Control-Forming Domains are Not Only Phases: Evidence for Probe Horizons
Jurij Božič, McGill university

I. Introduction. A general assumption is that syntactic Phases (Chomsky 2001) delimit probing domains. Keine (2016, to appear) proposes that probes can have search-restrictions, which he terms ‘Horizons’, introducing a new type of locality in addition to phases. This paper examines the domains of Control-formation in Slovenian (South Slavic) and determines that Keine’s system is needed: Control must be constrained with a combination of phasal boundaries and probe Horizons. We will show that, with TP-embeddings, Subject Control is possible, but that only Object Control is possible with CP-embeddings. Since the embedding is a weak phase in both cases, it follows that the X0 responsible for assigning the external θ-role must be unable to probe across C0. In other words, C0 must be a Horizon for this probe.

II. Basic Data. Slovenian has two types of Control constructions: one type embeds an infinitival TP, and the other an infinitival CP when the matrix verb is a perception predicate. To facilitate the discussion of the latter, we must first consider cases where perception verbs embed regular, finite CPs. In those cases, an extra DP object can occur in the matrix clause:

1. Otroci vidjo/slišjo, da Janez videl trava.
   The children see/hear that John mows lawn
   ‘The children see/hear that John is mowing the lawn.’
   (No extra internal θ)

2. Otroci vidjo/slišjo Janeza, da pro, kosi trava.
   The children see/hear John that mows lawn
   ‘The children see/hear John mowing the lawn.’
   (Extra internal θ!)

(2) is a type of prolepsis with a base-generated DP-object, involving an embedded pro.1 The embedded clause is fully finite with no tense or φ-deficiency. Also, the matrix DP ‘John’ must be the ‘object of perception’ in (2), but it is not in (1). This implies that perception verbs bear an optional internal θ-role.

Perception verbs can also embed infinitival CPs, which are instances of Object Control, as in (3):

3. Otroci vidjo Janeza [i CP PRO, kosi trava].
   The children see John mow-INF lawn
   ‘The children see John mowing the lawn.’

Why should (3) be analyzed as Control and not as ECM/SUBJ-to-OBJ raising? Since Slovenian is a negative concord language, negative words such as nobody require Neg0 to be a clause-mate. If nobody is a Controller in typical TP-embedding constructions, Neg0 cannot occur low (4), and exactly the same is true of the perceptual Control construction (5), which suggests nobody is θ-marked in the matrix clause:

4. *Noben, poskuša [TP PRO, kosi trava].
   Nobody tries not mow-INF lawn
   ‘Nobody is trying to mow the lawn.’

5. *Otroci vidjo nobenga, [i CP PRO, kosi trava].
   The children see nobody not mow-INF lawn
   ‘The children don’t see anyone mowing the lawn.’

Note that Neg0 can occur low, with a different scope reading, but only when the Controller is not a negative word. Den Dikken (2017) makes similar observations about supposed cases of Hyper-Raising in Finnish and Hungarian, and notes that we do not expect the pattern in (4)–(5) to obtain with raising. In fact, subject scrambling and A′-movement reveal no such restriction in Slovenian. This implies matrix θ-marking of the object in (3), and hence Control, since it must be a clause-mate with Neg0.

We must also justify the TP/CP-difference in embedding in (4)–(5). A fairly standard diagnostic for the presence of a CP in Slovenian is clitic climbing (Marušič 2005). While clitic climbing is possible if not obligatory with regular Control (6), it is not possible with perceptual Control (7):

6. Otroci so jį jį probal [TP PRO, t darilu].
   The children tried to give her a present.
   \[ \begin{array}{c} jį \end{array} \]

7. Otroci so (jį jį) videl Janeza, [i CP PRO, t darilu].
   The children saw John give her a present.
   \[ \begin{array}{c} jį \end{array} \]

1This must be the case. I will present data in which a φ-mismatch between the matrix object and the embedded pro may occur; and the matrix object – if it is a coordination – can also serve as a ‘split antecedent’ for the embedded subject and object.
III. The Problem. As noted above, the perceptual verbs participating in the (1)–(2) constructions host an optional internal \( \theta \)-role that gets assigned to the matrix DP-object when present. We expect this \( \theta \)-optionality to work in the same way when perceptual Control is formed. However, derivations without the additional internal \( \theta \)-role always crash when embedding CP\(_{\text{INF}}\):

\[
(8) \quad \text{*Otroci, vidjo } [\text{CP } \text{PRO}, \text{kosit travo}].
\]

\begin{itemize}
\item children see
\item mow-INF lawn
\end{itemize}

\textit{int.} ‘The children see (themselves) mowing the lawn.’

