

Eliminating Strong/Weak Parameter on T*

Nobu Goto

This squib consists of two parts: the first part (section 1) identifies potential flaws in the labeling theory proposed in Chomsky (2015); and the second part (section 2) suggests a system that eliminates the flaws.

1 EPP, ECP, and Labeling

Chomsky (2015) drops the assumption that EPP (Extended Projection Principle) is a universal property of language, i.e., the requirement that clauses must have subject, or more precisely, that SPEC-T be filled with a(n) (overt) subject (cf. Chomsky 1981). To ensure this principle, it has been assumed that in non-null subject languages (henceforth, non-NSLs) like English, EPP is satisfied by an overt subject, whereas in null subject languages (henceforth, NSLs) like Italian, EPP can be satisfied by a null pro expletive subject. However, dropping the long-standing assumption, and adopting Rizzi's (1982) idea that the parametric difference is attributed to richness of agreement, Chomsky (2015) proposes to parameterize the strength of T regarding a labeling algorithm LA (cf. Chomsky 2013): T in English-type non-NSL, with weak agreement, is too "weak" to serve as a label, hence there must be an overt subject in SPEC-T to label the SPEC-TP construction as $\langle \varphi, \varphi \rangle$ by the agreeing features, whereas T in Italian-type NSLs, with strong agreement, is "strong" enough to serve as a label, hence there is no need to have an overt subject in SPEC-T to label the SPEC-TP construction. Under the present theory, therefore, whether SPEC-T needs to be filled with an overt subject is attributed to the strength of T: if T is weak, the SPEC-T

requires an overt subject, but if T is strong, it doesn't. It follows that EPP is no longer a universal property of language; SPEC-T does not have to be filled neither with an overt subject nor a null pro expletive subject if T is strong.

The strong/weak parameter on T captures the long-standing observation that further *wh*-extraction from SPEC-T is impossible in non-NSLs like English, but possible in NSLs like Italian (see Perlmutter 1971). The classical explanation of the difference was attributed to ECP (Empty Category Principle), i.e., the requirement that traces must be properly governed (see Chomsky 1981): in non-NSLs like English, subject traces fail to be properly governed, so that further *wh*-extraction is impossible, whereas in NSLs like Italian, subject traces are properly governed by the null pro expletive subject, so that further *wh*-extraction is possible. Dropping this ECP-based explanation in the GB era, Chomsky (2015) offers a unified account of the parametric difference in ECP under the strong/weak parameter on T: in non-NSLs like English, T is too weak to serve as a label, hence the *wh*-subject must stay in SPEC-T to keep the SPEC-TP construction labeled as $\langle \varphi, \varphi \rangle$, but in NSLs like Italian, T is strong enough to be able to label the SPEC-TP construction, hence the *wh*-subject doesn't have to stay in SPEC-T and can raise further.

Given that memory is phase-level (cf. Chomsky 2000, 2001, 2004, 2005, 2007, 2008), Chomsky also provides an argument for why *wh*-extraction from SPEC-T is allowed in English when the phase head C undergoes deletion (see Perlmutter 1971): if the SPEC-TP construction has already been labeled as $\langle \varphi, \varphi \rangle$ at a phase level, then the information is available and stays labeled at the phase level, so that even if the *wh*-subject raise further to SPEC-C, no labeling failure occurs. In this particular case, he notes that phasehood of C is inherited by T along with φ -feature inheritance (cf.

Richards 2007) and accordingly activated on T by C-deletion. Thus, T, not C, functions as a phase head here, and as a result, unlike the usual way, the interior of T is identified as a transferred domain, observing the Phase Impenetrability Condition (PIC) (see Chomsky 2000, 2001, 2004). For this reason, when the phase head C undergoes deletion, a PIC violation by further *wh*-extraction from SPEC-T can be circumvented. When C is not deleted, on the other hand, C retains phasehood, so that the interior of C, TP, informally speaking, is transferred as usual, along with a *wh*-subject in SPEC-T, and therefore yielding the *that*-trace effect. In this way, EPP, ECP, and their parametric differences can be unified under the labeling theory once we adopt the strong/weak parameter on T.

