

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

The Variables of VP Ellipsis

**Permalink**

<https://escholarship.org/uc/item/1q275271>

**Author**

Turnbull-Sailor, Craig William

**Publication Date**

2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA  
Los Angeles

## **The Variables of VP Ellipsis**

A dissertation submitted in partial satisfaction  
of the requirements for the degree  
Doctor of Philosophy in Linguistics

by

**Craig William Turnbull-Sailor**

2014

© Copyright by  
Craig William Turnbull-Sailor  
2014

ABSTRACT OF THE DISSERTATION

## The Variables of VP Ellipsis

by

**Craig William Turnbull-Sailor**

Doctor of Philosophy in Linguistics

University of California, Los Angeles, 2014

Professor Timothy A. Stowell, Chair

A constituent containing the main predicate of a clause can go unpronounced, as in *Mary will leave before John will* [-], when certain syntactic, semantic, and discourse conditions are met. This process has come to be known as “VP Ellipsis” (VPE), but this term is misleading: it implies that non-verbal predicates cannot be omitted in the same fashion (they can be), and that VP is the constituent undergoing the operation in question elsewhere (it isn’t). This dissertation focuses on the second point, recast here as a research question that has received surprisingly little attention in an otherwise robust literature: exactly what constituent(s) does VPE operate on?

I argue that VPE is a non-uniform operation: two distinct “sizes” of VPE can be diagnosed according to the amount and variety of material that can be omitted under identity with some salient antecedent. I provide a handful of diagnostics that reveal this distinction in VPE size, and I show that, surprisingly, this distinction tracks a previously-known but ill-understood observation in the VPE literature: namely, in certain environments, VPE can apply within a clause whose grammatical voice does not match that of its antecedent (e.g. passive vs. active), while in other environments, such instances of VPE are unacceptable. The diagnostics of VPE size that I present suggest that smaller instances of VPE correlate with those environments that allow voice-mismatches in VPE, whereas larger instances

of VPE correlate with configurations that resist such voice-mismatches. I argue that this follows if grammatical voice is encoded in the syntax on a dedicated functional projection located at the edge of the main predicate, and the different sizes of VPE are distinguished by whether the ellipsis site is large enough to include this functional projection (“high-VPE”) or is not large enough to include it (“low-VPE”), such that this head remains intact, and thus free to differ featurally from its antecedent (leading to voice-mismatch). This analysis posits a principled, fine-grained distinction in ellipsis size where only coarser distinctions (e.g. VP vs. TP ellipsis) were thought to exist.

The dissertation of Craig William Turnbull-Sailor is approved.

Andrew Simpson

Carson T. Schütze

Anoop K. Mahajan

Timothy A. Stowell, Committee Chair

University of California, Los Angeles

2014

*To my parents, for everything*

## TABLE OF CONTENTS

<b>1</b>	<b>Principles of Variation in VP Ellipsis Size</b>	<b>1</b>
1.1	Introduction	1
1.2	Setting up the research question	2
1.3	Background and assumptions	7
1.3.1	The inflectional domain	7
1.3.2	VP ellipsis	11
1.3.3	Discourse coherence	16
1.4	Diagnosing VPE size	18
1.4.1	Methodology	18
1.4.2	The distribution of strict identity in VPE	20
1.4.3	Preferential (non-)recovery of antecedent VP-modifiers	23
1.4.4	Aspectual mismatch and auxiliary omission	27
1.4.5	Summary and consequences	36
<b>2</b>	<b>On the Fine Structure of Low-VPE</b>	<b>38</b>
2.1	Introduction	38
2.2	<i>Being</i> being stranded where it shouldn't be (*being)	39
2.3	Choices of Voice <sup>0</sup> s	42
2.3.1	Raising your Voice: On apparent variation in low-VPE size	44
2.4	Supporting and extending the analysis: GET-passives	54
2.4.1	Support: GET can't get stranded, but BE can (*get)	54
2.4.2	Extensions: GET-passives and the identity condition on ellipsis	57



2.4.3	Summary . . . . .	62
2.5	Low-VPE vs. high-VPE: The Principle of Antecedent-Containment Avoidance . . . . .	63
2.5.1	On the locus of subordination . . . . .	64
2.5.2	Background on antecedent-contained deletion . . . . .	68
2.5.3	Subordination can create ACD configurations . . . . .	70
2.5.4	The costs of antecedent-containment . . . . .	72
2.6	Summary: On the uniform size of low-VPE . . . . .	74
<b>3</b>	<b>Case Studies in High-VPE: Retorts and Tag Questions . . . . .</b>	<b>77</b>
3.1	Introduction . . . . .	77
3.2	Retorts . . . . .	78
3.2.1	Introduction . . . . .	78
3.2.2	Characteristic properties of retorts . . . . .	87
3.2.3	The derivation of retorts . . . . .	99
3.2.4	On some differences between CIP and LEP retorts . . . . .	110
3.2.5	On subjectless retorts . . . . .	115
3.2.6	Summary of retorts . . . . .	121
3.3	Retorts and tag questions as high-VPE configurations . . . . .	122
3.3.1	Neophobia . . . . .	122
3.3.2	Pronominal subjects . . . . .	122
3.3.3	Ellipsis . . . . .	125
3.3.4	Summary . . . . .	131
3.4	Microvariation in high-VPE . . . . .	131
3.5	Toward an analysis of high-VPE . . . . .	136

<b>4</b>	<b>On Grammatical Privilege</b>	<b>138</b>
4.1	Introduction	138
4.2	The grammatical privilege of VoiceInflP	139
4.2.1	Argument 1: “VP” topicalization	140
4.2.2	Argument 2: Reduced relative clauses (“whiz-deletion”)	141
4.2.3	Argument 3: Pseudoclefting	143
4.2.4	Argument 4: Predicate wh- questions	145
4.2.5	Argument 5: Fragment answers to predicate wh- questions	146
4.2.6	Argument 6: Participle preposing	147
4.2.7	Afterword: Confirming the results using GET-passives	148
4.2.8	Summary	150
4.3	On the theoretical status of privileged XPs	150
4.3.1	Ellipsis	153
4.3.2	Propositionality	155
4.3.3	Summary and outlook	157
	<b>References</b>	<b>158</b>

## ACKNOWLEDGMENTS

Some dissertation acknowledgements begin with pithy remarks about the writing process, or lamentations about the document's difficult birth, or poetry. This one doesn't.