In the absence of the extra internal \( \theta \)-role, we expect Subject Control to occur, as in (8), which does for instance occur in English: cf. \textit{John asked} Mary, [\text{PRO, to leave}] vs. \textit{John, asked} [\text{PRO, to leave}]. But this is not the case in Slovenian, even though subject Control is possible with TP-embeddings, cf. (6). One might consider a \textit{lexical solution} to this problem: perhaps CP\(_{\text{INF}}\) can only be c-selected by a \( v^0 \) that obligatorily (and not optionally) specifies an extra internal \( \theta \)-role. This hypothesis does predict (8). However, every perceptual verb that participates in (1)–(2) (e.g. \textit{videt} ‘see’, \textit{slisat} ‘hear’, \textit{opazt} ‘notice’, \textit{zaznat} ‘detect’, etc.) can also form perceptual Control. The selection analysis would predict that this generalization is a \textit{lexical accident}, since some verbs could easily select for CP\(_{\text{INF}}\) while keeping their internal \( \theta \)-role optional. If we wish to derive this systematic behaviour of perception verbs, we need to find a different solution.

IV. Role of Phases. Could the strong/weak Phase (Chomsky 2001) distinction be invoked to solve the discussed problem? It cannot. The presence of a \textit{weak phase boundary} between the matrix and embedded clauses is a pre-condition for Control-formation to begin with (Boeckx et al. 2010; Gallego 2010), and we present extra evidence to support this. The embedded CP\(_{\text{INF}}\) is indeed a \textit{weak phase}: it allows long-distance NPI-licensing (9), it is tense/\( \phi \)-deficient (10), it licenses scope inversion with a matrix Q (11):

\[
(9) \quad \text{Otroci niso videl Janeza} \_ \text{ACC } [\text{CP s prstom migt, da bi pomagov}].
\]

\begin{itemize}
\item children AUX-not see
\item John with finger
\item move-INF that would help
\end{itemize}

‘The children didn’t see John lift a finger to help.’

\[
(10) \quad \text{Otroci so \textit{vjeri} videl Janeza} \_ \text{ACC } [\text{CP jutri kosit travo}].
\]

\begin{itemize}
\item children AUX\(_{\text{PAST}}\) yesterday see
\item John with finger
\item mow-INF lawn
\end{itemize}

\textit{int.} ‘The children yesterday saw John mowing the lawn tomorrow.’

\[
(11) \quad \text{Otroci \textit{vidjo} en ga policaja} \_ \text{ACC } [\text{CP stat na vsakmu kriziscu}].
\]

\begin{itemize}
\item children see
\item one policeman
\item stand-INF on every crossroads
\end{itemize}

‘The children see one policeman standing at every crossroads.’

(9) and (11) are not possible across strong phases in Slovenian, while (10) is grammatical \textit{only} across strong phases. CP\(_{\text{INF}}\) is inevitably \textit{weak}, which means that this alone cannot derive (8).

V. Horizons. Keine (2016, \textit{to appear}) proposes that \textit{probes} may have restrictions on ‘search’, imposing a new constraint on probing, in addition to strong phasal spell-out. I propose that this is what we require in order to derive (8). Specifically, Voice\(^0\), which is responsible for assigning the external \( \theta \)-role, always \textit{terminates} probing when it encounters a C-feature; in Keine’s terms, C is a Horizon for Voice\(^0\) (‘Voice\(^0\)-C’):

\[
(12) \quad \theta\text{-assigning heads in Slovenian:}
\]

\begin{itemize}
\item Voice\(^0\) \(\vdash\) C, \(v^0\) \(\vdash\) \(\varnothing\)
\end{itemize}

This means that Voice\(^0\) will never be able to probe past C\(^0\), but \(v^0\) will, since it has no Horizon restriction. This directly derives the distinction between the two Control types in Slovenian:

\[
(13) \quad [TP \ldots \text{Voice}^0_u \theta \ldots v^0_u \theta \ldots [TP \text{ DP} \ldots ]] \quad [TP \ldots \text{Voice}^0_u \theta \ldots v^0_u \theta \ldots [CP \text{ DP} \ldots ]
\]

The upshot of this is that Subject Control is possible with CP-embeddings, but only Object Control is possible with CP-embeddings. It should be noted that it does not matter which theory of Control we subscribe to: under the Movement Theory of Control (Boeckx et al. 2010), the \(u\theta\)-probe on Voice\(^0\) will not be able to probe past C\(^0\) to discover the DP in the embedded SpecTP, while under Landau’s (2004) approach, Voice\(^0\) will not be able to probe past C\(^0\) to license the \([-R(efer)]\) feature on the embedded PRO. Under both views, the derivation without an internal \( \theta \)-role will crash.

Conclusion. In sum, Horizons are required to constrain probing in addition to the locality domains set by strong phasal spell-out. This distinction gives a principled account of why CP-embedding Control constructions do not permit forming Subject Control in Slovenian, but the TP-embedding ones do. Voice\(^0\) must terminate probing as soon as it encounters a C-feature, while \(v^0\) has no such restriction.