ADJ}) \quad Z \quad XP \quad YP (= \text{ADJ}) \]

3.3. \textbf{Deriving Invisibility of Adjuncts}

3.3.1. Island

(22) The fact that ‘adjuncts are generally invisible to extraction’ follows because once “two-peaked” structures are generated, adjuncts undergo Transfer immediately.

\[
\begin{array}{c}
\text{no extraction} \\
\quad CP \\
\quad C \quad TP \\
\quad T \quad vP \quad PP (= \text{ADJ}) \\
\quad \ldots Wh \ldots
\end{array}
\]

3.3.2. Intervention

(23) The fact that ‘adjuncts are generally invisible to intervention’ follows because once “two peaked” structures are generated, adjuncts undergo Transfer immediately.

\[
\begin{array}{c}
\text{no intervention} \\
\quad TP \\
\quad T \quad VP \quad PP (= \text{ADJ}) \\
\quad \ldots DP \ldots
\end{array}
\]

3.3.3. Reconstruction

(24) The fact that ‘adjuncts are generally invisible to binding after fronting’ follows because once “two peaked” structures are generated, adjuncts undergo Transfer immediately.

\[
\begin{array}{c}
\text{no binding} \\
\quad TP \\
\quad DP_{[+pro]} \quad TP \quad PP (= \text{ADJ}) \\
\quad \ldots DP_{[+R]} \ldots
\end{array}
\]
3.3.4. Scrambling
(25) The fact that ‘adjuncts are generally invisible to long-distance scrambling’ follows because once “two-peaked” structures are generated, adjuncts undergo Transfer immediately.

3.3.5. Ellipsis
(26) The fact that ‘adjuncts are generally invisible to ellipsis’ follows because once “two-peaked” structures are generated, adjuncts undergo Transfer immediately.

3.4. “SIMPL” at Transfer? - binding transparency of adjuncts
(27) Chomsky (2004) proposed the operation called “SIMPL” applied at Transfer to convert Pair-Merge/“separate plane” into Set-Merge/“primary plane”, which is necessary to capture an inherent asymmetry of adjuncts between extraction and binding.

a. Condition C:
*Hei cried [\_ADJ after Mary hit John].
b. Variable binding:
\_Every companyi was pleased [\_ADJ because John bought itsi product].
c. Wh-in-situ:
\_Who left [\_ADJ because John bought what]? d. NPI-licensing:
\_He didn’t fall asleep [\_ADJ during any of the talks].

(28) Stipulation: Following Obata (2010), chunks undergoing Transfer are “re-assembled” into one representation for the global computation like Condition C to take place at the semantic interface, where binding-type dependencies could apply.


3.5. Summary

(29) We have shown that the first descriptive generalization about invisible adjuncts can be deduced from simplest Merge and Labeling Algorithm without Late-Merge and Pair-Merge. In §4 and §5, we argue that this kind of reductionism makes correct empirical predictions about visible adjuncts.

4. Visibility of Adjuncts in Syntax

In this section, we present the second descriptive generalization that adjuncts entering Feature-Sharing become visible to syntax. Here again, core data are from island (4.1.), intervention (4.2.), reconstruction (4.3.), scrambling (4.4.), and ellipsis (4.5.).

4.1. Island

(30) Adjuncts entering Feature-Sharing are visible to extraction.


b. √What did John drive Mary crazy [ADJ trying to fix <what>?] (Trueswell 2007)

c. Hungarian:

 √Kinél szivott [ADJ nagyobbat <kinél>]?  
 who-to smoke-Pst large-Cpr-Acc  
 ‘He smoked more than who?’ (den Dikken 2012)

d. Japanese:

 √[CP Yamada-sensei-gashinsatsu-shita-yori(-mo)] Tanaka-sensei-ga kanja-o  
 Dr.Yamada-Nom examination-did-than(-also) Dr.tanaka-Nom patient-Acc  
 [ADJ <CP> oozei] shinsatsu-shita.  
 many examination-did  
 ‘Dr.Tanaka examined more patients than Dr.Yamada examined.’ (Miyamoto 2012)

‘Truswell (2011:157) accounts for these kinds of contrasts with an appeal to his Single Event Grouping Condition, which says that ‘an instance of wh-movement is legitimate only if the minimal constituent containing the head and the foot of the chain can be construed as describing a single event grouping’ (original italics). It seems to me that this ‘event grouping’ can plausibly be syntactically grouped with an appeal to an Agree relation between v and the adjunct: whenever there is ‘event grouping’, there is such an Agree relation (presumably an Agree relation for an event-structural/aspectual feature present on v), and concomitantly, the adjunct is transparent.” (den Dikken 2012: 10)

