

A new generalization about gapping*

Maziar Toosarvandani
University of California, Santa Cruz

June 9, 2015

Gapping removes the finite element (T and its host), possibly along with additional material, in the second and subsequent coordinates of a coordination structure, leaving behind two remnants. (The material that has gone missing—in 1, the finite auxiliary *had* and the main verb *ordered*—is represented with ‘ Δ ’.)

- (1) Some had ordered mussels, and others Δ swordfish.

Building on a long tradition of earlier work, Johnson (2009:293) identifies three unique properties that distinguish gapping from superficially similar elliptical constructions, such as pseudogapping, e.g. *Some had ordered mussels, and others had Δ swordfish*.

First, gapping is restricted to coordination structures (2) (Jackendoff 1971:22, Hankamer 1979:18f.). Second, the gap in gapping cannot be embedded (3) (Hankamer 1979:19). Third, the antecedent in gapping cannot be embedded (4) (Hankamer 1979:20).

- (2) * Some had eaten mussels, because others Δ shrimp.
(3) * Some had eaten mussels, and she claims that others Δ shrimp.
(4) * She’s said Peter has eaten his peas, and Sally Δ her green beans, so now we can have dessert.
Intended: ‘She has said that Peter has eaten his peas; Sally has eaten her green beans.’

(Johnson 2009:293)

Crucially, 4 is ungrammatical under an interpretation in which only the antecedent clause—not the gapped clause—is embedded. (The sentence is grammatical with a different interpretation, one that is not relevant here, where the entire conjunction is embedded.)

*I am immensely grateful to Nate Clair, Elizabeth Coppock, Annahita Farudi, Danny Fox, Kyle Johnson, Laura Kertz, Ben Merikli, David Pesetsky, Craig Sailor, Bern Samko, and audiences at Cornell University, the Massachusetts Institute of Technology, Northwestern University, the University of California, Santa Cruz, the University of Rochester, and Wayne State University for their questions and suggestions. This research was assisted by a New Faculty Fellowship from the American Council of Learned Societies, funded by the Andrew W. Mellon Foundation.

There are a number of theories of gapping that successfully account for at least the first two of these properties (Coppock 2001, Lin 2001, Johnson 2004, 2009). These all share one analytical ingredient: they appeal to low coordination of small verbal constituents, such as vPs, under a single T head. (See Kubota and Levine, to appear, though, for a different approach within Categorical Grammar.)¹ T goes missing in the second and subsequent coordinates because it was never present inside them to begin with. Some other mechanism gets rid of any additional material that goes missing: either some kind of ellipsis (for Coppock and Lin) or across-the-board movement (for Johnson).

As Johnson (2009:296–300) observes, low coordination easily derives the first property of gapping, because two or more vPs can be embedded under a single T head only through coordination. It also derives the second property: a single T head cannot be shared with a vP that is embedded *inside* a coordinate. But, deriving the third property is less straightforward. Under any theory of gapping that uses low coordination, the notion of an ‘antecedent’ is a complex one. Since T is removed through a different mechanism than any additional material that goes missing, the ‘antecedent’ of T is analytically distinct from the ‘antecedent’ of additional missing material.

I will take a closer look here at the third property of gapping in light of low-coordination theories of gapping. In Section 1, I show that while the syntax of low coordination can account for why the antecedent of T cannot be embedded inside the first coordinate, it has nothing to say about why the antecedent of any additional missing material also cannot be embedded. Then, in Section 2, I propose a new generalization about what can be embedded inside the first coordinate, which I call the No Correlate Embedding Generalization, whose source does not lie in the syntax of low coordination. In Section 3, I suggest that this generalization arises from a constraint on the focus structures of the coordination structures in which gapping appears. Finally, in Section 4, I examine a prediction of this account: Low-Coordinate Parallelism should be satisfied when the embedding material in the first coordinate is itself contained inside a focus.