Although these arguments are suggestive, unfortunately, there is a major hurdle to be overcome before we accept the parameter as a principled property of the language system. The Minimalist Program (MP) for the past 20 years tells us that the notion of “strength” runs the risk of stipulation (see, among others, Broekhuis 2008). In fact, Chomsky (1998: 127) himself has suggested the following research agenda for the MP, pointing out that such a notion should be eliminated in favor of the optimal design of language:

(1) “optimal design should eliminate such strange and difficult properties as strength”

In terms of this, the strong/weak parameter on T should be eliminated.

In relation to this, one of the conceptual problems of the parameter in question is why it is only T that undergoes such parameterization. In contrast to T, adopting the idea that roots R are unspecified as to category in the lexicon and their category

derives from merger with a functional element (see, among others, Borer 2005a, b, 2013), Chomsky claims that “R is universally too weak to label.” This asymmetry between T and R with respect to the strength of labeling itself may be an interesting question, but if it is dispensable, it should be abandoned, in favor of (1).

The parameterization of Italian T as “strong” also poses a fundamental question for Chomsky’s (2013) elegant argument for why the external argument (SUBJ) must raise from the v*P-internal position: the syntactic object $SO = \{SUBJ, v^*P\}$, neither a head, is invisible to LA and cannot be labeled, and hence the subject must raise (Chomsky 2013: 44; cf. Chomsky 2008, fn. 34): by raising the subject to SPEC-T, what is visible to LA becomes v* alone, as in $\{t, v^*P\}$, *t* the lower copy of the subject, so that the SO is labeled as v*, as required. The reason for the invisibility of the lower copy of the subject to LA, Chomsky argues, is that it is part of a discontinuous element, which is created when some element (here, SUBJ) in a domain D1 (here, SPEC-v*) goes to a different domain D2 (here, SPEC-T).

This argument is also suggestive, but once we accept the strong/weak parameter on T, unfortunately, the impact of the argument is also undermined. That is, if Italian T alone can label the SPEC-TP construction due to its strength, then it is not necessary for the subject to raise to SPEC-T, allowed to stay in situ. The question is, then, how the predicate-internal subject construction is labeled in such a situation. Chomsky (2015) recognizes this point as a potential problem of the labeling analysis of Italian EPP, putting aside about just where the subject and v*P are when the subject does not raise. However, according to Belletti’s (2001, 2004) series of studies of the relevant structure (“subject inversion structures,” in her terms), she concludes that when the subject does not raise, either the subject remains in its original v*P-

internal position or it raises to a very low position in the clause structure, lower than the lowest SPEC hosting an adverb (cf. “Cinque’s hierarchy”). If this is the case, it follows that the subject that does not undergo raising to SPEC-T amounts to stay around in its base position, i.e., the v*P-internal position; hence still visible to LA. The construction in question is therefore expected to result in a labeling failure, contrary to fact: as we have seen above, for it to be labeled as v*, the subject must raise (at best up) to SPEC-T so that the lower copy of the subject is invisible to LA. Then, the question remains how the predicate-internal subject construction is labeled when the subject does not raise to SPEC-T, or more relevantly, when T is strong.

What remains, then, is the impression that the strong/weak parameter on T may be a stipulation, and so should probably be eliminated from the theory if possible. Having said that, Chomsky’s fine unification of EPP and ECP in terms of labeling should be defended, in the light of recent developments of the labeling theory (for arguments for Chomsky’s labeling theory, see, among others, Epstein, Kitahara, and Seely 2014, 2015, Goto 2013, to appear). Especially, given that everything unfortunate starts from the radical assumption that EPP is not universal, it is tempting to ask whether it is indeed on the right track and the traditional assumption that EPP is a universal property of language cannot be maintained in a way that fits well into the present theory. Considering that EPP has received a significant amount of attention in the literature to date, there should be a breakthrough solution. If there is an independently motivated way to satisfy EPP, or to ensure $\langle \varphi, \varphi \rangle$ labeling without recourse to the strength of T, then it would be worth considering the possibility of the theory that incorporates the way.

