A Labeling Approach to Constraints on Coordinate Structure Syntax

Ryan Walter Smith

Draft: 11/9/2015

Abstract

Cross-linguistically, coordinate structures are subject to two major constraints: the Coordinate Structure Constraint (CSC), which prevents both individual conjuncts and parts of conjuncts from moving, and the Coordination of Like Categories Constraint (CLC), which requires that conjuncts match in syntactic and semantic category. These constraints are typically considered to be causally independent of one another, and despite their cross-linguistic robustness, there is as yet no agreed upon theory of their derivation. In this paper, I propose a unified analysis of these constraints by means of the Labeling Algorithm of Chomsky (2013), in particular by treating coordinate structures as effectively XP-YP structures that can be labeled by the category features that they share. Given a dynamic definition of maximal projections, the immobility of individual conjuncts then follows from the restriction that only maximal projections may move.

1. Introduction

Coordinate structures are known to be subject to a number of syntactic constraints cross-linguistically. Particularly well-known among these is the Coordinate Structure Constraint (CSC), first formulated in Ross (1967) as follows.

1) Coordinate Structure Constraint

> In a coordinate structure, no conjunct may be moved, nor may any element contained in that conjunct be moved out of that conjunct.

The first part of the CSC, termed the Conjunct Constraint, prevents the generation of sentences like those in (2), in which one conjunct has moved out of the coordinate structure, and the second part, termed the Element Constraint, prevents ones like those in (3), in which an element of one conjunct has moved out of the coordinate structure (Grosu 1973; Zhang 2009).

(2) a. *What did you buy and bread?

b. *What did you buy bread and?

(3) a. *Who did you see and talk to Jeff?

b. *Who did you see Mary and talk to?

Another well-attested constraint on coordinate structures is the Coordination of Like Categories (CLC) constraint, due originally to Schachter (1977), who formulated it as in (4).

(4) Coordination of Like Categories Constraint

The constituents of a coordinate structure must belong to the same syntactic category and fulfill the same semantic function.

This constraint permits sentences like the one in (5a), but prohibits those like the one in (5b).

(5) a. I believe [the man and the woman] (DP & DP)

b. I believe [him and that he is nice] (DP & CP)

Although there are a number of counterexamples to the Element Constraint even in English (Lakoff 1986), as well as a few to the CLC (Borsley 1997), the Conjunct Constraint has withstood empirical scrutiny, and has even been referred to as “the most empirically robust island constraint known” (Postal 1998). However, compared to other

island constraints, there has been little work on unifying CSC effects with other cross-linguistically robust restrictions on extraction in natural language. Moreover, the CSC and the CLC are typically taken to be completely independent syntactic constraints, with the latter typically following from selectional restrictions placed on the coordinator (Johanneson 1998; Zhang 2009).

In this paper, I aim to unify Conjunct Constraint effects with CLC effects through the mechanism of labeling as developed by Chomsky (2013) and the restriction on moving only maximal projections utilized in the analysis of criterial freezing due to Rizzi (2014). The paper will proceed as follows. First, I will review Chomsky's (2013) approach to the Labeling Algorithm, as well as Rizzi's (2014) application of the labeling algorithm to the phenomenon of criterial freezing effects, in which an XP bearing a discourse relevant feature such as Q, Topic, or Focus, can no longer move once it reaches the edge of a head containing the same feature. I then turn to Chomsky's (2013) approach to coordination, and show that it does not make correct predictions about coordinate structure syntax. Finally, I attempt to derive both Conjunct Constraint and CLC effects from the labeling algorithm by treating coordinate structures as XP-YP structures that are labeled through feature sharing, the immobility of individual conjuncts being due to the aforementioned restriction on movement of non-maximal projections.

2. Labeling, maximality, and freezing

Chomsky (2013) proposes a conception of the language faculty in which Merge is totally free, applying blindly by simply combining two syntactic units together without labeling them. In order to provide labels for the structures built by Merge, which must occur before they are transferred to the interfaces, Chomsky proposes a Labeling

Algorithm, defined succinctly as in (6), with a notion of Closeness defined as in (7) (Rizzi 2014).

(6) Labeling Algorithm

The category created by Merge receives the label of the closest head

(7) Closeness

H_1 is the closest head in a structure α iff

i. α contains H_1, and

ii there is no H_2 such that α contains H_2 and H_2 c-commands H_1

With this definition of the labeling algorithm, it is straightforward to label asymmetric structures of the form [X YP], in which a head X is merged with a maximal projection YP; because X is a head, it c-commands the head of Y, and no other head in the structure c-commands it, the entire structure is labeled X. However, since Merge is completely free on this conception of the language faculty, there is nothing preventing the merger of two phrases or two heads, resulting in structures like [XP YP] and [X Y] respectively. Because both heads in both structures are equally close in terms of (7), the labeling algorithm cannot assign a label to the larger structure of which they are a part, resulting in a crash at Spell Out. Focusing on structures of the form [XP YP], there are two solutions to this problem. The first involves moving, or internally merging, one of either XP or YP in order to break the symmetry of the structure in a manner analogous to that in Moro (2000), resulting in a configuration as in (8) (Chomsky 2013, 12).