1. Taking a closer look at the third property

In its traditional formulation, the third property of gapping is usually stated along the following lines: The antecedent in gapping cannot be embedded inside the first coordinate (see, for instance, Hankamer 1979:20). Theories of gapping that appeal to low coordina-

¹As a consequence, the subject of the first coordinate must raise asymmetrically to Spec-TP, while the subjects of other coordinates stay in situ. This accounts for Siegel’s (1987) observation that the subject of the first coordinate c-commands—and hence can bind into—the subject of subsequent coordinates.

- (v) No woman₁ can join the army, and her₁ girlfriend Δ the navy.

(Johnson 2009:293)

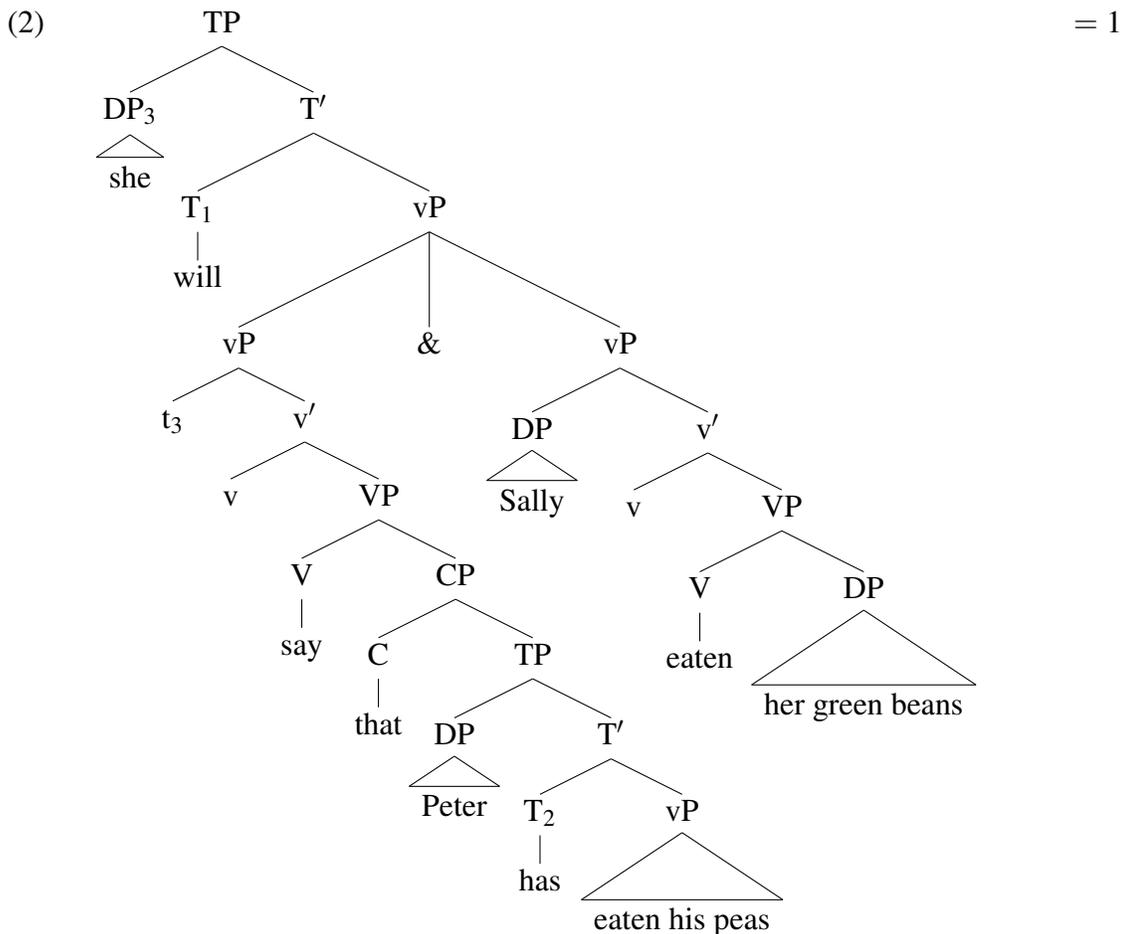
Lin (2002:58–94) offers one way of understanding this asymmetrical subject movement (see also Johnson 2004:41–49). If the Coordinate Structure Constraint holds of LF representations (Fox 2000:51–58), and if A-movement can reconstruct, it will not be subject to the island constraint in the first place.

tion remove T and any additional material in different ways. Consequently, the antecedent of T is analytically distinct from the antecedent of any additional missing material.

