Probe Horizons in Control Formation

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: Voice^0 may license PRO. But *only Object Control* is possible with CP-embeddings, where only v^0 can license PRO:

(1) $[_{TP}$ . . .$\text{Voice}^0{}_{u\theta}$. . . $\text{v}^0{}_{u\theta}$ . . . $[_{TP}$ WEAK PRO ....]] $[_{TP}$ . . .$\text{Voice}^0{}_{u\theta}$. . . $\text{v}^0{}_{u\theta}$ . . . $[_{CP}$ WEAK PRO ...]]

We will show that the embedding must be a *weak phase* in both cases and therefore cannot account for the distinction. If phases cannot account for the absence of SUBJ-Control with CP-embeddings, it then follows that the X^0 responsible for assigning the external θ-role (viz. Voice^0) must be unable to probe across C^0. In other words, C^0 must be a Horizon for Voice^0 in Slovenian.

II. Basic Data. Slovenian has two types of Control constructions: one type embeds an infinitival **TP**, and the other an infinitival **CP** if 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:

(2) Otroci vidjo/slišjo, [da Janez$_{NOM}$ kosi travo.]
children see/hear that John mows:3SG lawn
'The children see/hear that John is mowing the lawn.' *No extra matrix internal θ*

(3) Otroci vidjo/slišjo **Janeza**$_{ACC i}$, [da pro_i kosi travo.]
children see/hear John that mows:3SG lawn
'The children see/hear John mowing the lawn.' *Extra matrix internal θ!*

(3) 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 (3) but is not in (2). Perception verbs bear an *optional internal* θ-role.

The same set of perception verbs can also embed **infinitival** CPs, which are instances of Object Control, as shown in (4):

(4) Otroci vidjo Janeza$_{ACC i}$ $[_{CP}$ PRO_i kosit travo].
children see John mow-INF lawn
'The children see John mowing the lawn.'

Why should (4) be analyzed as Control and not as ECM/SUBJ-to-OBJ raising? Firstly, the matrix OBJ must be the direct 'object of perception', which implies matrix θ-marking. Secondly, an embedded *idiom* receives an obligatory *compositional* reading when its subject is replaced by PRO:

(5) #General je vidu **vse karte**$_{ACC,i}$ $[_{CP}$ PRO_i **bit na mizi**].
general AUX saw all cards be-INF on table
int. 'The general saw that the situation became clear.'

We must also justify the presence of the CP-layer 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_i so $\boxed{\text{ji}_j}$ probal $[_{TP}$ PRO_i dat t_j darilo].
children AUX $\text{her}_{CL:DAT}$ try give-INF present
'The children tried to give her a present.'

(7) Otroci so (*ji_j) videl Janeza$_{ACC i}$ $[_{CP}$ $\boxed{\text{ji}_j}$ PRO_i dat t_j darilo].
children AUX $\text{her}_{CL:DAT}$ see John $\text{her}_{CL:DAT}$ give-INF present
'The children saw John give her a present.'

[1] As corroborating evidence, 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 (2)–(3) constructions with a CP_{+FIN} host an optional internal θ-role that gets assigned to the matrix DP-object when present. We expect this θ-optionality to work in the same way when perceptual Control is formed. However, derivations without the additional internal θ-role always crash when embedding CP_{-FIN}:

(8) *Otroci$_i$ vidjo [$_{CP}$ PRO$_i$ kosit travo].
children see mow-INF lawn
int.'The children see (themselves) mowing the lawn.'

In the absence of the extra internal θ-role, we expect Subject Control to occur, as in (8), which does for instance occur in English: cf. *John asked$_\theta$ Mary$_i$ [PRO$_i$ to leave]* vs. *John$_i$ asked [PRO$_i$ 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 *lexical solution* to this problem: perhaps CP_{-FIN} can only be c-selected by a v^0 that obligatorily (and not optionally) specifies an extra internal θ-role. This hypothesis does predict (8). However, every perceptual verb that participates in the prolepsis alternation (2)–(3) (*videt* 'see', *slišat* 'hear', *opazt* 'notice', *zaznat* 'detect', etc.) can also form OBJ-Control but not SUBJ-Control. The selection analysis predicts that this generalization is a *lexical accident*, since some verbs could easily select for CP_{-FIN} while keeping their internal θ-role optional. If we wish to derive this systematic behaviour of perception verbs, we need a different solution. One possibility is that (8) is due to some semantic restriction on perceptual verbs – perhaps they cannot form 'reflexive' readings. However, that is not the case, as an overt OBJ DP hosting a reflexive in (8) (and binding PRO) renders the example grammatical. The restriction hence seems to be syntactic.

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 *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_{-FIN} is indeed a *weak phase*: it allows long-distance NPI-licensing, as shown in (9).

(9) Otroci niso videl Janeza$_{ACC}$ [$_{CP}$ **s prstom mignt**, da bi pomagov].
children AUX-not see John with finger move-INF that would help
'The children didn't see John lift a finger to help.' [NPI IS BOLDED]

NPIs are not licensed across strong (finite) phasal domains in Slovenian: Neg^0 and the NPI must be clause-mates, but this is not the case in (9). However, could it be that the CP is spelt out as soon as vP is constructed, thus allowing v^0 to probe into the embedded clause, but not $Voice^0$? Such an account is not compatible with (9) either, as matrix Neg^0 needs to be generated above VoiceP:

(10) Otroci$_i$ ne jejo **vsi**$_i$ zelenjave.
children:PL.M not eat all:PL.M vegetable
'Not all children eat vegetables.'

(10) shows a floating quantifier scoping over the subject, which is usually taken as evidence for the VP(VoiceP)-internal subject hypothesis, and this has also be argued for Slovenian (Ilc & Milojević-Sheppard 2001). Since Neg^0 is fairly high in the structure, this means that there can be no (strong) phasal boundary between VoiceP and vP. Additional evidence for CP_{WEAK} in (9) includes an embedded tense-deficiency and scopal interactions between matrix and embedded Qs.

V. Horizons. Keine (2016, *to appear*) proposes that *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 θ-role, always *terminates* probing when it encounters a C-feature:

(11) Θ-assigning heads in Slovenian:
$Voice^0 \Vdash C$, $v^0 \Vdash \varnothing$ $\rightsquigarrow$ *C is a Horizon for Voice^0, but v^0 has no Horizon*

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 shown in (1). 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.

A combination of PHASAL DOMAINS and HORIZONS is needed to derive the Slovenian data.

Selected references: Keine, S. (2016). *Probes and their Horizons.* PhD Thesis, Umass, AM. • Keine, S. (*to appear*). Selective Opacity. *Linguistic Inquiry.*