On Pseudogapping in HPSG

Dong-woo Park
(Seoul National University)

Park, Dong-woo. 2009. On Pseudogapping in HPSG. SNU Working Papers in English Linguistics and Language 8, 1-15. This study investigates the constraints imposed on the pseudogapping in the framework of Head-Driven Phrase Structure Grammar (HPSG). Based on the existing schema to account for coordination and gapping, a new pseudogapping schema in coordination structures is proposed in this paper. In the process of capturing the constraints, new DOM lists are added and an existing DOM list is divided into two DOM lists depending on the feature of elements in each domain. Furthermore, new features SEP and INC are introduced. SEP is used for distinguishing prepositions which should be located in the same domain with the following NPs from those which can be separated from the following NPs. INC feature determines whether overlapping adverbs are in non-empty lists or not. Pseudogapping occurs not only in coordination structures, but in comparative or subordination structures. Thus, this paper introduces a pseudogapping schema that can be applied to all structures mentioned above. (Seoul National University)

Keywords: pseudogapping, HPSG, DOM lists, SEP feature, INC feature, pseudogapping schema.

1. Introduction

English has ellipsis like these examples;

(1) a. Sluicing:
   She read something, but she won't say what \[vP\].
   b. Verb Phrase Ellipsis:
   She read something and he did \[vP\] too.
   c. Pseudogapping:
   She'll read something to Sam, but she won't \[vP\] to Billy.
   d. Gapping:
   Some read something to Sam and others \[vP\] to Billy.
2

2. Previous studies


Takahashi (2003) compares two existing approaches to explain pseudogapping. The first one is the Heavy NP Shift (HNPS)
approach (Jayaseelan 1990). This approach captures pseudogapping as the result of vP deletion, which applies right after the application of HNPS. In (4), the paper first moves to the right, out of vP, just like a brand-new toy in (3), and then vP is elided.

(4) We gave t₁ to John on Friday [a brand-new toy].

(5) Although John wouldn't give to Bill the book, he would [vP give[to-susan] the paper].

This method is faced with two obstacles. The first one is that the first object, an indirect object, in double object constructions cannot undergo HNPS. Despite this, the grammaticality of (6) is not degraded.

(6) Although John wouldn't give Bill the book, he would [give-the-book to-susan] Susan.

The second obstacle is that more than one item cannot undergo HNPS in a clause. (7) is ungrammatical, since both indirect object and direct object undergo HNPS. In contrary, (8) is grammatical even though both indirect object, Susan, and direct object, a paper, seem to have undergone HNPS.

(7) *John gave tt₁ tt₂ yesterday [the tall man ]₁ [the book written by the professor at MIT]₂.

(8) ?Although John would give Mary a book, he wouldn't give Susan a paper.

The second approach to pseudogapping is the Object Shift approach (Lasnik 1999). Unlike the HNPS approach, Object Shift is a leftward movement. In (9), Susan moves to the left, out of vP and the rest of the vP deletes.

(9) ... and he would [Susan [vP give-the-book]]

(10) ... and he would [the paper [vP give-Bill]]

1) Lasnik (1999) claims that this sentence is ungrammatical because a direct object
However, a direct object cannot cross over an indirect object by Object Shift (OS). Even though (10) should be ungrammatical under the OS approach, it is grammatical and cannot be accounted for with OS.

To overcome the insufficient explanation, Takahashi (2003) suggests an Eclectic Approach. This approach is the union of Object Shift and HNPS and proper to explain the remnants, as in (11).

(11) Although he wouldn’t give the book to Bill, he would [VP the paper [VP - - -] to Susan]]

Takahashi (2003) explains pseudogapping by allowing Object Shift. But one thing he has to prove is the position the object moves to. A new phrase might be inserted between TP and vP. It does not matter what that phrase is. Suppose that position is in the FocP (Focus Phrase). In order for Foc to attract the object, there should be at least one EPP in Foc.

(12) a. Mary gave Susan a lot of money and John will give Bill a lot of advice.

b. Mary gave Susan a lot of money and John will give Bill a lot of advice.

