

# A Revised Labeling Approach to Long-Distance Agreement

Shigeru MIYAGAWA<sup>†</sup>, Danfeng WU<sup>†</sup>, and Masatoshi KOIZUMI<sup>‡</sup>

<sup>†</sup>MIT Department of Linguistics and Philosophy, Massachusetts Institute of Technology  
77 Massachusetts Avenue, Room 32-D808, Cambridge, MA 02139, USA

<sup>‡</sup> Department of Linguistics, Tohoku University  
3-12-4 Kawauchi, Aoba-ku, Sendai, 980-8576 Japan

E-mail: <sup>†</sup>{miyagawa,dfwu}@mit.edu, <sup>‡</sup>koizumi@tohoku.ac.jp

**Abstract** The present study points out a critical issue of the labeling theory (Chomsky 2013) where it undergenerates. It fails to predict the grammaticality of long-distance agreement, a phenomenon that has been attested in many languages such as English, Icelandic and Kaqchikel. A revision of the theory is proposed, which successfully accounts for the data that the original theory failed to explain. Furthermore, the revision parallels a previous proposal by Saito (2016) and suggests that agreement and case are two different ways to resolve the problem of projection.

**Key words** Labeling, Problems of Projection,  $\varphi$ -agreement, projection licenser, expletive *there*, locative inversion, long-distance agreement, case, syntax

## 1. Introduction

Chomsky's (2013) labeling theory has provided a novel and uniform account of why syntactic movement and agreement exist in languages. According to this theory, movement and agreement exist to resolve Problems of Projection (POP) that arise from merging two heads or two phrases. In particular, when a noun phrase and a TP are merged, they agree with each other, so that the common agreement feature can supply its label to the merged phrase, and solve the POP. We note a critical problem of this theory and present attested data across three languages that it fails to generate. These data on long-distance agreement suggest that agreement does not necessarily have to occur between the pair of phrases that cause the POP. Rather, as long as one of the phrases has agreement morphology, that suffices as a solution to the POP. So we revise the theory and argue that in a POP situation, the constituent marked by agreement morphology is licensed to project that agreement feature as the label of its mother node, and that agreement morphology does not have to result from agreement with its sister. By relaxing the criterion of what form of agreement counts as a solution to POP, we can now successfully account for the data that the original theory failed to explain. Furthermore, our revision is not ad-hoc, but in fact parallels Saito's (2016) argument that case morphology is also a solution to POP. If this is true, it suggests that agreement and case morphology are reflexes of each other and two different ways to resolve a POP.

## 2. Original Theory of Labeling

Chomsky's (2013) general approach lays out a theory of how structures are built in narrow syntax, and particularly, how

each structure gets its label. The theory not only provides the necessary theoretical tools to build the basic structures of syntax, but it also makes predictions about when movement must occur — movement occurs to address Problems of Projection (POP) — and it also identifies the role of agreement in language.

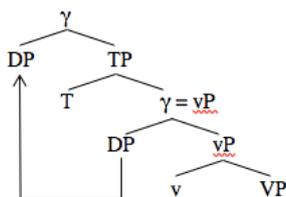
To begin, we will review details of the theory of labeling. According to this theory, Merge applies to two objects  $\alpha$  and  $\beta$ , and forms a new object  $\gamma = \{\alpha, \beta\}$ .  $\alpha$  and  $\beta$  are existing objects available to this operation, while  $\gamma$  is a new construct that must be given a label. The nature of human language is that each phrase is headed, meaning that either  $\alpha$  or  $\beta$  provides the label for  $\gamma$ ; the question is, which is chosen,  $\alpha$  or  $\beta$ , to project itself to give the label for  $\gamma$ ? There are three possible configurations that result from Merge.

- (1) a.  $\gamma = \{H, \alpha P\}$
- b.  $\gamma = \{\alpha P, \beta P\}$     Problem of projection
- c.  $\gamma = \{H_1, H_2\}$     Problem of projection

In (1a) Merge has combined a head (H) and an XP; in (1b) two XPs have been combined; and in (1c) two heads have been combined. In order for  $\gamma$  to find its label, it undertakes search within its local domain, which is the binary structure that it directly dominates, and picks one of the pairs to project to furnish the label for  $\gamma$ . This search, called a labeling algorithm, must result in an unambiguous choice. In (1a) the search does in fact result in a unique choice, since the members of the pair,  $\{H, \alpha P\}$ , are distinct, and H is the closest target of the search. In contrast, neither of the remaining two results in a unique search because the two members are XPs ((1b)) or they are both heads ((1c)). Unless they are altered in some fashion, neither of these merged structures would be labeled, thus it will create a POP and fail as a structure in language.

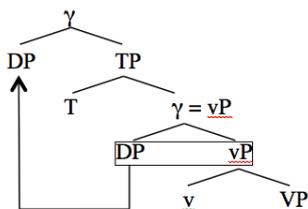
Chomsky identifies two additional operations that can apply to ambiguous merged structures, movement, which is a form of Merge, and agreement. We illustrate both in the derivation of the external argument. There are two POPs that arise in this structure, both of the form in (1b) above, in which Merge has paired two XPs. The first of these is the pair  $\{DP, vP\}$ .

(2)



A way to provide a unique label for  $\gamma$  is to have one of the members move out of the structure, leaving just one member of the pair for  $\gamma$ . This is what we see in (3); having the DP vacate its original position leaves  $vP$  as the sole member, allowing it to project and giving  $\gamma$  an appropriate label.