(8) XP...[<XP> YP]

Because Internal Merge renders the lower copy in a chain invisible for the purposes of labeling, the structure in (5) is no longer symmetric, and the structure can now receive the label of YP.

The second possibility for labeling XP-YP structures involves labeling the structure with an interpretable feature that both structures share (Chomsky 2013, 13). That is, given a structure [XP YP] and a feature F, if XP and YP share the feature F, then the label of the structure becomes F. Chomsky (2013) takes advantage of this possibility in deriving successive cyclic movement. For example, consider the sentence in (9).

(9) Which book did you say you read?

Movement of an interrogative phrase to the edge of a CP results in an XP-YP structure, as in (10).

(10) [[which$_Q$ book] [C you read <which book>]]

Here, neither phrase shares a feature, so the structure remains unlabelable. As such, movement occurs again in order to break the symmetry, this time raising to the edge of the next CP up.

(11) [[which$_Q$ book] [did$_Q$ you say [<which book> C you read <which book>]]]

Here, the phrase headed by 'did' and the interrogative DP share a feature, namely Q, thus allowing the structure to be labeled Q.

Rizzi (2014) further shows the utility of this possibility for labeling by applying it, with some additional constraints, to *criterial freezing*[1], in which phrases moved to certain

[1] The term 'criterial' here refers to a position with a particular scope-discourse property, such as the interrogative feature Q or the Focus feature (Chomsky 2001), called a *criterion* by Rizzi (2006; 2007). *Freezing* refers to a more general class of phenomena in which extraction out of moved constituents cannot take place. See Corver (2006) for an in-depth discussion of freezing effects and analyses thereof.

positions, termed *criterial positions*, can move no further. An example of this, again involving interrogatives, can be seen in the pair of sentences in (12).

(12) a. John wonders which book Bill read

b. *Which book does John wonder Bill read?

As one can see, 'which book' must stop at the edge of the embedded clause, and ungrammaticality results otherwise. To derive this result, Rizzi first formulates a constraint, as in (13), that states that only movement of maximal projections is permitted, with the notion of maximal projection determined dynamically via the definition in (14) from Chomsky (1995)[2].

(13) Maximality (Rizzi 2014, 20)

Only maximal objects with a given label can be moved

(14) Maximal projection

α is a maximal projection iff α is labeled, and the node immediately dominating α does not have the same label.

In the derivation of (12a), 'which book' moves to the edge of the embedded clause, resulting in an XP-YP structure. Labeling applies, resulting in the structure being labeled Q. However, given the definition of a maximal projection in (14), 'which book' no longer qualifies as a maximal projection, and therefore can move no further, thus deriving criterial freezing. Rizzi applies the same logic to restrictions on extraction of subjects, as well as to floating quantifiers, thus showing the wide range of effects for which the proposal accounts.

[2] See also Speas (1986) and Carnie (1995; 2000) for dynamic treatments of maximality/X-bar status.

There does, however, seem to be a slight technical problem with Rizzi's account, in that it seems to require that the interrogative C and the interrogative DP have exactly the same label. However, the labeling algorithm as it applies to XP-YP structures does not work this way; rather than labeling the node containing XP and YP with the *label* that XP and YP share, it labels the containing node with a (bundle of) *feature(s)* that they share[3]. As such, it would still be possible to label the result of merging an interrogative DP and an interrogative CP with the interrogative Q feature even if the DP and CP are not so labeled[4]. Because of this, if we wish to maintain Rizzi's explanation for criterial freezing, we will need to slightly redefine the definition of a maximal projection to make reference not to identity of the labels of two nodes but to the features contained in each node, as in (15) below.

(15) Maximal projection (revised)

> α is a maximal projection iff α is labeled, and the label of the node immediately dominating α is not a subset of the features of α.

On this revised definition, it no longer matters whether two phrases share the same label to be rendered non-maximal when the node immediately dominating them is labeled with one of their features; they are non-maximal because the label of that node contains a subset of the features of each node. This definition preserves Rizzi's analysis of criterial freezing while removing the need for the two nodes to have identical labels.

Having introduced the major concepts and tools to be used in the analysis, in the next section, I discuss and critique Chomsky's (2013) proposal for coordination, and then move onto my own analysis.