According to Takahashi (2003), both (12a) and (12b) are grammatical. In (12a), Bill undergoes OS, and a lot of advice HNPS, followed by vP ellipsis. In (12b), every derivation step is identical with (12a), except that Bill does not move to the higher maximal projection. That means Foc in (12a) has an EPP, while that in (12b) does not. Of course, cross-linguistically, whether an EPP is in certain phrase or not depends on each language. For example, in languages which allow wh-movement such as English, C has an EPP. On the other hand, wh-in situ languages like Chinese do not have an EPP in C. However, it is difficult to think out the case that existence of EPP in certain lists wholly depends on its

is a remnant. However, it can be grammatical under certain circumstances according to Baltin (2003) and Bowers (1998).
circumstances. That is too arbitrary. What if (12b) does not have FocP? This approach also confronts the same situation, arbitrariness. That is, it is not clear when TP selects FocP as a complement.


In the framework of HPSG, Beavers and Sag (2004) propose a schema, outlined in (13) to explain all kinds of coordination structures.

\[
\text{(13) } \quad \begin{array}{c}
\text{AB}_{n-1} \quad \ldots \quad B_2 B_1 CB_1 D \\
\text{AB}_n D \\
\text{AB}_{n-1} \quad \ldots \quad B_2 B_1 CB_1 D \\
\text{AB}_{n-1} \\
\vdots \\
\text{AB}_2 B_1 CB_1 D \\
\text{AB}_2 D \\
\text{AB}_2 CB_1 D \\
\text{AB}_1 D \\
C \\
\text{AB}_1 D
\end{array}
\]

This schema suggests branching \( n \)-ary coordinate structures. A, B, and D are strings and C is a coordinator. Furthermore, A and D can be either empty or non-empty. When A and D are all empty, this case produces (14a), the constitute coordination. If A is non-empty and B is empty, (14b), the Argument Cluster Coordination is produced. And Right Node Raising is produced when only A is empty as in (14c). At last, (14d) shows the case that both A and B are non-empty.

\[
\text{(14) a. Constituent Coordination} \\
\text{John, Bill and Mary} \\
\text{b. Argument Cluster Coordination (ACC)} \\
\text{gave a dog a bone and a policeman a flower} \\
\text{c. Right Node Raising (RNR)}
\]
Sandy cooked and Mary ate a pizza.

And Beavers and Sag (2004) employ the DOM list device, which was first suggested in linearization theory (Reape 1994). DOM list was devised to allow elements in sentences to change their positions. Furthermore, it can be used to make it possible to enable some elements in the daughter’s DOM lists not to be present in the mother’s DOM lists.

A new single HPSG schema is introduced to explain all phenomena in (14). It can be encoded in (15) as follows.

\[(15) \text{ conj-ext} \Rightarrow \]

As shown above, A in the mother’s DOM list comes from the first conjunct, while D comes from the second conjunct. And the different elements in each conjunct \(B_1\) and \(B_2\) in the daughters’ DOM lists are preserved in the mother’s DOM list. A coordinator is represented as C and the right conjunct with a coordinator has \([\text{CRD +}]\). On the other hand, the left conjunct has \([\text{CRD -}]\). This material prevents us from predicting sentences in (16) correctly ungrammatical.
(16) a. "Jan walks. chews gum."
   b. "And Jan walk, and Jan chews gum."

2.3. Chaves (2005)

Even though Beavers and Sag (2004) made a single comprehensive schema in order to account for coordination structures including CC, ACC, and RNR, Chaves (2005) points out that it fails to explain sentences in (17), called gapping, since shared elements can be located in non-peripheral positions.2)

(17) a. John will bring dessert, and Mary, wine.
   b. Yesterday we traveled sixty miles, and on the day before, fifty.
   c. Ann reads stories to her kids, and Maris, to the students.

(18) conj-cat

\[
\begin{align*}
\text{MTR} & \quad \text{DOM} \quad \text{SYN} \quad \text{DTR} \quad \text{CRD} \\
& \quad \left[ \begin{array}{c}
A_1 \otimes L_1 \otimes R_1 \otimes F \otimes [L_2 \otimes R_2] \otimes ne - list \otimes D_1 \\
\quad \text{DOM} \quad \text{SYN} \quad \text{CRD} \end{array} \right] \\
& \quad \left[ \begin{array}{c}
[D_1 \otimes (\text{SYN conj}) \otimes \text{DTR} \otimes \text{CRD}] \quad \text{SYN} \quad \text{CRD} \\
\quad \text{DOM} \quad \text{SYN} \quad \text{CRD} \end{array} \right]
\end{align*}
\]

(18) shows the constraints imposed on gapping. The big difference between (15) and (18) is that there are paired elements sharing their HEAD and SEM and the latter elements are elided, except the peripheral elements A and D. They are I_1 and I_2. Gapping happens when the shared non-peripheral lists I_1 and I_2 are resolved as non-empty.