Does the third property of gapping hold for both antecedents? We can start by looking at an example in which only T goes missing. In 1, the antecedent of T cannot be located inside an embedded clause in the first coordinate. (Here and below, the antecedent of T is underlined.)

- (1) * She will say that Peter has eaten his peas, and Sally Δ eaten her green beans.
 Intended: ‘She will say that Peter has eaten his peas; Sally has eaten her green beans.’

It is easy to see how this follows from the syntax of low coordination. T₂, contained within the embedded clause in the first coordinate, cannot be shared with the second vP coordinate.



To test whether the antecedent of any additional missing material can be embedded, we need an example like Johnson’s (2009) from 4 above. In the simplified version below, the additional missing material in the second coordinate—the main verb *eaten*—is identical to

material that is contained within an embedded clause in the first coordinate. (The antecedent of additional missing material is boxed.)

- (3) * She has said that Peter has eaten his peas, and Sally Δ her green beans.
Intended: ‘She has said that Peter has eaten his peas; Sally has eaten her green beans.’

Since 3 is ungrammatical under the intended interpretation, it appears that any additional material that goes missing also cannot be embedded inside the first coordinate.

Importantly, the syntax of low coordination does *not* account for why this should be. In 3, the second coordinate can share the T of the matrix clause. So, the restriction on embedding the antecedent of additional missing material must arise from another source. This might, in principle, be a product of the operation that removes this material, i.e. ellipsis or across-the-board movement. But I suggest below that it is in fact a specific case of a more general constraint on gapping, which prevents *correlates* from being embedded inside the first coordinate.

2. The No Correlate Embedding Generalization

With this understanding from the syntax of low coordination for why the antecedent of T cannot be embedded inside the first coordinate, the sentence in 4—which is parallel to 3 except that the main verb does not go missing in the second coordinate—should be fully grammatical.

- (4) * She has said that Peter has eaten his peas, and Sally Δ drunk her milk.
Intended: ‘She has said that Peter has eaten his peas; Sally has drunk her milk.’

Since T of the *matrix* clause in the first coordinate—the auxiliary *has*—selects for participial morphology, it can be shared with the second coordinate. (In more traditional terminology, the T that has gone missing can find as its antecedent the T of the matrix clause in the first coordinate.)

So why isn’t 4 grammatical? I propose that the traditional generalization about embedding the antecedent of gapping should be subsumed by a more inclusive generalization. The remnants stand in contrastive relationships with elements in the first coordinate: *Peter* and *eaten his peas* contrast with the correlates *Sally* and *drunk her milk*. As shown in 5, each of these can bear a pitch accent.

- (5) * She has said that [PETER] has [eaten his PEAS], and [SALLY] Δ [drunk her MILK].
Intended: ‘She has said that Peter has eaten his peas; Sally has drunk her milk.’

I propose that it is the *correlates* that cannot be embedded under material that is present only in the first coordinate.

- (6) *No Correlate Embedding Generalization*
The correlates in gapping cannot be embedded inside the first coordinate.

It is not clear how the syntax of low coordination could ever account for the No Correlate Embedding Generalization, as the correlates do not form a natural class syntactically. They are simply the elements in the first coordinate that stand in a contrastive relationship with the remnants.

3. The role of parallelism

What else could the source for the No Correlate Embedding Generalization lie? I suggest that it arises from the same grammatical principles that enforce a contrastive relationship between each remnant and its correlate. As many have observed, gapping has a characteristic information structure that is reflected in its prosody (Kuno 1976:310, Hankamer 1979:183f., Sag 1976:287, Levin and Prince 1986, Hartmann 2001:162–166, Kehler 2002:81–100, Winkler 2005:191–194, Repp 2009:83–148).