I'd first like to acknowledge to those who helped me most during the authorship of this work. Foremost on that list are my committee members, Tim Stowell, Carson Schütze, Anoop Mahajan, and Andrew Simpson (who was kind enough to step in as an outside member on short notice). I would particularly like to thank my committee chair, Tim Stowell, for his support, his encouragement, his friendship, and his house, which I was fortunate enough to live in for the better part of a year in the capacity of house-sitter while Tim was away on sabbatical. I would also like to extend my deepest gratitude to Carson Schütze, whose patience never waned and whose and thoughtful, comprehensive feedback never subsided, despite receiving countless scatterbrained e-mails, abstracts, and drafts from me, often at the last minute before some looming deadline. He has been an invaluable resource to me. Other current or former UCLA faculty whom I thank for helpful comments on versions of this work include Keir Moulton, Jessica Rett, Yael Sharvit, and Maziar Toosarvandani.

I must also thank my friend, officemate, collaborator, and (at times unwilling)  $\LaTeX$ -troubleshooter Byron Ahn, who has been a reliable source of insightful feedback, irresponsible speculation, and pure entertainment over the years. I feel lucky to have been surrounded by colleagues at UCLA who are both lovely people and damned intelligent. Robyn Orfitelli is another excellent example of such a human, and someone who provided me with equal parts slapstick entertainment and rigorous intellectual exercise. Others include members of my beloved cohort, particularly Grace Kuo, my favorite among them (sorry Natasha Abner). Thanks also to Mel Bervoets, Adam Chong, Nancy Finifter, Laura Kalin, Margaret Kroll, Vic-

toria Mateu, Chad Vicenik, Kaeli Ward, Lauren Winans, and many others whom I've been fortunate enough to call friends and peers during my time in graduate school. It would be terribly remiss of me to neglect thanking Melanie Lynn (née Levin), our former student-affairs officer, who was a true student advocate and all-around life-saver.

Outside of UCLA, I must thank Gary Thoms, Maziar Toosarvandani, and Harold Torrence for their friendship, their helpful discussion of many topics, and for generally making this field a lot of fun to work in. I've benefitted from conversations with many people during the writing process, including Lobke Aelbrecht, Bronwyn Bjorkman, Liliane Haegeman, Will Harwood, Kyle Johnson, Jim McCloskey, and Jason Merchant, among others. Sara Rosen at the University of Kansas deserves prominent mention here as well for convincing me to become a linguist in the first place. Outside of academia, I thank and acknowledge my oldest and closest friends, Doug Hamilton, Dustin Honeyman, and Nick Tallmon. Finally, endless love and respect to my parents and my brother, who have always supported me.

## VITA

- 2005            B.A. (Linguistics, English) University of Kansas.
- 2005-2007      Graduate Research Assistant to Dr. Jie Zhang, Linguistics Department, University of Kansas. Experimental investigations of the ‘tone circle’ in Taiwanese (Southern Min).
- 2007            M.A. (Linguistics), University of Kansas.
- 2008–2009      Graduate Research Assistant to Dr. Timothy Stowell, Linguistics Department, UCLA. The syntax of parenthetical constructions.
- 2008–2013      Teaching Assistant, Linguistics Department, UCLA. Taught sections of Linguistics 20 (Introduction to Linguistics), 120b (Introduction to Syntax), 127 (Syntactic Typology), among others.

## PUBLICATIONS AND PRESENTATIONS

Sailor, Craig. Submitted. The typology of head movement and ellipsis: a reply to Lipták and Saab (to appear).

Sailor, Craig and Gary Thoms. To appear. On the non-existence of non-constituent coordination and non-constituent ellipsis. To appear in *Proceedings of WCCFL 31*. Cascadilla Proceedings Project.

Ahn, Byron and Craig Sailor. 2014. Obligatory object gaps in infinitival clauses. Presented at LSA 88. Minneapolis, MN.

Sailor, Craig and Anoop Mahajan. 2013. Toward a derivational typology. In *Proceedings of NELS 42*. GSLA, UMass.

Sailor, Craig. 2013. Questionable negation. Presented at LSA 87. Boston, MA.

Sailor, Craig. 2012. On embedded polar replies. Presented at *Workshop on the Syntax of Answers to Polar Questions*. Newcastle University.

Sailor, Craig. 2012. The size of silence: diagnosing the fine structure of VP ellipsis. Presented at the *Talks in Linguistics* series, University of Illinois at Chicago.

Sailor, Craig. To appear (submitted 2011). VP ellipsis in tag questions: a typological approach. To appear in *Proceedings of CLS 46*. University of Chicago Press.

Zhang, Jie, Yuwen Lai, and Craig Sailor. 2011. Modeling Taiwanese speakers' knowledge of tone sandhi in reduplication. *Lingua* 121:181-206.

# CHAPTER 1

## Principles of Variation in VP Ellipsis Size

### 1.1 Introduction

A constituent containing the main predicate of a clause can go unpronounced, as in *Mary will leave before John will* [-], when certain syntactic, semantic, and discourse conditions are met. This process has come to be known as “VP Ellipsis” (VPE), but this term is misleading: it implies that non-verbal predicates cannot be omitted in the same fashion (they can be), and that VP is the constituent undergoing the operation in question elsewhere (it is not). This dissertation focuses on the second point, recast here as a research question that has received surprisingly little attention in an otherwise robust literature: exactly what constituent(s) does VPE operate on?

I argue that VPE is a non-uniform operation: two distinct “sizes” of VPE can be diagnosed according to the amount and variety of material that can be omitted under identity with some salient antecedent. I provide a handful of diagnostics that reveal this distinction in VPE size, and I show that, surprisingly, this distinction tracks a previously-known but ill-understood observation in the VPE literature: namely, in certain environments, VPE can apply within a clause whose grammatical voice does not match that of its antecedent (e.g. passive vs. active), while in other environments, such instances of VPE are unacceptable. The diagnostics of VPE size that I present suggest that smaller instances of VPE correlate with those environments that allow voice-mismatches in VPE, whereas larger instances

of VPE correlate with configurations that resist such voice-mismatches. I argue that this follows if grammatical voice is encoded in the syntax on a dedicated functional projection located at the edge of the main predicate, and the different sizes of VPE are distinguished by whether the ellipsis site is large enough to include this functional projection (“high-VPE”) or is not large enough to include it (“low-VPE”), such that this head remains intact, and thus free to differ featurally from its antecedent (leading to voice-mismatch). This analysis posits a principled, fine-grained distinction in ellipsis size where only coarser distinctions (e.g. VP vs. TP ellipsis) were thought to exist.

## 1.2 Setting up the research question

An E(llipsis)-clause can differ in grammatical voice from its A(ntecedent)-clause in certain “VP” ellipsis (VPE)<sup>1</sup> contexts (Sag 1976:75, fn. 2), but never in sluicing contexts (Merchant 2001; examples taken from Merchant 2013b):

*VPE: Voice mismatches possible (sometimes)*

- (1) a. The janitor must remove the trash whenever it is apparent that it should be ~~removed~~.
- b. This guy’s tape obviously should be scrutinized more than you did ~~scrutinize it~~.