---

2) Remind the schema in Beavers and Sag (2004). The sharing and slipped elements are located only in peripheral positions, A and D. It fails to capture the phenomenon that non-peripheral elements can be slipped.
Furthermore, he mentioned additional constraints, \( h_f \_share \) and \( h_s \_share \) as follows.

\[(19) \quad h_f \_share([1], [2]) \leftarrow ([1]=()) \land ([2]=()) \lor \]
\[= \left[ \text{FRM}\mid \text{SYN}\mid \text{HD} \right] ([1]) \land \left[ \text{SYN}\mid \text{HD} \right] ([2]) \land h_f \_share([1],[2]) \]

\[(20) \quad h_s \_share([1], [2]) \leftarrow ([1]=()) \land ([2]=()) \lor \]
\[= \left[ \text{SYN}\mid \text{CAT}\mid \text{HEAD} \right] ([1]) \land \left[ \text{SYN}\mid \text{CAT}\mid \text{HEAD} \right] ([2]) \land h_s \_share([1],[2]) \]

In (21), the non-peripheral DOM list \( I_2 \) is not empty.

(21) John likes caviar, and Mary, beans.

\[
\begin{align*}
\text{MTT \mid DOM} & \leftarrow \left[ \text{FRM}\mid \text{SYN}\mid \text{HD} \right] ([\text{John}]) \land ([\text{likes}]) \land ([\text{caviar}]) \land \\
\text{DTRS} & \left[ \text{DOM}([\text{John}]) \land ([\text{likes}]) \land ([\text{caviar}]) \right] \land \\
& \left[ \text{DOM}([\text{Mary}]) \land ([\text{likes}]) \land ([\text{beans}]) \right]
\end{align*}
\]

A larger gap can be represented as follows;

(22) Mia can help me today, and Jess, tomorrow.

\( l_2 = \langle \text{[can]}, \text{[help]}, \text{[me]} \rangle, \quad l_2 = \langle \text{[tomorrow]} \rangle \)

In (22), overlapping elements can, help, me in \( I_2 \) can be elided.

(22) is sufficient for explaining continuous gapping. However, Jackendoff (1971) and others point out that there are another kind of gapping - discontinuous gapping as in (23).

(23) a. John kissed Susan at the party, and Peter, Mary.
    b. Dexter wants Watford to win, and Warren, Ipswich.
    c. Peter took Susan home, and John, Wendy.

Chaves (2005) modifies his first schema to account for discontinuous gapping by introducing the shuffle ‘O’ operator at the right
periphery of the internal sub-sets.

\[(24)\] \(c nj-ex \Rightarrow \)

\[
\begin{align*}
\text{MTR} & \left[ \text{DOM} \left[ A_1 \circ L_1 \circ L_2 \circ L_3 \circ C_2 \right] \left[ \text{ne-list} \oplus D_2 \right] \right] \\
\text{SYN} & \left[ \right] \\
\text{CRD} - & \\
\text{DTRS} & \left[ \text{DOM} \left[ \left( \text{SYN} \ c nj \right) \right] \left[ A_2 \circ L_2 \circ L_3 \circ C_3 \right] \left[ \text{ne-list} \oplus D_2 \right] \right] \\
\text{SYN} & \left[ \right] \\
\text{CRD} + & \\
\wedge \text{h.f.share} \left( A_1, A_2 \right) & \wedge \text{h.f.share} \left( D_1, D_2 \right) \\
\wedge \text{h.s.share} \left( T, T \right) & \wedge \text{h.s.share} \left( P, P \right) \\
\wedge D_2 = \text{ne-list} & \Rightarrow \left[ \text{SYN} \left[ \text{HD verb, MRK none} \right] \right] \in \left[ D \right]
\end{align*}
\]

3. Proposals

3.1. Additional constraints on pseudogapping

Chaves (2005) already pointed out the problems of Beavers and Sag (2004). His alternative is to add the new non-peripheral list \( H_2 \) that consists of reduplicated elements, which can be elided. His gapping schema (18) has no specific constraint — what can be the remnants? Thus, his proposal incorrectly predicts that sentences below are all grammatical.

\[(25)\] a. *You feel relieved, but I do jubilant
b. *Rona sounds annoyed, and Sue did frustrated.