(7) [SOME] had ordered [MUSSELS], and [OTHERS] ~~ordered~~ [SWORDFISH].

In 7, the remnants *others* and *swordfish* bear a pitch accent and contrast with the corresponding phrases in the first coordinate. These correlates are often new information and can also be realized with pitch accents, as *some* and *mussels* are in 7.

I propose that this contrastive relationship arises from a more general constraint on the focus structures of the coordination structures in which gapping occurs. This constraint can be stated using Rooth’s (1985, 1992) alternative semantics for focus.

(8) *Low-Coordinate Parallelism*
For vPs α and β , if α and β are coordinated, $\llbracket \alpha \rrbracket \in ALT(\beta)$ and $\llbracket \beta \rrbracket \in ALT(\alpha)$.

The alternative set for a linguistic expression, given by the function *ALT*, is the set of ordinary meanings derived by substituting focus-marked constituents with every expression of the same type. Low-Coordinate Parallelism thus requires that the coordinates in a low-coordination structure be alternatives to one another.

As a consequence, assuming that each remnant contains a focus, the coordinates in a low coordination will have to have parallel focus structures. In addition, any nonfocused material they contain will have to be semantically identical. To illustrate this, the focus structure for 7 is given below.

(9) had $[_{VP1}$ [SOME]_F ordered [MUSSELS]_F] and $[_{VP2}$ [OTHERS]_F ~~ordered~~ [SWORDFISH]_F]
= 7

Low-Coordinate Parallelism is satisfied, because the first coordinate is in the alternative set of the second coordinate, and vice versa.²

²Importantly, this calculation requires the subject of the first coordinate to be interpreted within the first coordinate, as shown in 7. But for the purposes of variable binding, it must also take scope over the entire coordination, so that it can bind into the second coordinate (see footnote 1). Erlewine (2014:101–105) offers one way of resolving this apparent paradox. He argues that it is possible to interpret a lower copy

- (1) a. $\llbracket \text{vP}_1 \rrbracket = \mathbf{order}(\mathbf{mussels})(\mathbf{some}) \in ALT(\text{vP}_2) = \{\mathbf{order}(x)(y) \mid x, y \in D_e\}$
 b. $\llbracket \text{vP}_2 \rrbracket = \mathbf{order}(\mathbf{swordfish})(\mathbf{others}) \in ALT(\text{vP}_1) = \{\mathbf{order}(x)(y) \mid x, y \in D_e\}$

In fact, the alternative sets for the two coordinates— $ALT(\text{vP}_1)$ and $ALT(\text{vP}_2)$ —are the same, since they have the same focus structure and the nonfocused material in each coordinate is semantically identical.

It should now be clear how Low-Coordinate Parallelism derives the No Correlate Embedding Generalization. If the first coordinate contains any nonfocused material that is not present in another coordinate, it will not be satisfied. To illustrate this, the focus structure for the ungrammatical gapping sentence in 5, in which the correlates are embedded in the first coordinate, is given below.

- (2) * has $\llbracket \text{vP}_1 \text{ she said that } [\text{PETER}]_F \text{ has } [\text{eaten his PEAS}]_F \rrbracket$, and $\llbracket \text{vP}_2 [\text{SALLY}]_F [\text{drunk her MILK}]_F \rrbracket = 5$

There is some nonfocused material present in the first coordinate that is not found in the second coordinate. The sentence thus violates Low-Coordinate Parallelism and is ungrammatical.

- (3) a. $\llbracket \text{vP}_1 \rrbracket = \mathbf{say}(\mathbf{eat}(\mathbf{his-peas})(\mathbf{peter}))(\mathbf{she}) \in ALT(\text{vP}_2) = \{f(x) \mid x \in D_e \wedge f \in D_{\langle e,t \rangle}\}$
 b. $\llbracket \text{vP}_2 \rrbracket = \mathbf{drink}(\mathbf{her-milk})(\mathbf{sally}) \notin ALT(\text{vP}_1) = \{\mathbf{say}(f(x))(\mathbf{she}) \mid x \in D_e \wedge f \in D_{\langle e,t \rangle}\}$

Since there is a focus on the entire VP remnant, the first coordinate is an alternative to the second coordinate. The proposition that she said that Peter ate his peas is a proposition of the form ‘ $x f$ ’, where x is some individual and f is some property of individuals. However, the second coordinate is *not* an alternative to the first coordinate: it is not a proposition of the form ‘she said that $x f$ ’.