---

[3] I thank Heidi Harley (pc) for bringing this point to my attention.

[4] In fact, in Chomsky (2014), the label of an XP-YP structure is the *pair* of features that the two structures share, so it would never be exactly the same as the label of either phrase contained in it.

3. Coordination in Chomsky (2013)

Chomsky (2013) proposes that coordinate structures have the structure in (16).

(16) [Conj[XP YP]]

Here, coordinate structures are essentially small clauses containing two phrases, reminiscent of Moro's (2000) treatment of copular clauses and Ott's (2012) approach to split topics in German. Chomsky claims that conjunctions, much like acategorial roots, are too weak to act as the label of a phrase, and therefore (16) is not labelable as such. In order to label the structure, one of XP or YP must move, resulting in the configuration in (17).

(17) [XP Conj [<XP> YP]]

Here, XP has raised to the "specifier" of Conj, allowing the lower structure to receive the label of YP, and the entire coordinate structure to receive the label of XP.

Although this proposal reflects what Chomsky refers to as the "semantic symmetry" of the conjuncts in a coordinate structure (p. 41), and demonstrates another empirical application of symmetry breaking to labeling, it faces a few challenges. First of all, although the movement of XP does result in a symmetry breaking, allowing for the labeling of the lower structure, there is nothing that stops XP from raising even further beyond the coordinate structure, in violation of the Conjunct Constraint. Furthermore, strictly speaking, nothing precludes XP and YP from having completely different categories; as long as labeling can be achieved, the structure is well-formed, meaning that, barring additional assumptions about the interface with semantics, sentences with unlike category coordination as in (5b) above should be grammatical across the board.

It is apparent that Chomsky's proposal derives none of the syntactic properties of coordinate structures on its own. It seems, then, that more technology is necessary. In the next section, I will make use of the tools discussed in Section 2 to derive both Conjunct Constraint and Conjunction of Like Category constraint effects.

4. Deriving the CC and CLC from labeling

For my analysis, I will assume a more standard approach to the syntax of coordinate structures (Johannessen 1998, Munn 1993, Zhang 2009, etc.) in which the conjunction is structurally associated with one of the conjuncts, rather than Chomsky's small clause approach.

(18) [XP [Conj YP]]

I follow Chomsky in proposing that Conj is too weak to provide a label, though I go perhaps one step further in proposing that Conj is so weak that it is essentially ignored by the labeling algorithm. This means that the embedded structure containing Conj and YP can be trivially labeled with the label of YP, but that this results in an XP-YP structure, and thus is unlabelable as it stands. This is where the effect of feature agreement as used by Rizzi (2014) to derive freezing effects comes into play. Since the structure in (18) results in an XP-YP structure, if the two phrases share any set of interpretable features, including but not limited to category features, the structure can be labeled with that set of features. As such, for a sentence containing two interrogative DPs, the two phrases share the category D and the discourse feature Q, and thus can be labeled with either one or both of them. In so doing, however, because the structure has been labeled with a feature the two phrases share, given the definition of maximal projection provided in (15) above, they no longer count as maximal projections, and thus can no longer move independently

under Rizzi's notion of Maximality. However, because the entire coordinate structure is still a maximal projection, it *can* be moved, thus allowing for sentences like (19), with coordinated interrogative DPs.

(19) Who and what did you talk about last night?

Because the label of the entire structure is determined by the category that the conjuncts share, this derives the CLC, and because of the Maximality requirement on movement and the fact that, on this approach, the individual conjuncts no longer count as maximal projections, the Conjunct Constraint is derived.

A further prediction that this approach makes is that a coordinate structure consisting of an interrogative phrase and a non-interrogative phrase will not be able to undergo wh-movement. This is because the coordinated phrases do not share a Q feature, and thus Q cannot be the label of the coordinate structure, meaning that, if the structure is moved to the edge of a +Q CP, minimal search will not be able to identify a feature the two structures have in common. This prediction is correct; as (20) shows, wh-movement does not occur when an interrogative is coordinated with a non-interrogative phrase.

(20) a. *Who and John did you talk to last night?

b. *John and who did you talk to last night?

This fact also provides support for the redefinition of maximal projection provided in (15) over Rizzi's original definition. Assuming Rizzi's original definition in (14), it would in fact be possible for 'who' in the structure to be labeled differently, thus allowing it to move independently of 'John,' thus violating the CC. Given the revised definition, however, because the label of the entire coordinate structure contains a subset

of the features contained in ‘who,’ ‘who’ does not count as a maximal projection, and thus cannot move under Maximality.