(Lasnik 1999)

These adjectival remnants can be ruled out through Jayaseelan(1990) and Lasnik's (1999) analysis. They all assume NP movement — HNPS and O6. However, relieved and jubilant are not nouns but adjectives so that they do not undergo HNPS and O6. Adjectives, the complements of verb are elided along with the vP ellipsis.
Without any constraint on the part of speech, it is not enough to apply the gapping schema to pseudogapping. The fact that AP cannot be the only remnant in pseudogapping constructions is also supported by the examples below.

(27) a. I made John happy and she did make Mike upset.
    b. *I made John happy but she did make John upset.

Roughly, the definition of pseudogapping is generally assumed to be the deletion of vP except an auxiliary verb and a argument or arguments. However, the definition of pseudogapping should be more specific in that not all kinds of arguments can be the remnant of vP-ellipsis as shown above.

According to Lasnik (1999), one of the pseudogapping puzzles which the gapping schema cannot explain but we have to solve is the difference between (28a, b) and (28c, d). In (28a) and (28b), prepositions that two adjuncts have in common are elided as the gapping schema predicts. However, in (28c) and (28d), prepositions are not elided even though they are in the same context.

(28) a. John speaks to Bill and Mary should speak to Susan.
    b. John talked about linguistics and Mary will talk about philosophy.
    c. John swam beside Bill and Mary did swim beside Susan.
    d. John stood near Bill and Mary should stood near Susan.

The examples in (28) show that the object of some prepositions can be a remnant of pseudogapping, while that of others cannot. This distinction is closely related to two different kinds of prepositions. One is argument-marking prepositions and the other is predicative prepositions. The former does not contribute anything to the
meaning of sentences semantically, i.e. its RESTRICTION is empty, sharing the values of MODE and INDEX with those of its complement. And the latter has its own MODE value, INDEX value, and non-empty REST. Thus, this indicates that argument-marking prepositions are transparent and can be elided with verb, while predicative prepositions are non-transparent and cannot be elided with verb. This difference is shown in a variety of constructions including passives and wh-questions.

(29) a. Bill was spoken to by John.
   b. Linguistics was talked about by John.
   c. * Bill was swum beside by John.
   d. * Bill was stood near by John. (Lasnik 1999)

(30) a. Who does John speak to?
   b. Who did John talk about?
   c. * Who did John swim beside?
   d. * Who did John stand near?

(29) and (30) indicate that the object of argument-marking prepositions can move to the subject position in pseudo-passive and even in wh-question constructions, while that of predicative prepositions cannot be extracted in both constructions. In addition, the different distributions of reflexives support the justification of distinguishing them. However, in some pseudo-passive sentences, the object of predicative preposition can be extracted to the subject position. To classify them clearly, a new additional feature distinguishing the two group is needed. I will call it SEP(arable) feature. All argument marking prepositions have [+SEP], while some predicative prepositions have [+SEP] and the others have [-SEP].

When predicative prepositions with [-SEP] are paired in the coordinated clause, they cannot be elided in pseudogapping constructions. And this constraint should be specified. Otherwise, some pseudogapping schema cannot rule out sentences such as (28c, d).

3.2. Pseudogapping schema
Pseudogapping is more complex and more peculiar than gapping due to additional and specific constraints to explain the phenomena. By adopting the DOM list device, we can solve the problem, arbitrariness because the daughter's DOM lists need not realize in the mother's node, when there are reduplicated elements in clause. The problem is how we can rule out the case where overlapping elements cannot be elided.