Low-Coordinate Parallelism imposes the same parallelism constraint that Rooth’s (1992:102–107) squiggle operator (\sim) does on so-called bare remnant ellipsis (or stripping). For him, one of these operators is adjoined to the clause containing the ellipsis and another is adjoined to the antecedent clause, thus requiring the two clauses to be alternatives to one another. For this reason, it is possible to view Low-Coordinate Parallelism as simply a statement about where squiggle operators *must* adjoin: there has to be one operator adjoined to each vP coordinate in a low-coordination structure. This derives the parallelism between the coordinates in a more general fashion, in terms of the distribution of Roothian squiggle operators.

of movement, like the one the subject leaves in Spec-vP, for calculating the presuppositional meaning of the focus sensitive operator *even*. This does not affect the sentence’s at-issue content, which arises through composition with the highest copy of the subject, as usual. By analogy, Low-Coordinate Parallelism could be treated as a presupposition that is calculated at the vP-level, using the copy of the subject present within the first coordinate.

However this parallelism constraint arises, it probably holds only of low-coordination structures. As Johnson (2009:293) shows, the antecedent of VP-ellipsis in pseudogapping can be embedded relatively easily inside the first coordinate.

- (4) ?She's said Peter has eaten his peas, and Sally has Δ her green beans, so now we can have dessert.
 'She has said that Peter has eaten his peas; Sally has eaten her green beans.'
 (Johnson 2009:293)

Clearly, two full clauses are coordinated here, since a finite auxiliary (*has*) is present in the second coordinate. Low-Coordinate Parallelism must therefore not apply to full clausal coordinations, since 4 is not nearly as bad as the minimally different gapping sentence in 4.

4. A prediction

Low-Coordinate Parallelism would seem to predict, for the gapping sentence in 5, that if the embedding material in the first coordinate were contained inside a focus, it would become well-formed. For instance, the correlate in the first coordinate could be the higher subject DP—replaced by the proper name *Mike* so it bears a pitch accent more easily—and the larger VP *said that Peter has eaten his peas*. (Broad foci like this one are often associated with more than one pitch accent, even though it is in principle possible for a single pitch accent to project focus onto such a large phrase.)

- (5) ?? [MIKE]_F has [SAID that PETER has eaten his PEAS]_F, and [SALLY]_F [drunk her MILK]_F.
 Intended: 'Mike has said that Peter has eaten his peas; Sally has drunk her milk.'

With this focus structure, Low-Coordinate Parallelism is satisfied. The embedding predicate in the first coordinate is contained inside a correlate, so that there is no nonfocused material in either coordinate that is not semantically identical.

- (6) a. $\llbracket \text{vP}_1 \rrbracket = \text{say}(\text{eat}(\text{his-peas})(\text{peter}))(\text{mike}) \in$
 $ALT(\text{vP}_2) = \{f(x) \mid x \in D_e \wedge f \in D_{\langle e,t \rangle}\}$
 b. $\llbracket \text{vP}_2 \rrbracket = \text{drink}(\text{her-milk})(\text{sally}) \in$
 $ALT(\text{vP}_1) = \{f(x) \mid x \in D_e \wedge f \in D_{\langle e,t \rangle}\}$

Each coordinate here expresses a proposition of the form '*x f*', which it is easy to verify is the case.

But is 5 actually felicitous? Four native speakers of English who I consulted do not generally accept it. On a seven-point scale, three rated it as 1 or 2, and the fourth as 4 (where 1 is completely ungrammatical and 7 is completely grammatical). This is somewhat surprising since 5 has the same focus structure as a less controversial 'simple gap' (see also Siegel 1987:54, Johnson 2004:33). (Three of the same speakers rated the sentence in 7 as 6 or 7, including the fourth one above; the final speaker rated it as 5.)