In what follows, I explore such a possibility, attempting to learn new things

from the past, with appropriate updating of the relevant assumptions. My claim is, thus, as has traditionally been assumed (see Chomsky 2008), that EPP is universal, holding for both phases, CP and v*P, and for all languages, and, consequently, the strong/weak parameter on T can be, hence should be, eliminated from the theory. More specifically, I argue that it is indeed possible to eliminate the stipulation if we adopt and elaborate on the widely accepted view that EPP can be satisfied by V-raising (for a theoretical precursor, see Alexiadou and Anagnostopoulou 1998, and for further developments in interesting ways, see, among others, Biberauer and Roberts 2008, Richards and Biberauer 2005).¹ Then, in section 2, I adopt the idea of EPP-satisfaction by V-raising, and elaborate on it, particularly in terms of the simplest Merge-based system, advocated by Chomsky (2013, 2015), under which an SO is constructed by simplest Merge, defined as follows:

$$(2) \text{ Merge}(\alpha, \beta) \rightarrow \{\alpha, \beta\}$$

Just as an instance of (2), I argue that Merge of a head to SPEC (henceforth, Merge H-to-S) is possible, and demonstrate that the labeling theory that allows the possibility can not only maintain the spirit of Chomsky's labeling analysis of EPP and ECP but also bring significant consequences for the halting problem (cf. Rizzi and Shlonsky 2007) and for the feature inheritance mechanism (cf. Richards 2007).

2 Labeling by Merge H-to-S and Inheritance for Labeling

Alexiadou and Anagnostopoulou (1998) offer an interesting approach to EPP-satisfaction: EPP can be satisfied by either DP-raising to SPEC-T or V-raising to T.

They basically assume, as Chomsky does, that the parametric difference is attributed to richness of agreement, proposing that in non-NSLs like English, with weak ϕ -agreement marking on V, EPP is satisfied by DP-raising to SPEC-T, whereas in NSLs like Italian, with rich ϕ -agreement marking on V, EPP can be satisfied by V-raising to T (as well as by DP-raising to SPEC-T). Since it is easy to find morphological evidence for the postulation of such a difference, I assume their approach to EPP-satisfaction in the labeling theory as a principled way to satisfy EPP, or more precisely, to ensure $\langle \phi, \phi \rangle$ labeling of the SPEC-TP construction. Then, following Chomsky's (2013: 45) suggestion that so-called SPEC-Head agreement be mandatory for $\langle \phi, \phi \rangle$ labeling, I assume Toyoshima's (2001) Head-to-SPEC movement hypothesis, according to which all movement is to a SPEC position, whether the movement is of a head or of a phrase (see Fukui and Takano 1998 for a similar view). This hypothesis also seems to me to make sense conceptually, in particular in the context of the simplest Merge-based system: if Merge applies freely, as in (2), and if Merge H-to-S is just an instance of the general application of Merge, then there is nothing to preclude the possibility, as is the case of Merge of a phrase to SPEC: once we accept the simplest Merge-based system, anything goes; rather, any stipulation that blocks it carries a burden of proof.

Only with these assumptions can Chomsky's labeling analysis of EPP be maintained: in non-NSLs like English, the SPEC-TP construction is labeled as $\langle \phi, \phi \rangle$ by raising DP to SPEC-T, as Chomsky argues, whereas in NSLs like Italian, it is done by raising of V to SPEC-T with Merge H-to-S.²

One of the important differences between Chomsky's analysis and ours is that the universality of EPP is not maintained in the former, but in the latter is: in our

analysis, EPP is just a universal property of T, and, therefore, must be observed in all languages by means of the way each language has acquired. Thus, in our theory, there is no need to stipulate the strong/weak parameter on T to satisfy EPP, or to ensure $\langle \varphi, \varphi \rangle$ labeling, which would be a welcome development, in favor of (1).

Given this consideration, the potential problem of asymmetry between T and R with respect to the strength of labeling disappears: T is inherently weak, just like R; hence no need to be parameterized to begin with. Note that “weak” here is nothing other than a cover term for the universal property of the nonphase heads. That is, the term “weak” suggests that T and R are inherently unspecified as to category and their categorial status are derivative from the phase heads, C and v^* . It follows that the nonphase heads lack their labels unless they are selected by the phase heads. This, in turn, allows a uniform characterization of the phase heads and the nonphase heads: the phase heads are the functional elements that are specified as to category, whereas the nonphase heads are the substantive elements that are unspecified as to category. This would be a desirable result for feature inheritance to which I will return.

We conclude, then, that EPP is a universal property of language, holding for both phases CP and v^*P .