*Sluicing: Voice mismatches never possible*

---

<sup>1</sup>As Sag (1976) and others have noted, the term “VP” ellipsis/deletion is potentially misleading for a couple of reasons: first, it implies that only verbal predicates can undergo this operation, which is false – in principle, any predicate type can; and, second, it implicates a very specific syntactic projection, even though it is an open question whether VP is in fact the relevant constituent undergoing the operation (even when a verbal predicate is involved). In fact, the position I take in this dissertation – to be justified in detail in chapter 2 – is that a projection larger than VP (and even *v*P) is in fact the relevant constituent affected by this operation. Nevertheless, for simplicity, I use the standard appellation “VPE” throughout this dissertation (rather than Sag’s own suggestion of “post-auxiliary ellipsis”).



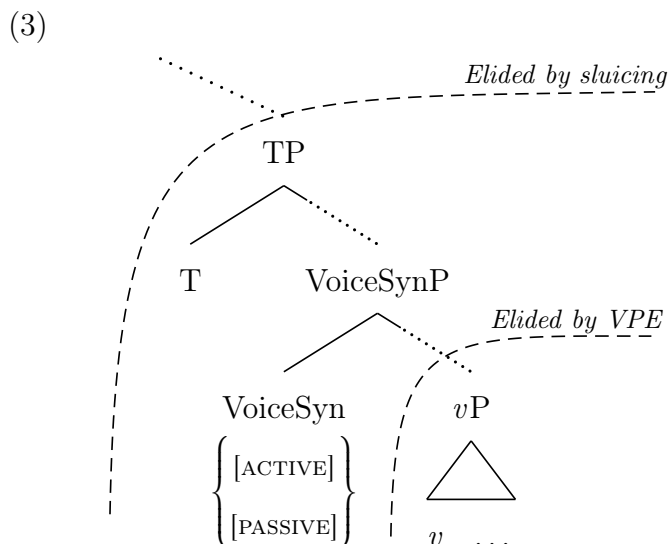
- (2) a. \*Someone murdered Joe, but we don't know who by [~~Joe was murdered~~].  
 b. \*Joe was murdered, but we don't know who [~~murdered Joe~~].

Merchant (2013b) argues that this asymmetry in the acceptability of voice mismatch (VMM) with VPE vs. sluicing finds a straightforward explanation if (i) grammatical voice is encoded syntactically as features on a head – call it the VoiceSyn(tax) head<sup>2</sup> – that selects a projection of the main verb (Kratzer 1996, Sailor and Ahn 2010, a.o.), and (ii) ellipsis is subject to a syntactic identity requirement, i.e. the elided structure must be featurally identical to its antecedent (see Merchant 2013a for a survey of ellipsis identity). Under these assumptions, the pattern above is simply a by-product of the difference between the sizes of structure that each operation elides: VMM would only be possible when the E-clause VoiceSyn<sup>0</sup> is excluded from the elided material, allowing it to vary freely from its analogue in the A-clause. If VoiceSyn<sup>0</sup> is part of the elided material, the identity requirement prohibits its features from differing from those on the A-clause's VoiceSyn<sup>0</sup>, thereby preventing VMM. This, Merchant argues, is the source of the VMM asymmetry between VPE and sluicing: VPE elides a “low” structure that properly excludes VoiceSyn<sup>0</sup>, while sluicing elides a “high” constituent that necessarily includes it:<sup>3</sup>

---

<sup>2</sup>Anticipating aspects of the coming analysis, the choice of name here is intended to distinguish this head from the head associated with voice inflection (e.g. passive *-en*), which I will label “VoiceInfl(ection)”. A detailed justification for distinguishing these two heads can be found in chapter 2.

<sup>3</sup>The dotted branches in (2) represent the presence (or availability) of additional structure that I have obscured for clarity. Later, I argue that VPE actually elides a constituent slightly larger than *vP*. Here and throughout I leave aside discussion of middle voice; see Merchant (2013b).



In Merchant’s (2013b) terms, an ellipsis operation is “low” if it excludes (i.e. is lower than) VoiceSynP, and “high” if it includes (i.e. is higher than) VoiceSynP. The distinction hinges entirely on the inclusion or exclusion of VoiceSyn<sup>0</sup> in the ellipsis site; that is, the (non-)availability of VMM is dictated by the size of ellipsis alone.

If this were the whole story, then, in principle, VMM would be available in every instance of VPE (and unavailable in every instance of sluicing). Importantly, though, VMM is not available in every instance of VPE (cf. (1)):

*Ungrammatical voice mismatches in VPE*

- (4) a. \*The janitor removed the trash, but the recycling wasn’t ~~removed~~.  
 b. \*This guy’s tape should be scrutinized by John, and Bob also should ~~scrutinize it~~.

If VPE is a uniform operation, the ungrammaticality of examples such as this does not follow from Merchant’s (2013b) structural account: VPE is a “low-ellipsis” phenomenon (VoiceSynP is not included in the ellipsis site), meaning VMM should never be responsible for a violation of the syntactic identity requirement. Merchant

(2013b) acknowledges that VMM is ungrammatical in environments such as these (cf. Sag 1976:17), but he does not offer an explanation; rather, he sets such cases aside as “not representative of the full class of relevant data” (p. 80), and focuses his discussion strictly on the grammatical examples, as in (1).

In this dissertation, I argue that VPE is not a uniform phenomenon. I claim that Merchant’s “low” versus “high” distinction in ellipsis size should be extended to capture the VMM asymmetry within VPE (i.e., (1) vs. (4)). In other words, I argue that VMM diagnoses (at least) two different sizes of VPE: in environments that allow VMM, VoiceSyn<sup>0</sup> is outside the ellipsis site; in environments that prohibit VMM, VoiceSyn<sup>0</sup> is inside the ellipsis site. This implies that the span of structure that VPE elides in examples such as (1) is demonstrably smaller than the one VPE elides in e.g. (4), a prediction I confirm with independent evidence. VPE is therefore not a homogeneous operation, as Merchant and others have assumed. Ellipsis in the verbal domain is either “low-VPE” (VoiceSynP is left intact) or “high-VPE” (VoiceSynP is elided, à la sluicing).