(7) [PETER]_F has [eaten his PEAS]_F, and [SALLY]_F [drunk her MILK]_F.

Just as in 5 above, one remnant in 7 is a VP that corresponds to a correlate that is a VP. It satisfies Low-Coordinate Parallelism in the same way, but it is clearly well-formed.

There are two plausible sources for why speakers identify a contrast between 5 and 7, both of which are independent of Low-Coordinate Parallelism. First, as one reviewer suggests, gapping might require small correlates and remnants, each bearing a narrow focus. This might even be a more general constraint on information structure, since it evokes a constraint like Schwarzschild's (1999) AVOIDF, which favors that a focus be as small as possible. Second, it seems likely to me that the VP *eaten his peas* is a more salient and/or more frequent alternative to the remnant VP *drunk her milk* than the larger VP *said that Peter has eaten his peas*. The availability of such an alternative in the first coordinate in 5 might interfere with a speaker's ability to infer the necessary contrastive relation, contributing to its degraded status.

After taking a closer look at third property of gapping, I argued that it exhibits a hitherto unnoticed property: the No Correlate Embedding Generalization. The correlates cannot be embedded just inside the first coordinate. I proposed that this generalization might arise because of a constraint on the focus structures of the coordination structures in which gapping appears. This constraint, which I called Low-Coordinate Parallelism, requires vP coordinates to be focus alternatives to one another. It remains now to figure out why low coordinations might be subject to such an information-structural constraint in the first place, and how this constraint would interact with other constraints on focus.

References

- Coppock, Elizabeth. 2001. Gapping: In defense of deletion. *Chicago Linguistic Society (CLS)* 37:133–148.
- Erlewine, Michael Yoshitaka. 2014. Movement out of focus. Ph.D. Dissertation, Massachusetts Institute of Technology.
- Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, MA: MIT Press.
- Hankamer, Jorge. 1979. *Deletion in coordinate structures*. New York: Garland Publishing.
- Hartmann, Katharina. 2001. *Right node raising and gapping: Interface conditions on prosodic deletion*. Amsterdam: John Benjamins.
- Jackendoff, Ray S. 1971. Gapping and related rules. *Linguistic Inquiry* 2:21–35.
- Johnson, Kyle. 2004. In search of the English middle field. Ms., University of Massachusetts, Amherst, URL http://people.umass.edu/kbj/homepage/Content/middle_field.pdf.
- Johnson, Kyle. 2009. Gapping is not (VP-)ellipsis. *Linguistic Inquiry* 40:289–328.
- Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI.

- Kubota, Yusuke and Robert Levine. To appear. Gapping as hypothetical reasoning. *Natural Language and Linguistic Theory* .
- Kuno, Susumu. 1976. Gapping: A functional analysis. *Linguistic Inquiry* 7:300–318.
- Levin, Nancy S. and Ellen F. Prince. 1986. Gapping and causal implicature. *Papers in Linguistics* 19:351–364.
- Lin, Vivian. 2001. A way to undo A-movement. *West Coast Conference on Formal Linguistics (WCCFL)* 20:358–371.
- Lin, Vivian. 2002. Coordination and sharing at the interfaces. Ph.D. Dissertation, Massachusetts Institute of Technology.
- Repp, Sophie. 2009. *Negation in gapping*. Oxford: Oxford University Press.
- Rooth, Mats. 1985. Association with focus. Ph.D. Dissertation, University of Massachusetts, Amherst.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.
- Sag, Ivan A. 1976. Deletion and logical form. Ph.D. Dissertation, Massachusetts Institute of Technology.
- Schwarzschild, Roger. 1999. GIVENness, AVOIDF and other constraints on the placement of accent. *Natural Language Semantics* 7:141–177.
- Siegel, Muffy E. A. 1987. Compositionality, case, and the scope of auxiliaries. *Linguistics and Philosophy* 10:53–76.
- Winkler, Susanne. 2005. *Ellipsis and focus in Generative grammar*. Berlin: Walter de Gruyter.