Let’s return to the crucial difference between Chomsky’s analysis and ours, i.e., the way of satisfying EPP in Italian-type NSLs: for Chomsky, strong T has the key, but for ours, it is V with rich φ -agreement. Consider the scheme in (3), t a copy of V:

(3) $[\beta V [T [v^* [\alpha DP, t]]]]$

In (3), as we have just said above, β is labeled as $\langle \varphi, \varphi \rangle$ by raising V to SPEC-T.³

Note here that the normal labeling procedure can apply to α : by raising V to SPEC-T, what is visible to LA in α becomes D alone, hence α is labeled D, as required. Thus, in our analysis, the problem of how the predicate-internal subject construction is labeled when the subject stays in situ can be circumvented. One may wonder what would happen if the construction in question is transferred to the interfaces with the label of D, not that of v^* , contrary to the standard assumption. However, there is no conceptual argument against this; hence no reason to reject what seems to be different from the standard. Rather, labeling the structure in question as D may shed new light on an important observation by Belletti (2001, 2004) that a postverbal subject is interpreted as new information focus: as in (3), what is prominent in the remaining structure is the label of the postverbal subject D, so that the postverbal subject might be showing a strong tendency to be interpreted as new information focus at the interfaces.

One of the consequences of our theory gives a straightforward answer to the observation that the *that*-trace effect is obviated by a presence of the expletive *there*. Consider (4) below:

- (4) a. *What_i do you think that [_{α} t_i is in the box]?
- b. What_i do you think that [_{α} there is t_i in the box]?

The ungrammaticality of (4a) (i.e., the *that*-trace effect) has already been explained by Chomsky's theory: the *wh*-subject (here, *what*) must stay in SPEC-T to keep α labeled as $\langle \varphi, \varphi \rangle$; otherwise, the derivation results in a labeling failure or a PIC violation (see section 1). He does not provide an explanation for the grammaticality of an example like (4b), but once we incorporate Merge H-to-S, it becomes possible to account for it

without extra stipulation: since α stays labeled as $\langle \varphi, \varphi \rangle$ thanks to Merge of *there* to SPEC-T, further *wh*-extraction is allowed. Note that the analysis is compatible with the independently motivated assumption made in Sabel (2000) and Hornstein and Witkoś (2003) that the expletive *there* is a head, and shares the values of its φ -features with its associate by forming a constituent underlyingly, as in $\{D, NP\}$. Details aside, Merge H-to-S actually permits a loophole for further extraction from the criterial position, here, SPEC-T.

Significantly, this is in fact reminiscent of Rizzi and Shlonsky' (2007) "skipping strategies" – that is, XP can raise further, if the criterial position is occupied by another element. Thus, without recourse to the strength of T, we can now reinterpret Chomsky's labeling analysis of ECP in Italian-type NSLs as follows: the SPEC-TP construction stays labeled as $\langle \varphi, \varphi \rangle$ thanks to Merge of V to SPEC-T (see β in (3)), hence further *wh*-extraction is allowed. The same analysis is true of the following *que-qui* contrast in French:

(5) a. *Quelle étudiante_i crois-tu que t_i va partir?

‘Which student do you believe that is going to leave?’

b. Quelle étudiante_i crois-tu qui t_i va partir?

‘Which student do you believe QUI is going to leave?’

(5a) is ungrammatical for the same reason as (4a): a labeling failure or a PIC violation. Also, the ungrammaticality of (5b) can be accounted for in the same way as (4b), particularly in collaboration with Taraldsen's (2001) proposal that French *qui* consists of *que* and expletive: thanks to Merge of the expletive *-i* to SPEC-T, the relevant

SPEC-TP construction stays labeled as $\langle \varphi, \varphi \rangle$, and hence further *wh*-extraction is allowed.

In this way, once we adopt the possibility of labeling by Merge H-to-S, EPP, ECP, and their parametric differences in Chomsky (2015), as well as the phenomena that can be captured by skipping strategies in Rizzi and Shlonsky (2007) can be unified under the labeling theory,

Let us turn finally to the indispensable device, feature inheritance, which has been a well-accepted mechanism in phase theory ever since Richards (2007) provided a clear rationale for it. He claims that the necessity of feature inheritance from C to T and v^* to V (“inheritability,” in his terms) follows by conceptual necessity from two basic assumptions about the phase-based derivational system: “interpretability” (6a) and “impenetrability” (6b) (= his (1) and (2), respectively):

(6) a. Value and Transfer of uFs must happen together.

b. The edge and nonedge (complement) of a phase are transferred separately.

(6a) is the interface condition (cf. Epstein and Seely 2002) and (6b) is the effect of PIC (cf. Chomsky 2000, 2001, 2004). By way of illustration, consider the following CP phase:

(7) [C [α DP [T [...]]]]