Crucially, this variability in VPE size is detectable even when the E(llipsis)-clause and A(ntecedent)-clause match in their voice features. The implication is that VPE size is not dictated solely by identity: often, E-clause material that is identical to A-clause material is nevertheless left stranded outside the VPE site. In support of this, I introduce a handful of other diagnostics of VPE size later in this chapter. The first diagnostic I discuss involves the distribution of strict identity interpretations with instances of VPE whose antecedents contain reflexive anaphors: here, high-VPE resists such interpretations, while low-VPE is amenable to them. Next, I define a diagnostic involving antecedent VP modifiers, showing that they are preferentially recovered in high-VPE and preferentially excluded from the recovery of low-VPE. Lastly, I discuss patterns of auxiliary (non-) omission as a diagnostic of VPE size, showing that high-VPE licenses the omission of particular inflectional auxiliaries that low-VPE does not. I show that each of

these diagnostics distinguishes low-VPE from high-VPE in a manner similar to Merchant’s use of VMM to distinguish low-VPE from sluicing. We will see that high-VPE and sluicing pattern alike with respect to each of these diagnostics, while low-VPE patterns very differently.

All of this raises an important question: what linguistic factors influence VPE size? A debate over this question is ongoing. Recent proposals have implicated a range of linguistic factors in the distinction between (1) and (4), including coherence relations (Kehler 2002, SanPietro et al. 2012; see §1.3.3), information structure (Kertz 2010), conversational implicatures (Grant et al. 2012), and other factors (Arregui et al. 2006, Clifton and Frazier 2010, Kim et al. 2011). Kennedy (2003:26) speculates that the factor(s) responsible for the distribution of VMM in VPE “would not govern the syntax of ellipsis *per se*, but rather would govern the felicity of particular uses of ellipsis”, citing a phenomenon where an extra-syntactic factor (information structure) had been argued to “crucially determine [the phenomenon’s] felicity...but not its syntactic well-formedness”. Below, I will argue – contra Kennedy (and Kehler 2002), but in keeping with the spirit of Kim and Runner (2011) – that the most immediate factor dictating the (un)availability of VMM is syntax. If its availability or unavailability is indeed a function of VPE size, as I argue, then the aforementioned factors do not simply dictate when VMM is possible vs. impossible, they dictate when VPE can be structurally low vs. high. In other words, if extra-syntactic factors are at work, then they influence the syntax of VPE; they do not bypass it. This position must be taken into consideration in future work dealing with the syntactic identity condition on VPE.

Beyond laying out the relevant background and assumptions, my focus in this chapter lies with establishing some diagnostics of VPE size. In chapter 2, I focus on low-VPE, arguing that it targets the maximal projection of the inflectional head associated with passive morphology (*-en*), which I take to be projected in the structure even in non-passive clauses. Given that VMM is possible with low-VPE,

this necessitates the projection of two voice-related heads in the verbal domain: a lower one associated with inflection that elides in low-VPE, and a higher one associated with the clause’s grammatical voice features that survives low-VPE. In chapter 3, I turn to high-VPE, appealing to a case study of that phenomenon in English involving what I call *retorts*, a speech act rejecting a prior assertion. I show that retorts share with tag questions a particular set of properties relating to ellipsis and anaphoricity that makes them the ideal high-VPE environments, and I discuss some consequences for the analysis of high-VPE. Chapter 4 considers some consequences of these findings for constituency, both within the verbal domain and in general.

## 1.3 Background and assumptions

### 1.3.1 The inflectional domain

The English inflectional domain comprises a layered array of aspectual auxiliaries and the bound affixes associated with them, bookended by tense and modals at the top of the array, and the main verb at the bottom.

- (5) The cheesecake should have been being eaten.
- (6) modals > *-ed* (tense ‘PAST’) > *-en* (asp. ‘PERF’) > *-ing* (asp. ‘PROG’) > *-en* (voice ‘PASS’)

Since Tenny (1987), it is generally accepted that this fixed order of inflectional material reflects (part of) a universal hierarchy of functional projections in syntax (cf. Cinque 1999). Because each of these heads can have morphological exponence within a single clause, I will refer to them with evocative names, following various authors (Cinque 1999, Bjorkman 2011, a.o.). I will refer to the projection that is associated with passive inflectional morphology (*-en*) as “VoiceInfl(ection)P”, to

distinguish it from the head described above that controls the syntax of grammatical voice. To collapse the two without justification would be to assume an analysis that may not be supported by the facts;<sup>4</sup> so, for now, I will assume they are separate, though the head associated with the syntax of grammatical voice will not be relevant for us until much later in the discussion. For now, our focus is with the inflectional hierarchy, which I represent below in (7):

(7) ... TP > PerfP > ProgP > VoiceInfl(ection)P > vP ...

Importantly, the bound affix associated with a particular inflection type is not realized on the auxiliary associated with that inflectional head; instead, it is realized on the next verbal element down, a state of affairs that came to be known as affix “hopping/lowering” following Chomsky’s (1957) initial observation. As the name implies, this was thought of as downward movement (lowering) of the affixes from the heads they are associated with to those they are realized on:<sup>5</sup>

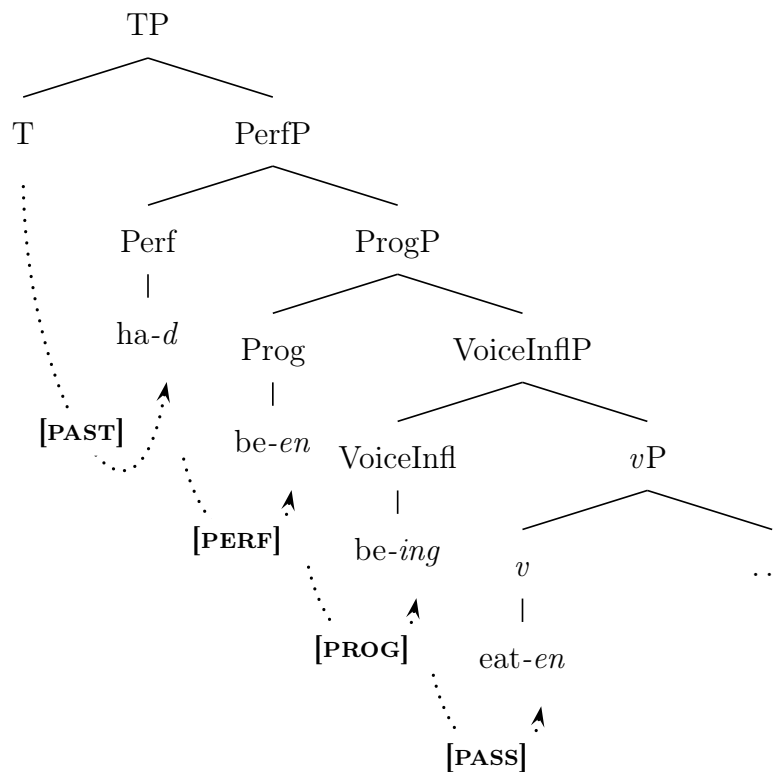
*“Affix-lowering” in the inflectional hierarchy*

(8) ...had been being eaten

---

<sup>4</sup>Indeed, in chapter 2, I argue that VoiceInflP is distinct from, and just below, the head that drives the syntax of grammatical voice.