5. Additional Issues

Although my analysis derives the effects of the CC and the CLC, it is important to mention some residual issues. For one, there do seem to be some counter-examples to the CLC, particularly in copular contexts like the example in (21) (Borsley 2005).

21) John is a lawyer and proud of his daughter

Here, there is an apparent violation of the CLC, in that the coordination appears to involve coordination of a DP and an AP. There are two possible ways to solve this problem for the proposal at hand. The first possibility is to propose that a PredP in fact dominates such structures, as proposed by Bowers (1993). Another is to have the two phrases share a Pred feature. In both cases, the labeling algorithm will be able to pick out the features that the two phrases share, and thus eliminate the apparent counter-example. More research needs to be done to determine whether such a solution is generally available in cases like this.

Another possible issue is presented by the language Gitksan, which seems to violate the CC, as (22) shows.

(22) Gwihl gubis Henry ganhl miyup? (Gitksan)

What eat Henry and rice

‘What did Henry eat and rice?’

Forbes (2013), who documents this case, argues that the apparent coordinator here is actually a comitative preposition. However, she notes, if this is the case, then Gitksan possesses no true coordinators. She also points to additional work that suggests that the

CC can be violated with disjunction as well, which provides a more compelling case for the argument that Gitksan is a true case of a CC-violating language. I will leave this issue to future research.

6. Conclusion

In this paper, I have argued that two major constraints on coordinate structure syntax, the Conjunct Constraint and the Coordination of Like Categories Constraint, follow from independently motivated considerations of labeling and movement of maximal projections as established in Chomsky (2013) and Rizzi (2014). If this approach is on the right track, then not only do two constraints previously considered unrelated to one another receive a unified explanation, but their explanation is further tied to other well-established syntactic constraints and effects. In particular, the Conjunct Constraint can be treated along the same lines as freezing effects and other island constraints, such as those treated in Bošković (2015), rather than simply being stipulated. Finally, the analysis supplied here provides further support for the viability of syntactic explanation in terms of labeling.

References

Borsley, R. D. (2005). Against ConjP. *Lingua*, *115*(4), 461-482.

Bošković, Ž. (2015). On the timing of labeling: deducing comp-trace effects, the subject condition, the adjunct condition, and tucking in from labeling. Ms. Retrieved from http://ling.auf.net/lingbuzz/002452.

Bowers, J. (1993). The syntax of predication. *Linguistic inquiry*, 591-656.

Carnie, A. (1995). *Non-verbal predication and head-movement* Doctoral dissertation, Massachusetts Institute of Technology.

Carnie, A. (2000). On the Definition of X0 and XP. *Syntax*, *3*(2), 59-106.

Chomsky, N. (2013). Problems of projection. *Lingua*, *130*, 33-49.

Chomsky, N. (2014). Problems of projection: Extensions. Ms.

Corver, N. (2006). Freezing effects. *The Blackwell companion to syntax*, *2*, 383-406.

Forbes, C. (2013). Coordination and number in the Gitksan nominal domain. MA thesis, University of Toronto.

Grosu, A. (1973). On the nonunitary nature of the coordinate structure constraint. *Linguistic Inquiry*, 88-92.

Johannessen, J. B. (1998). *Coordination*. Oxford University Press.

Lakoff, G. (1986). *Frame semantic control of the coordinate stucture constraint*. Cognitive Science Program, Institute of Cognitive Studies, University of California at Berkeley.

Moro, A. (2000). *Dynamic antisymmetry* (No. 38). Cambridge, MA: MIT Press.

Munn, A. B. (1993). *Topics in the syntax and semantics of coordinate structures.* Doctoral dissertation, The University of Maryland.

Ott, D. (2012). *Local instability: Split topicalization and quantifier float in German* (Vol. 544). Walter de Gruyter.

Postal, P. M. (1998). *Three investigations of extraction* (Vol. 29). MIT Press.

Rizzi, L. (2006). On the form of chains: Criterial positions and ECP effects. *CURRENT STUDIES IN LINGUISTICS SERIES*, *42*, 97.

Rizzi, L. (2007). On some properties of criterial freezing. *Studies in linguistics*, *1*, 145-158.

Rizzi, L. (2014). Cartography, criteria, and labeling. *Beyond Functional Sequence: The Cartography of Syntactic Structures,|*, *10*, 314.

Ross, J. R. (1967). *Constraints on variables in syntax.* Doctoral dissertation, Massachusetts Institute of Technology.

Schachter, P. (1977). Constraints on coordination. *Language*, 86-103.

Speas, M. J. (1986). *Adjunctions and projections in syntax*. Doctoral dissertation, Massachusetts Institute of Technology.

Zhang, N. N. (2009). *Coordination in syntax* (Vol. 123). Cambridge University Press.