Here, since C is a phase head, it follows from (6b) that what is transferred is α . In this situation, if C retains φ -features without feature inheritance, Value and Transfer of φ -

features of C and those of DP cannot happen together at α , failing to satisfy (6a); but on the other hand, if C hands φ -features onto T with feature inheritance, Value and Transfer of φ -features of T and those of DP can happen together at α , being able to satisfy (6a). Hence, feature inheritance is necessary, for Richards, to satisfy two assumptions.

Bearing this argument in mind, let us turn to Chomsky's explanation of why *wh*-extraction from SPEC-T is allowed when C is deleted, as illustrated in (8):

(8) Who do you think \emptyset [α t T [β read the book]] (“who do you think read the book?”)

Of particular relevance here is the shift of transfer domain: to circumvent a PIC violation, Chomsky notes that phasehood of C is inherited by T along with φ -feature inheritance and is activated on T due to C-deletion. Thus, in (8), T is a phase head, and β is transferred, in conformity with PIC/(6b), so that a PIC violation by further *wh*-extraction (cf. the trace of *who* in α) is circumvented (for labeling of α , see section 1). It is important to note here that Richards' argument for feature inheritance from PIC is untenable: in (8), Value and Transfer of φ -features of T and those of DP (here, *who*) cannot happen together, just because what is transferred is β , for the reason just given. Thus, unfortunately, Richards' argument for feature inheritance is undermined.

However, this does not mean that feature inheritance is an unwarranted stipulation in the labeling theory; rather it receives a general rationale under our theory. Recall that in the previous section, I have suggested a uniform characterization of the phase heads and the nonphase heads: the phase heads are category-specified elements, and the nonphase heads are category-unspecified

elements. This characterization suggests that the phase heads C and v* inherently have the ability of labeling (“labelability”) but the nonphase heads T and R do not:

(9) C and v* have labelability but T and R do not.

Note that this view presents a new rationale for feature inheritance in collaboration with Chomsky’s (2013: 45-46) interpretation of interpretability in terms of labeling, defined as follows:

(10) “all SO that reach the interfaces must be labeled” (p. 45); “a label is required for interpretation at the interfaces” (p. 46)

Given (10), T and R must be labeled before they are sent to the interfaces. It follows that feature inheritance is necessary for them to be labeled in conformity with (10): without feature inheritance, they would be unable to qualify for labeling, but with it, they could qualify for it. Thus, in our theory, the inheritability follows by two basic principles about the labeling theory – that is, “labelability” (9) and “interpretability” (10):

(11) Feature inheritance is required for labeling for interpretation at the interfaces.

In this way, the status of feature inheritance is maintained and enhanced, provided with a simple principled explanation under the labeling theory. This would be a desirable result for the phase-based system.

References

- Alexiadou, Artemis and Elena Anagnostopoulou. 1998. Parameterizing Agr: Word order, V-movement and EPP checking. *Natural Language and Linguistic Theory* 16, 491-539.
- Belletti, Adriana. 2001. "Inversion" as focalization. In *Inversion in Romance and the theory of universal grammar*, ed. by Aafke Hulk and Jean-Yves Pollock, 60-90. Oxford: Oxford University Press.
- Belletti, Adriana. 2004. Aspects of the low IP area. In *The Structure of IP and CP: The cartography of syntactic structures, vol. 2*, ed. by Luigi Rizzi, 16-52. New York: Oxford University Press.
- Richards, Marc and Theresa Biberauer. 2005. Explaining expl. In *The function of function words and functional categories*, ed. by Marcel den Dikken and Christina Tortora, 115-154. Amsterdam: Benjamins
- Biberauer, Theresa and Ian Roberts. 2008. Subjects, tense, and verb movement in Germanic and Romance. *Cambridge Occasional Papers in Linguistics (COPiL)* 3, 24-43.
- Borer, Hagit. 2005a. *In name only: Structuring sense*, Volume I. Oxford: Oxford University Press.
- Borer, Hagit. 2005b. *The normal course of events: Structuring sense*, Volume II. Oxford: Oxford University Press.
- Borer, Hagit. 2013. *Taking form: Structuring sense*, Volume III. Oxford: Oxford University Press.
- Broekhuis, Hans. 2000. Against feature strength: The case of scandinavian object shift. *Natural Language and Linguistic Theory* 18, 678-721.

- Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures*. Berlin: Mouton de Gruyter.
- Chomsky, Noam. 1998. Some observations on economy in generative grammar. In *Is best good enough? Optimality and competition in syntax*, ed. by P. Barbosa et al, 115-127. Cambridge, Mass.: MIT Press.
- 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, Mass.: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1-52, Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and beyond: The cartography of syntactic structures*, Vol. 3, ed. by Adriana Belletti, 104-131, Oxford: Oxford University Press.
- Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36, 1-22.
- Chomsky, Noam. 2007. Approaching UG from below. In *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, ed. by Uli Sauerland and Hans-Martin Gärtner, 1-29, Berlin: Mouton de Gruyter.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. by Robert Freiden, Carlos Peregrino Otero & Maria Luisa Zubizarreta, 133-166. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130, 33-49.
- Chomsky, Noam. 2015. Problems of projection: Extensions. In *Structures, strategies and beyond – studies in honour of Adriana Belletti*, ed. by Elisa Di Domenico, Cornelia Hamann & Simona Matteini, 3-16. Amsterdam/Philadelphia: John

Benjamins.

- Epstein, Samuel David, and Daniel Seely. 2002. Rule applications as cycles in a level-free syntax. In *Derivation and explanation in the minimalist program*, ed. by Samuel David Epstein and T. Daniel Seely, 65–89. Oxford: Blackwell.
- Epstein, Samuel David, Hisatsugu Kitahara, and Daniel Seely. 2014. Labeling by minimal search: Implications for successive-cyclic A-movement and the conception of the postulate “phase.” *Linguistic Inquiry* 45, 463-481.
- Epstein, Samuel David, Hisatsugu Kitahara, and Daniel Seely. 2015. *Explorations in maximizing syntactic minimization*. New York and London: Routledge.
- Fukui, Naoki, and Yuji Takano. 1998. Symmetry in syntax: merge and demerge. *Journal of East Asian Linguistics* 7, 27-86.
- Goto, Nobu. 2013. Labeling and scrambling in Japanese. *Tohoku: essays and studies in English language and literature* 46, 39-73.
- Goto, Nobu. to appear. Labelability = extractability: Its theoretical implications for the free-merge hypothesis. In *Proceeding of NELS 46*.
- Hornstein, Norbert and Jacek Witkoś, J. 2003. Yet another approach to existential constructions. In *Grammar in focus: Festschrift for Christer Platzack*, ed. by L.-O. Delsing, C. Falk, G. Josefsson, and Halldór Ármann Sigurðsson, 167-184. Lund: Wallin and Dalholm.
- Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.
- Richards, Marc. 2007. On feature inheritance: An argument from the phase impenetrability condition. *Linguistic Inquiry* 38, 563-572
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

- Rizzi, Luigi and Ur Shlonsky. 2007. Strategies of subject extraction. In *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, ed. by Hans Martin Gärtner and Uli Sauerland, 115-160. Berlin: Mouton de Gruyter.
- Sabel, Joachim. 2000. Expletives as features. In *Proceedings of the 19th West Coast Conference on Formal Linguistics*, ed. by R. Billerey & B.D. Lillehaugen, 411-424. Somerville, MA: Cascadilla Press.
- Taraldsen, Knut T. 2001. Subject extraction, the distribution of expletives and stylistic inversion. In *Subject inversion in Romance and the theory of universal grammar*, ed. by Aafke Hulk and Jean-Yves Pollock, 163-182. New York: Oxford University Press.
- Toyoshima, Takashi. 2001. Head-to-Spec movement. In *The minimalist parameter: Selected papers from the Open Linguistics Forum, Ottawa, 12–23 March 1997*, ed. by Galina M. Alexandrova and Olga Arnaudova, 115–136. Amsterdam: John Benjamins.

Notes

*Acknowledgements to be added.

¹ V here is a synonym for R. Hereafter I will use it without distinction from R.

² It is well known that French is a non-NLS like English but has V-raising to T like Italian. Perhaps, this follows given that ϕ -agreement marking on French V is weak (see Biberauer and Roberts 2008).

³ V in SPEC-T must be a more complex element like [R-v*]. The nature of the amalgam has to be clarified, but the argument presented here is perfectly compatible with Chomsky's (2015) assumption that "although R cannot label, the amalgam [R-v*] can."