‘the accusative-marked adverbial modifier is in an Agree relation with v and that this Agree relation makes the projection of the adverbial modifier transparent” (den Dikken 2012: 11)

‘what makes some object-oriented secondary predicates transparent for extraction is the fact that object-oriented FQs can enter into an Agree relationship with Asp via the NP they modify during the course of the derivation” (Miyamoto 2012: 365)
4.2. Intervention
(31) Adjuncts entering Feature-Sharing are visible to intervention.

\[ \text{Complementizer agreement in Hellendoorn:} \]
\[ \text{dat}^{*}\text{datte} [_{\text{ADJ}} \text{op den wa`rmsten dag van `t joar}] \text{wiev} \text{tegen oonze wil ew`rkt hebt} \]
\[ \text{that/that-PL} \text{on the warmest day of the year we against our will worked have} \]
\[ \text{‘that on the warmest day of the year we have worked against our will’ (Carstens 2003)} \]

“For the adjunct to “count” as a possible goal in this relation, it must have a relevant feature; I suggest this is a Case feature.” (Carstens 2003: 399)

4.3. Reconstruction
(32) Adjuncts entering Feature-Sharing are visible to binding even after fronting.

a. \[^{*}_{\text{ADJ}} \text{In Ben,`s office}, \text{he, lay on his desk.} \]
b. \[^{*}_{\text{ADJ}} \text{With John,`s computer}, \text{he, began to write a book.} \]
c. \[^{*}_{\text{ADJ}} \text{To Ben,`s office}, \text{he, takes the bus.} \]
d. \[^{*}_{\text{ADJ}} \text{For Mary,`s brother}, \text{I heard she, was given some clothes.} \]

Speas (1991) argues that \(\theta\)-relation plays an important role here for adjuncts to behave like arguments in terms of reconstruction. Here we assume that \(\theta\)-relation is encoded with \(\theta\)-features in syntax (Hornstein 1999).

4.4. Scrambling
(33) Adjuncts entering Feature-Sharing are visible to long-distance scrambling.

a. \[^{\sqrt{}}_{\text{ADJ}} \text{Naze], Mary-ga [_{t_{\text{John-ga sono setu-o sinziteiru ka}] sitteiru.} \]
Why Mary-Nom John-Nom that theory-Acc believes Q knows
‘Mary knows why John believes in that theory.’ (Bošković & Takahashi 1998)
b. \[^{\sqrt{}}_{\text{ADJ}} \text{Kyuuni-sika], Mary-ga [John-ga}_{t_{1 \text{nakeda-nak-atta to}] itta.} \]
Suddenly-NPI Mary-Nom John-Nom start to cry-Neg-Past that said
‘Mary said that John only suddenly started crying.’ (Boeckx & Sugisaki 1999)

Bošković & Takahashi (1998) proposes that these instances of long-distance scrambling of adjuncts are licensed (as LF-lowering) by features like Wh or Neg.

4.5. Ellipsis
(34) Adjuncts entering Feature-Sharing are visible to ellipsis.

John-wa zibun-no tukue-ni hon-o oita. \quad \sqrt{\text{Bill-wa <PP> ronbun-o oita.}}
John-Top self-Gen desk-Dat book-Acc put. \quad \text{Bill-Top article-Acc put}
‘John put a book on John’s desk.’ \quad ‘Bill put an article on Bill’s desk.’

The adjunct involved here is the so-called “pseudo-argument” undergoing selection by V, whose abstract relationship can in turn be implemented with selection-features (Müller 2010).

4.6. Summary
(35) **Descriptive Generalization**: Adjuncts entering Feature-Sharing become \textit{visible} to syntax.
5. \{XP, YP\} + Feature-Sharing = “One-Peaked” Structure

In the section, we account for the descriptive generalization (35) with crucial appeal to Chomsky’s (2013) Labeling Algorithm. In 5.1. we first construct the proposal and then in 5.2. the facts on visible adjuncts are explained.