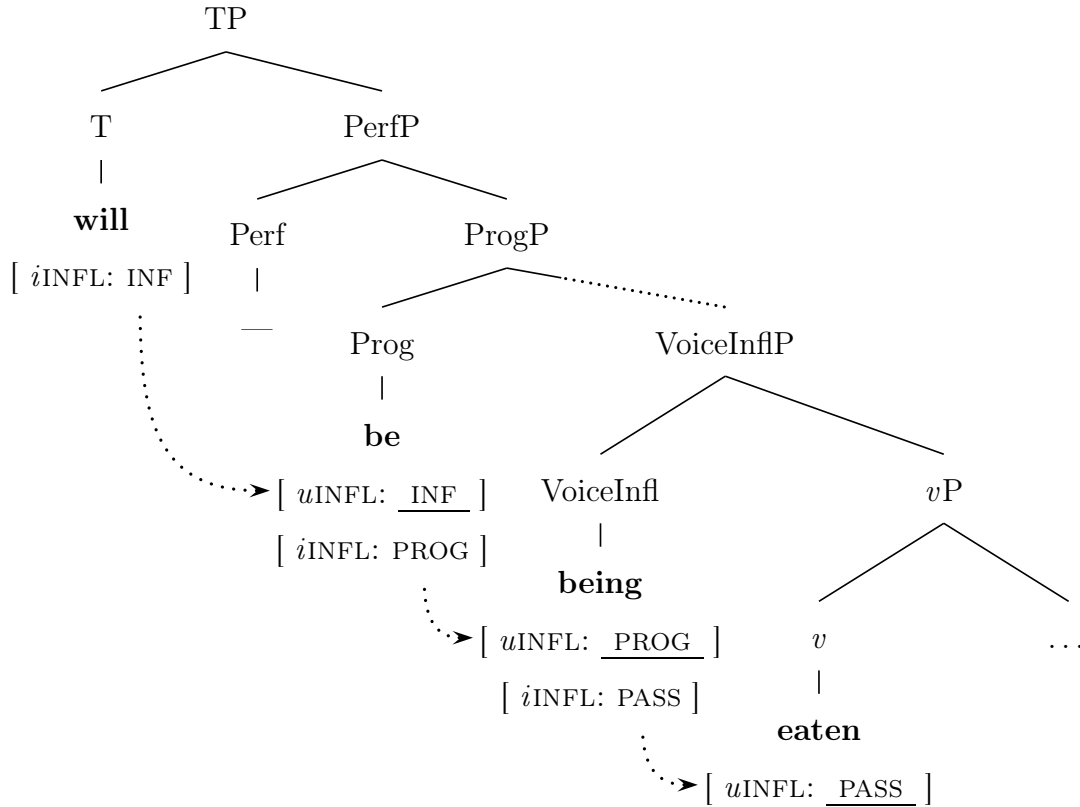
<sup>5</sup>Under standard assumptions, the highest auxiliary (*had*) raises to T<sup>0</sup> at some point; I leave this aside here.



Recently, Bjorkman (2011) has argued that the affix-lowering effect does not arise via movement, but via *Agree*: a head bearing an interpretable inflectional feature with a value  $F$  ( $[i\text{INFL}: F]$ ) must undergo *Agree* with another, lower inflectional head bearing an unvalued uninterpretable inflectional feature,  $[u\text{INFL}: \_\_]$ . This newly-valued uninterpretable feature  $[u\text{INFL}: \underline{F}]$  becomes a target for Vocabulary Insertion at the end of the derivation; thus, the morpheme corresponding to the feature value  $F$  (e.g. *-ing* if  $F = \text{PROG}$ ) ends up being pronounced one head below its position of interpretation. An illustration of this system is in (9), where dashed lines represent instances of *Agree*.<sup>6</sup>

<sup>6</sup>Note that Bjorkman (2011) uses the label “AspP” for the constituent immediately dominating VoiceInflP (although AspP is still crucially distinct from PerfP in her system). Since progressive is the primary non-perfect aspect I discuss in this paper, I will refer to this constituent using a more evocative term, “ProgP”. Bjorkman’s VoiceInflP equivalent is simply “VoiceP” for her. The distinction between voice-related syntax (e.g. active vs. passive) and voice-related morphology (e.g. passive *-en*) will be critical to the coming analysis; thus, a distinction in labels for these voice-related heads is necessary from the outset. I use dotted branches here and throughout to indicate the presence of additional obscured structure.

(9) ...will be being eaten



This approach achieves the effect of affix-lowering without invoking movement (upward or downward). More importantly, it has wide crosslinguistic empirical coverage, a virtue that, as Bjorkman (2011:§2.6) argues, other analyses of inflectional systems do not enjoy (e.g. those of Thoms 2010, Harwood 2013a, Aelbrecht and Harwood 2013, a.o.). Note that this approach predicts that different inflectional realizations of BE<sup>7</sup> should be merged in different positions in the tree, since auxiliaries are inserted directly into just those inflectional heads that are “active” in the derivation (i.e., that contain bound inflectional morphology that requires a

<sup>7</sup>I use the notation BE, in small caps, to refer to any and all forms of the auxiliary verb “(to) be”, regardless of its surface morphological realization. Italics (e.g. *be*, *been*) are reserved for referring to specific surface morphological realizations. This carries over to subscripted instances of these notations following a syntactic head, e.g. “VoiceInfl<sub>BE</sub>” (a VoiceInfl head realized by any form of BE, as in any passive clause) versus “VoiceInfl<sub>be</sub>” (a VoiceInfl head realized with the surface form *be*, as in a modal/infinitive passive clause, e.g. *will be written*).



verbal host). For example, the BE that arises in perfect passives would be generated lower in the structure (in VoiceInfl) than the BE that arises in perfect progressives (in Prog), even though the two are homophonous (i.e. *been*). We will see evidence of this from ellipsis in the next chapter; see also Bjorkman (2011:§2.3.6) and references therein (and Harwood 2013b for recent counterarguments).

For concreteness, I will adopt Bjorkman’s Agree-based analysis of inflection throughout, leaving aside discussion of several technical details which will not be important here (including the decomposition of *have*, the “default” status of BE, and the directionality of Agree; see *op. cit.*). For ease of exposition, I occasionally refer to the empirical state of affairs seen above as “lowering”, recognizing that no affixal movement has actually occurred under an Agree-based approach. The choice of a lowering-type inflectional model such as this, as opposed to one of the auxiliary-raising type approaches cited above, is primarily driven by concerns of presentational clarity. While the analysis I develop throughout this dissertation will make heavy use of this model, it is crucial to point out that the key points of the analysis to come can be reconciled with e.g. an auxiliary-raising model (*mutatis mutandis*). Criticisms of the inflectional model I assume here therefore cannot provide serious challenges to the overall analysis to come (and certainly cannot challenge the empirical observations on which the analysis will be developed).