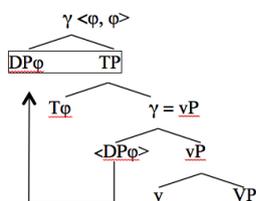
(3)



While movement of the external argument allows labeling of the lower  $\gamma$ , it leads to a second POP, at the landing site of

this movement, {DP, TP}. To avoid a POP, there is an additional element in this pair that makes labeling of  $\gamma$  possible, namely, agreement. As a result of agreement between DP and T, the two members of the pair, {DP, TP}, despite being distinct, nevertheless share the same feature  $\varphi$ . Searching {DP, TP}, then, the labeling algorithm finds the same prominent element  $\varphi$  in both terms, and can take that to be the label of  $\gamma$  (Chomsky 2013).

(4)



This theory successfully accounts for the external argument's movement to Spec, TP and its  $\varphi$ -agreement with T in English. Consider the following simple example:

(5) [ $_{\varphi P}$  John<sub>i</sub> will [<sub>vP</sub> t<sub>i</sub> buy books]].

The subject *John* moves out of Spec, vP so that the mother node of its base position can be labeled as vP. But when the subject moves to form a sister of TP, we run into a POP again. Here the subject DP agrees in  $\varphi$ -feature with T, so that their mother node can be labeled as a  $\varphi P$ .

## 2.1. A Problem

There is a problem with this approach. While it accounts for most cases of  $\varphi$ -agreement in English that involve local agreement of the subject and its sister, English has been argued to have long-distance agreement as well. Long-distance agreement is agreement not between sisters, but between two elements that are farther away.

English sentences involving the expletive *there* are claimed to involve long-distance agreement. *There* is generally assumed to be semantically vacuous and inserted only to fill Spec, TP (Chomsky 2000):

(6) There are many people in the room.

When the expletive phrase *there* is merged with TP and generates {*there*, TP}, a POP arises. The usual solution would be to have *there* and T agree in  $\varphi$ -feature, so that the merged node can bear the label of that  $\varphi$ . However, *there* is deficient in  $\varphi$ -features, according to Chomsky 2000. Having failed to agree with its specifier *there*, T then probes down and agrees with the associate DP instead. This is supported by the fact that the  $\varphi$ -features of T covary with those of the associate DP:

(7) a. There is a person in the room.

b. There are many people in the room.

The word order indicates that the associate DP follows finite T, so it can't be T's specifier. The definiteness restriction that the associate must be indefinite and nonspecific also suggests that it stays low (Milsark 1974). Assuming the low position of the associate DP, this suggests that agreement can occur between T and the associate phrase lower down instead of specifier of T.

The labeling approach fails to predict such long-distance agreement. Under this approach, the only solution to POP by agreement has to be agreement between XP and YP from which the POP arises, i.e. local agreement between sisters. It predicts long-distance agreement to be impossible and undergenerates.

Long-distance agreement is seen not only in *there* expletive sentences, but also locative inversion constructions in





While our paper is mostly concerned with  $\phi$ -agreement, the same approach may be extended to other forms of agreement such as *wh*-agreement. For instance, a *wh*-phrase has been shown to move successive-cyclically through intermediate phase edges before landing in Spec, CP of the interrogative C probe (c.f. van Urk and Richards 2013, for example). This is predicted by the labeling theory, as a POP arises when the *wh*-phrase is merged with an intermediate CP (*{wh*-phrase, CP}), and the solution is to move the *wh*-phrase out. It stops moving when the *wh*-word agrees with its sister, the interrogative CP, and the shared *wh*-feature projects.

Given the discussion of long-distance  $\phi$ -agreement in this paper, a natural question is whether there is long-distance *wh*-agreement in any natural language. Recall our definition of long-distance agreement as agreement that occurs not between sisters. We believe so. One example is English subject *wh*-questions, which has been shown to not involve any movement of the *wh*-subject:

(13) Who left?

With the *wh*-subject remaining in Spec, TP, the specifier-less interrogative C agrees with the *wh*-subject lower down. Unfortunately, since the CP has no specifier, we also don't run into a POP in such a situation, so we wouldn't be able to tease apart the original labeling theory and our revised version in the case of *wh*-agreement. Still, future research awaits to extend our modified labeling theory to other aspects of the grammar.

## 5. References

- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michaels and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130: 33–49.
- Chomsky, Noam. 2015. Problems of projection: Extensions. *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*, ed. by Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 3-16. Amsterdam/Philadelphia: John Benjamins.
- Imanishi, Yusuke. 2014. *Default ergative*. Ph.D. dissertation, MIT.
- Mateo Pedro, Pedro. 2009. Nominalization in Q'anjob'al (Maya). *Kansas Working Papers in Linguistics*, 46–63.
- Milsark, Gary. 1974. *Existential sentences in English*. Ph.D. dissertation, MIT.
- Saito, Mamoru. 2016. (A) Case for labeling: Labeling in languages without phi-feature agreement. *The Linguistic Review* 33:129-175.
- Thráinsson, Höskuldur. 2007. *The Syntax of Icelandic*. Cambridge: Cambridge University Press.
- van Urk, Coppe, and Norvin Richards. 2015. Two components of long-distance extraction: Successive cyclicity in Dinka. *Linguistic Inquiry* 46:113–155.