This problem can be solved by dividing $R_2$ into two DOM lists, $R_z$ and $R_{z'}$. Elements which can be included in $R_z$ are NPs and PPs, whose heads have $[\text{SEP}+]$. Meanwhile, APs and PPs, whose heads have $[\text{SEP}-]$ are not included in $R_z$ but $R_{z'}$.

Adverbs make the problem worse because of their position. It is well known that there are two kinds of adverbs - high adverbs and low adverbs. High adverbs are can be attached to positions higher than vP, while low adverbs within vP. Sentences in (31) include both a high verb and a low verb and the high verb cannot be elided.

(31) a. I gave John a book yesterday, fortunately, and Mary did give Tim a pen yesterday, fortunately.

b. *I gave John a book yesterday, fortunately, and Mary did give Tim a pen yesterday, fortunately.

(31a) suggests that the $L_2$, $L_2$, $R_2$, $P_2$ are not sufficient to account for complex sentences. In (28), yesterday is supposed to belong to $P_2$. However, fortunately needs a new DOM list, because $R_2$ includes only Tim a pen. I call a new list for fortunately $Q_2$, containing overlapping elements, which should not be elided because fortunately is not located within vP. In order to distinguish high adverbs from other adverbs, I will suggest high adverbs have INC(idental adverbs) feature and the others do not. Then only elements which have INC feature can be located in $Q_1$ and $Q_2$. $Q_1$ and $Q_2$ can be either empty or non-empty and can occur anywhere out of vP.

All constraints mentioned above put together, pseudogapping in coordination structure can be described as follows:
(32) cnj-ex

\[
\begin{align*}
\text{MTR} & : \text{DOM } [a \circ b \circ c \circ d \circ e \circ f \circ g \circ h \circ i \circ j \circ k \circ l \circ m \circ n \circ o \circ p \circ q \circ r \circ s \circ t \circ u \circ v \circ w \circ x \circ y \circ z \circ A \circ B \circ C \circ D \circ E \circ F \circ G \circ H \circ I \circ J \circ K \circ L \circ M \circ N \circ O \circ P \circ Q \circ R \circ S \circ T \circ U \circ V \circ W \circ X \circ Y \circ Z \circ \text{ne - list } @ D @] \circ @ Q \]
\text{SYN} & : 0
\end{align*}
\]

\[
\begin{align*}
\text{DTR} & : \text{DOM } [a \circ b \circ c \circ d \circ e \circ f \circ g \circ h \circ i \circ j \circ k \circ l \circ m \circ n \circ o \circ p \circ q \circ r \circ s \circ t \circ u \circ v \circ w \circ x \circ y \circ z \circ A \circ B \circ C \circ D \circ E \circ F \circ G \circ H \circ I \circ J \circ K \circ L \circ M \circ N \circ O \circ P \circ Q \circ R \circ S \circ T \circ U \circ V \circ W \circ X \circ Y \circ Z \circ \text{ne - list } @ D @] \circ @ Q \]
\text{CRD} & : \text{ne - list } @ D @] \circ @ Q
\end{align*}
\]

\[
\begin{align*}
\text{h_f_share}(a, b) \land h_f_share(b, d) \land h_s_share(e, f) \land h_s_share(f, g)
\end{align*}
\]

\[
\begin{align*}
\text{R}_2 = (\text{np}, P \text{[hd sep +]}), \quad \text{R}_2 = (\text{ap}, P \text{[hd sep +]}), \quad \text{INC -}
\end{align*}
\]

\[
\begin{align*}
\text{INC} = \text{INC} \lor \text{INC} \lor \text{INC} \lor \text{INC} \lor \text{INC} \lor \text{INC} \lor \text{INC}
\end{align*}
\]

\[
\begin{align*}
\text{ne - list } @ D @] \circ @ Q
\end{align*}
\]

(32) contains the new constraint \( R_2 \rightarrow R_2 \) which indicates \( R_2 \) can be remnant only when \( R_2 \) is remnant. That is, this prevents (27b) - AP is a unique remnant - from being predicted grammatical and allows (27a) to be judged as grammatical sentence, even though it sounds awkward.

(32) is within the limit of the coordination structures. However, pseudogapping also occurs in subordination or comparative structures, as in (2) and (3). Thus, in order to cover as many as phenomena, the pseudogapping schema which can be applied to any structure is needed. It can be represented as follows.