### 1.3.2 VP ellipsis

VPE is the licensed non-pronunciation of any main predicate, verbal or otherwise (making the name somewhat misleading: see fn. 1). Although VPE has enjoyed decades of research within the generative tradition – seminal works include Sag (1976), Lobeck (1995), Johnson (2001), Goldberg (2005), Aelbrecht (2010), among many others – several very basic questions remain open. What are the necessary and sufficient conditions for VPE to be licensed within an utterance, and, more broadly, for a language to have a VPE operation in the first place? What con-

straints are there on the recovery (interpretation, identification) of elided material? What is the syntactic status of the ellipsis site itself: does it lack internal structure like a null proform, or does it have articulated structure like its antecedent, only unpronounced? For recent detailed surveys of the expansive literature addressing these and other questions (both for VPE and other varieties of ellipsis, e.g. sluicing), see Merchant (2013a), van Craenenbroeck and Merchant (2013) and van Craenenbroeck (Forthcoming).

However, the question of precisely what category or categories are targeted by ellipsis (VPE or otherwise) has received almost no direct attention until very recently (Johnson 2004, Aelbrecht 2010, Thoms 2010, Aelbrecht and Harwood 2013, Bošković 2013). This VPE “size” question is particularly relevant for any analysis of the licensing condition on VPE – that is, the syntactic configuration(s) under which VPE may occur. Ideally, these two aspects of the theory of VPE would be reducible to a single property: the licensing configuration would directly determine the size of the ellipsis site.

To that end, two essentially opposing positions have been established in the literature: one which holds that VPE targets a fixed size of structure, and one which holds that it targets structures of varying sizes. The reduction of the “size” question to the licensing question appears feasible only for the latter position. I discuss these two camps in turn, below.

### **1.3.2.1 Camp #1: Size matters**

As the name “VP ellipsis” suggests, the earliest proposals – Akmajian and Wasow (1975), Sag (1976), a.o. – assumed that VP is the affected constituent in VPE (setting aside non-verbal predicate ellipsis). Following the development of a more detailed view of the extended verbal domain, more recent proposals (e.g. Johnson 2004, Aelbrecht 2010) have argued for a slightly larger elided constituent, i.e. *v*P.

Any proposal of this nature – one positing a uniform, fixed size of ellipsis under VPE – faces empirical challenges from the English INFL-domain, and an analytical reduction of the sort described above (whereby the licensing configuration and the choice of elided constituent have a common source) does not seem tractable for this approach. For example, Lobeck’s (1995) enduring analysis that VPE is only properly licensed under a lexically-filled  $T^0$  (or  $Neg^0$ : see Potsdam 1997) does *not* directly bear on the actual size of elided structure. As the VPE licenser, we might expect  $T^0$  to directly license ellipsis of its complement (cf. Merchant 2001); however, this can be easily ruled out by constructing examples that exploit the articulated English INFL-domain (boldface indicates the hypothesized VPE licenser, and the ellipse symbol “ $\circ$ ” indicates elided material):

(10) Mary didn’t leave, but John **might** $_{T^0}$  have $_{PERF^0}$   $\circ$ .

Here, the perfect auxiliary *have* is left outside the ellipsis site, despite the fact that the structurally-superior  $T^0$  is the presumed VPE licenser. In other words, the elided material is not local to its licenser, since *have* intervenes. Adding additional aspectual structure only makes the situation worse:

(11) Mary wasn’t leaving, but John **might** $_{T^0}$  have $_{PERF^0}$  been $_{PROG^0}$   $\circ$ .

If we maintain the intuition that  $T^0$  is indeed responsible for licensing VPE (see Lobeck 1995 for extensive argumentation), then it seems that the VPE “size” question is an entirely independent one.

Recognizing this, Aelbrecht (2010:§3.1) proposes a reconciliation that maintains  $T^0$  as the VPE licenser, and yet allows the size of VPE to be fixed as *vP*. She accomplishes this by appealing to an ellipsis feature [E] (following Merchant 2001) which is borne by both the licensing head,  $T^0$ , and the head taking *vP* as its complement. This feature is satisfied via Agree, potentially at a distance; this

makes the presence of any intervening inflectional material irrelevant.

While this approach achieves the desired effect, it raises a few problems. In particular, it simply stipulates that  $T^0$  and the head selecting  $vP$  undergo Agree, and only for the purpose of triggering ellipsis. Stating this problem differently: even if we could understand why  $T^0$  is the licenser of VPE, we would still be without an explanation for the significance of the head selecting  $vP$ , since (unlike  $T^0$ ) that head fails to play a meaningful empirical role in the data, and bears no special relationship to  $T^0$  in non-VPE clauses.

Another empirical challenge for the fixed-size camp comes from data indicating apparent variability in the amount of verbal material that can be included in the VPE site. From early on in the generative literature, it was noted that VPE appeared to be capable of applying to any one of a small array of constituents in the extended projection of the verb:<sup>8</sup>

- (12) John should have been studying, but
- a. ...Mary shouldn't have been.
  - b. ...Mary shouldn't have.
  - c. ...Mary shouldn't.

This observation helped fuel a lengthy debate about the constituency of the extended verbal domain. At the core of this debate was the category “AUX”, which Ross (1969) proposed to eliminate in favor of a simple proliferation of VPs, since, among other reasons, the latter state of affairs would be straightforwardly compatible with this apparent VPE variability: a single structural description could be stated for verbal ellipsis, and a plurality of VPs within a clause would provide ellipsis with as many potential domains of application (cf. Akmajian and Wasow 1975:237).

---

<sup>8</sup>See §1.4.4 on the interpretations that are (not) available for each of these outputs of VPE.

Although there are reasons to doubt that examples such as (12) reflect true fluctuation in VPE size (see §1.4.4), such data have nevertheless inspired analyses of VPE involving ellipsis sites of varying sizes, as I discuss next.