5.1. Proposal

(36) Frampton & Gutmann (2000): Agreement = Feature-Sharing

\[
\text{Agree (XP}_{[F, +]}, \text{YP}_{[F, +]}) = \text{XP}_{[F, +]}, \text{YP}_{[F, +]}
\]

(37) Chomsky (2013): Labeling via Feature-Sharing

\[
\text{\{XP}_{[+F]}, \text{YP}_{[+F]}\}
\]

\[\star \text{Minimal-Search} \]

\[\sqrt{\text{Feature-Sharing}}\]

(38) **Main idea**: Agreement triggers Labeling via Feature-Sharing.

\[
\text{XP}_{[+F]} \quad \text{YP} (= \text{ADJ})_{[+F]} \quad \sqrt{\text{LA (\{XP, YP\})}} \quad \text{FP} \\
\]

(39) Now since the adjunction structure \{XP, YP\} is labeled as FP, the next application of Merge of Z can target at the entire labeled object.

\[
\text{Z} \quad \text{FP} \quad \sqrt{\text{Merge (Z, FP)}} \quad \text{ZP} \\
\]

\[
\text{XP}_{[+F]} \quad \text{YP} (= \text{ADJ})_{[+F]} \\
\]

\[
\text{XP}_{[+F]} \quad \text{YP} (= \text{ADJ})_{[+F]} \quad \text{FP} \\
\]

(40) **Proposal**: Adjunction + Feature-Sharing have “one-peaked” structure.
5.2. Deriving Visibility of Adjuncts

5.2.1. Island

(41) The fact that ‘adjuncts entering Feature-Sharing are visible to extraction’ follows because adjuncts do not undergo Transfer to reduce “two-peaked” structure.

5.2.2. Intervention

(42) The fact that ‘adjuncts entering Feature-Sharing are visible to intervention’ follows because adjuncts do not undergo Transfer to reduce “two-peaked” structure.

5.2.3. Reconstruction

(43) The fact that ‘adjuncts entering Feature-Sharing are visible to binding even after fronting’ follows because adjuncts do not undergo Transfer to reduce “two-peaked” structure.
5.2.4. Scrambling
(44) The fact that ‘adjuncts entering Feature-Sharing are visible to long-distance scrambling’ follows because adjuncts do not undergo Transfer to reduce “two-peaked” structure.

5.2.5. Ellipsis
(45) The fact that ‘adjuncts entering Feature-Sharing are visible to ellipsis’ follows because adjuncts do not undergo Transfer to reduce “two-peaked” structure.

5.3. Summary
(46) We have shown that the second descriptive generalization about visible adjuncts, about which Late-Merge and Pair-Merge are silent, can also follow from simplest Merge and Chomsky’s (2013) Labeling Algorithm. That is, elimination of Late-Merge and Pair-Merge is desirable both empirically and theoretically.

6. Lebeaux vs. Lasnik on Anti-Reconstruction
(47) The current proposal makes the following bidirectional predictions:

   a. If an adjunct is invisible to extraction, it should be also invisible to binding.
   b. If an adjunct is visible to extraction, it should be also visible to binding.

(48) An apparent counterexample:
Both Relative Clauses (RCs) and Nominal Complement Clauses (NCCs) are an island for extraction (i.e. Complex NP Constraint), but they are different in binding.

   a. *He believes the claim [NCC that John is nice].
   b. *He likes the story [RC that John wrote].
   c. *Whose claim [NCC that John is nice] did he believe?
   d. √Which story [RC that John wrote] did he like?

(Lebeaux 1988)
(49) Lasnik (1998) argues that Lebeaux’s asymmetry between RCs and NCCs is an illusion.

a. √Which piece of evidence \([\text{NCC}]\) that John was guilty did he successfully refute?
b. √How many arguments \([\text{NCC}]\) that John’s theory was correct did he publish?
c. √Which proof \([\text{NCC}]\) that Mary’s theory is superior to John’s did she present?

(50) Donati & Cecchetto (2011) independently shows that NCCs are actually an adjunct based on three diagnostics such as 0-Criterion exemption, constituency, and islandhood.

7. Conclusion
(51) If the proposed system is on the right track, we may:

(i) eliminate structure-building operations specific to adjuncts such as Late-Merge (Lebeaux 1988; Stepanov 2001) and Pair-Merge/SIMPL (Chomsky 2004), keeping Merge simplest; Merge \((\alpha, \beta) = \{\alpha, \beta\}\)

(ii) capture Chomsky’s (2004) original insight that adjuncts are “on a separate plane” intuitively

(iii) derive inertness of adjuncts regarding Locality of Selection/c-command without Segment-levels

(iv) show that labeling through set-intersection is allowed by UG as one logical possibility (Citko’s 2008 Project Both; see Chomsky 1995, Ch.4 for the contrary view).

They all are a theoretically welcome result under the tenet of Minimalist Program, especially Bare Phrase Structure (Chomsky 1994).

References
Den Dikken, Marcel. 2012. The Phase Impenetrability Condition and successive cyclicity: A reconsideration. ms. CUNY.
Frampton, John and Sam Gutmann. 2000. Agreement is Feature Sharing. ms., Northeastern University.
Müller, Gereon. 2010. On Deriving CED Effects from the PIC. Linguistic Inquiry 41: 35-82.