(33) pseudogapping-ex

\[
\begin{align*}
\text{MTR} & : \text{DOM } [a \circ b \circ c \circ d \circ e \circ f \circ g \circ h \circ i \circ j \circ k \circ l \circ m \circ n \circ o \circ p \circ q \circ r \circ s \circ t \circ u \circ v \circ w \circ x \circ y \circ z \circ A \circ B \circ C \circ D \circ E \circ F \circ G \circ H \circ I \circ J \circ K \circ L \circ M \circ N \circ O \circ P \circ Q \circ R \circ S \circ T \circ U \circ V \circ W \circ X \circ Y \circ Z \circ \text{ne - list } @ D @] \circ @ Q \]
\text{SYN} & : 0
\end{align*}
\]

\[
\begin{align*}
\text{DTR} & : \text{DOM } [a \circ b \circ c \circ d \circ e \circ f \circ g \circ h \circ i \circ j \circ k \circ l \circ m \circ n \circ o \circ p \circ q \circ r \circ s \circ t \circ u \circ v \circ w \circ x \circ y \circ z \circ A \circ B \circ C \circ D \circ E \circ F \circ G \circ H \circ I \circ J \circ K \circ L \circ M \circ N \circ O \circ P \circ Q \circ R \circ S \circ T \circ U \circ V \circ W \circ X \circ Y \circ Z \circ \text{ne - list } @ D @] \circ @ Q \]
\text{CRD} & : \text{ne - list } @ D @] \circ @ Q
\end{align*}
\]

\[
\begin{align*}
\text{h_f_share}(a, b) \land (\text{np} \land P \text{[hd sep +]}) \land \text{h_f_share}(b, d) \land \text{h_s_share}(e, f) \land \text{h_s_share}(f, g)
\end{align*}
\]

\[
\begin{align*}
\text{R}_2 = (\text{np}, P \text{[hd sep +]}), \quad \text{INC} \rightarrow \text{INC} \rightarrow \text{INC} \rightarrow \text{INC} \rightarrow \text{INC} \rightarrow \text{INC} \rightarrow \text{INC}
\end{align*}
\]

\[
\begin{align*}
\text{INC} = \text{INC} \lor \text{INC} \lor \text{INC} \lor \text{INC} \lor \text{INC} \lor \text{INC} \lor \text{INC}
\end{align*}
\]

\[
\begin{align*}
\text{ne - list } @ D @] \circ @ Q
\end{align*}
\]

If \( \square \) in the first conjunct is occupied with subordinate conjunction,
it is pseudogapping in the subordinate construction. Furthermore, \textit{than} can be located in \(\mathcal{E}\) in the second conjunct, when it is pseudogapping in the comparative construction.

4. Unsolved Problems

Pseudogapping is a peculiar phenomenon related to semantics as well as syntax. Consequently we have to investigate how people can process the sentences with pseudogapping. Furthermore, we cannot understand this phenomenon, ignoring the context, or pragmatics because pseudogapping can occur in the sentences standing alone in certain context in discourse as in (34).

(34) a. A : Is she suing the hospital?
    B : She \textit{is suing} the doctor.

b. A : Has he sold his collection yet?
    B : He has sold some of his paintings; I'm not sure about the rest. \hspace{1cm} \textnormal{(Halliday and Hasan 1973)}

c. A : Gee, I've never seen you on campus before.
    B : Yeah! Neither have I seen you. \hspace{1cm} \textnormal{(Lasnik 1999)}

Sentences above cannot be explained with the pseudogapping schema I suggested in the previous section.

5. Conclusion

In this paper, I proposed a new pseudogapping schema based on the gapping schema mentioned in Chaves (2005). The new schema can capture the insufficient aspect dividing the DOM list \([I_2]\) into \([R_e]\) and \([R_{e''}]\). In addition, by introducing SEP feature, prepositions which can be contained in \(I_2\) are separated from those which cannot be contained in \(I_2\). Furthermore, a new DOM list \(Q\) is introduced for high adverbs that have INC feature.

However, as I mentioned in the previous section, semantic and pragmatic analysis is mandatory for the complete understanding of
pseudogapping.

References


Dong-woo Park
pdw83@snu.ac.kr