### 1.3.2.2 Camp #2: It's not the size that counts, but how you lose it

Other proposals have considered the possibility that the constituent elided by VPE varies according to the structure of the clause containing it. Recognizing the challenge posed by examples such as (10) and (11), such approaches have necessarily abandoned  $T^0$  as the (sole) licenser of VPE.<sup>9</sup>

However, abandoning this generalization means giving up a number of attractive generalizations (see Lobeck 1995). We seem forced into the conceptually-displeasing position of allowing an array of potential VPE-licensing heads (e.g.  $T^0$ ,  $\text{Perf}^0$ ,  $\text{Prog}^0$ , etc.), only the lowest of which (with phonological content) can be the actual licenser in a given clause.<sup>10</sup> Moreover, if we simply stipulate that certain auxiliaries bear the right features with no independent justification,<sup>11</sup> then we are in essence treating VPE as a construction, a notion which has no status in Minimalist syntax.<sup>12</sup>

To avoid these shortcomings, Thoms (2010) argues that VPE is not licensed by  $T^0$ , strictly speaking, and yet it is also not dictated by ad-hoc features on certain auxiliaries; instead, it is licensed by (auxiliary) verb movement. Under his analysis, auxiliaries and inflectional affixes alike correspond to independent heads

---

<sup>9</sup>One could imagine tweaking Aelbrecht's (2010) system so that it could allow for variable sizes of ellipsis, namely by postulating an array of heads which could bear the lower [E] feature that gets checked by  $T^0$ . I am not aware that such an approach has ever been developed, however, as it would require many stipulations.

<sup>10</sup>The situation gets even worse if we assume that the size of structure elided by VPE can vary, since we would then lose the generalization that the licenser is always the lowest INFL head in the clause. See ch. 2 for discussion of variable ellipsis size. For an analysis similar to the one described here with multiple potential VPE licensers, see López (1999).

<sup>11</sup>One might say that the relevant licensing features are INFL features, although then one would have to explain why it is always the lowest auxiliary (modulo *being*) that is adjacent to the ellipsis site.

<sup>12</sup>For similar criticisms regarding analyses of sluicing, see Mahajan (2005).

on the clausal spine; in order to achieve convergence, the auxiliaries must undergo movement to link up with the affixes, and it is precisely this movement which can license ellipsis (this is a crude simplification; see *ibid.* for details, which I leave aside here).

The core intuition behind Thoms' proposal is sound: ellipsis as a phenomenon ought to have a principled source; and, in particular, this mechanism ought to have grammatical status independent of the empirical phenomenon (VPE) it has been conscripted to capture. However, his analysis crucially relies on a very specific syntax of the INFL-domain – one which makes crucial use of several English-specific INFL-related heads, severely limiting the crosslinguistic applications of the system. See Bjorkman (2011:§2.5.1) for convincing arguments against such approaches to the INFL-domain.

### 1.3.3 Discourse coherence

Kehler (2002) discusses the VMM asymmetry in (1) versus (4) and argues that its source is pragmatic: only certain types of discourse will allow VMM to arise with VPE, while other discourse configurations prohibit VMM when VPE applies.<sup>13</sup> Kehler's argument is rooted in his general theory of discourse coherence, which is concerned with the relationships that hold between sentences in coherent (versus incoherent) discourse, and the grammatical effects such relationships have, such as influencing antecedent recovery for pronouns (and, obviously, ellipsis sites).

Kehler argues that two families of discourse relations are relevant to the VMM asymmetry: the *Resemblance* family of relations and the *Cause-Effect* family. Ut-

---

<sup>13</sup>Specifically, Kehler (2002) argues that certain discourses require only semantic identity to hold between the VPE site and its antecedent, thus allowing VMM, while other discourses additionally require syntactic identity, which blocks VMM (under the assumption that active and passive predicate structures are syntactically non-identical at a particular level of representation). Semantic vs. syntactic identity will not play a significant role in the present discussion, so I leave this aspect of Kehler's argument aside. For a survey of the literature on the ellipsis identity requirement, see Merchant (2013a).

terances that stand in a Resemblance relation share “commonalities and contrasts among corresponding sets of entities and relations” (Kehler 2002:15).<sup>14</sup> Within Resemblance, the *Parallel* relation describes those discourses involving the aforementioned “commonalities”, and the *Contrast* relation describes those involving the “contrasts” (and see *op. cit.* for discussion of additional Resemblance relations). Examples of each type (not involving ellipsis) are below (Kehler 2002:16):

*Resemblance relations*

- (13) a. Dick Gephardt organized rallies for Gore, and Tom Daschle distributed pamphlets for him. *Parallel*  
 b. Gephardt supported Gore, {but Arme y opposed him / but Arme y supported Bush}. *Contrast*

Sentences that stand in a Cause-Effect relation require “a path of implication connecting a pair of propositions” in the two sentences (Kehler 2002:20). Within Cause-Effect, a few types of relations exist; I exemplify the *Explanation* and *Denial of Preventer* relations here (*op. cit.*):

*Cause-Effect relations*

- (14) a. George is dishonest because he is a politician. *Explanation*  
 b. George is honest, even though he is a politician. *Denial of Preventer*

In (14a), coherence requires presupposing that being a politician implies being dishonest. In (14b), the same implication holds, but negation is involved.

The two families of relations can often be distinguished by the variety of ad-

---

<sup>14</sup>Kehler (2002:15) provides the following formal definition of Resemblance: “...the hearer identifies a relation  $p_1$  that applies over a set of entities  $a_1, \dots, a_n$  from the first sentence  $S_1$ , and a corresponding relation  $p_2$  that applies over a corresponding set of entities  $b_1, \dots, b_n$  from the second sentence  $S_2$ . Coherence results from inferring a common (or contrasting) relation  $p$  that subsumes  $p_1$  and  $p_2$ , along with a suitable set of common (or contrasting) properties  $q_i$  of the arguments  $a_i$  and  $b_i$ .”

verbials and discourse connectives that appear within them: for example, clauses standing in a Resemblance relation are typically coordinated with one another, while those in a Cause-Effect relation are typically connected asymmetrically, i.e. with one clause subordinated under the other. Kehler argues, however, that these characteristics are typical, but not definitional. We will return to this matter shortly.

Getting back to VMM, Kehler argues that VMM is impossible when the ellipsis clause and the antecedent clause stand in a Resemblance relation, as in (15a) (Kehler’s ex. (97)), whereas VMM is compatible with the Cause-Effect relation, as in (15b) (Kehler’s ex. (83)):

- (15) a. \*This problem was looked into by John, and Bob did too.  
*Resemblance*
- b. This problem was to have been looked into, but obviously nobody did.  
*Cause-Effect*

See SanPietro et al. (2012) for experimental confirmation of the effect of discourse coherence on VMM as reported by Kehler.

Having established some preliminaries regarding inflectional structure and VPE, I turn now to the task of diagnosing VPE size.

## 1.4 Diagnosing VPE size

The primary claims I put forth here are, first, that VPE size<sup>15</sup> is variable, and, second, that this variability is principled: it is mediated by the type of configu-

---

<sup>15</sup>It should be noted that although I consistently refer to the relevant phenomenon as one of “VPE size”, one might instead think of it as a phenomenon of “antecedent size” (as e.g. Moulton 2008 does for a subpart of the data discussed here). I am unaware of any way to distinguish the two on empirical grounds, so I put the matter aside; however, my structural claims about VPE size should, in principle, cross-apply straightforwardly to antecedent size should the latter be identified as the proper characterization in future work.



ration the E(llipsis)-clause and A(ntecedent)-clause are found in. Below, I first describe what is meant by “configuration” here alongside the general approach to diagnosing VPE size; then, I present a few such diagnostics.

#### 1.4.1 Methodology

To establish diagnostics of the hypothesized distinction in VPE size, we must construct examples informed by the diagnostic we already have from Merchant (2013b): that is, the distribution of VMM. In other words, assuming VMM accurately distinguishes between (at least) two different sizes of VPE, as I suggest, then we should begin the search for other diagnostics by changing as little as possible from what has already produced results with VMM. At the same time, though, if the hypothesized variability in VPE size is not simply a property of VMM itself, but in fact a more general property having to do with E-A configuration (to be made clear), then we require independent diagnostics that produce positive results even when the E- and A-clauses match in voice. In other words: if VPE can vary in size, and this variation is governed by independent factors, then it ought to be detectable even when VMM is not present in the data (see SanPietro et al. 2012). This poses somewhat of a challenge, since the extra-syntactic factors governing the (un)availability of VMM are contested in the literature, as mentioned above.

I attempt to overcome this challenge by adopting the following methodology. I start with the premise that the (un)availability of VMM reflects syntactic variability in VPE size, and that the factor(s) responsible for this variability are independent of VMM, but are otherwise unknown at present. Therefore, I compare examples whose E- and A-clauses stand in a configuration that would otherwise be compatible with VMM to those whose E-A configurations would otherwise be incompatible with VMM, even when such a mismatch is not present. The existence of pairs such as (1) vs. (4) – in which changing the E-A structural relationship from one of subordination (where VMM is possible) to one of coordination (where VMM

is bad) – allows us to test for similar fluctuations in VPE size in voice-matched examples without committing to an analysis for their ultimate cause (be it discourse coherence, information structure, etc.). To accomplish this, I appeal to exemplar environments that I take to be representative of each configuration type. These environments reflect the subordination/coordination distinction, which tracks the VMM data well.<sup>16</sup>

As a prototypical mismatch-compatible configuration, I discuss ellipsis clauses that are subordinated (or otherwise embedded) with respect to their antecedent clauses (e.g. *[John left]<sub>A</sub> even though [Mary didn't]<sub>E</sub>*; cf. (1)).<sup>17</sup> As a prototypical mismatch-incompatible configuration, I discuss unembedded ellipsis clauses that are coordinated with unembedded antecedent clauses (e.g. *[John left]<sub>A</sub> and [Mary didn't]<sub>E</sub>*; cf. (4)). I leave it to future work to establish the extent to which the conclusions we reach for these exemplar configurations apply to the full range of VMM (in)compatible configurations, and thus, the low- vs. high-VPE distinction. In the best-case scenario, the set of configurations licensing low-VPE and the set licensing high-VPE will form complementary natural classes, but this is an empirical question that must remain open for now.

The first non-VMM diagnostic we will consider is the (un)availability of strict identity when VPE takes an antecedent containing a reflexive.

#### 1.4.2 The distribution of strict identity in VPE

When the antecedent of a VPE site contains a reflexive pronoun, the interpretation of the elided material can, in certain cases, be ambiguous in its interpretation:

---

<sup>16</sup>Kehler (2002:§3.3.2) argues against this distinction as the relevant generalization for the distribution of VMM and other phenomena, citing a small class of exceptional data, e.g. grammatical examples in what appear to be coordination contexts. However, Kehler's coherence-based story is subject to its own class of exceptions, as discussed below.

<sup>17</sup>Throughout this dissertation, “subordinate clause” refers to the *central adverbial clause* of Haegeman (2012:§4.5). See *op. cit.* for definitions and discussion.

*Subordinated E-clause: “Strict” reading available (as well as “sloppy”)*

- (16) John slapped himself because Bill did.
- a. **Sloppy:** ...because Bill<sub>j</sub> slapped himself<sub>j</sub>.
  - b. **Strict:** ...because Bill<sub>j</sub> slapped him<sub>i</sub>.

The “strict” interpretation is marked, and is not as widely available as the “sloppy” interpretation (see Hestvik 1995 and references therein), though precisely characterizing the distribution of such strict readings has been a matter of debate for some time. Fiengo and May (1994) and Hestvik (1995) note that the asymmetry is sensitive to the syntactic relation connecting the E-clause to the A-clause: when the E-clause is subordinated, as in (16) above, strict identity is available; however, when the E-clause is coordinated, as in (17) below, the strict reading is degraded:<sup>18</sup>

*Coordinated E-clause: “Strict” reading unavailable*

- (17) Lea<sub>i</sub> will slap herself<sub>i</sub> today, and Jane<sub>j</sub> also will.
- a. **Sloppy:** ...and Jane<sub>j</sub> also will slap herself<sub>j</sub>.
  - b. **#Strict:** #...and Jane<sub>j</sub> also will slap her<sub>i</sub>.

Because this phenomenon tracks the same distinction that the distribution of VMM is sensitive to (namely coherence relations), it is therefore a reasonable candidate for a diagnostic of low- vs. high-VPE.

To confirm the diagnostic value of this phenomenon, it should be tested in concert with our other diagnostic, the distribution of VMM. If VMM and strict identity in VPE are both contingent upon low-VPE, then we expect each to be available whenever the other is. In other words, they should be able to co-occur within a sentence. Such sentences can be constructed, but are quite difficult to

---

<sup>18</sup>See Kehler (2002:ch. 3) for an alternative analysis based on discourse coherence theory.

judge due to their complexity; to the extent that such difficulties can be overcome, however, the judgments do seem to go the right way, at least in my own grammar:

*Strict reading available when VMM present*

- (18) The scientists should have shown the chimpanzee<sub>*i*</sub> to herself<sub>*i*</sub> because the bonobo<sub>*j*</sub> already had been.
- a. **Sloppy:** ...because the bonobo<sub>*j*</sub> already had been shown to herself<sub>*j*</sub>.
  - b. **Strict:** ...because the bonobo<sub>*j*</sub> already had been shown to her<sub>*i*</sub>.

The (un)availability of strict identity in VPE has the same distribution as our pre-existing diagnostic, suggesting the former has the same diagnostic power.

However, at this point we ought to wonder what this arises from: our story for the (un)availability of VMM was entirely structural, based on whether the E-clause VoiceSyn head was included or excluded from the VPE site. Why would the distribution of strict identity in VPE be sensitive to minor fluctuations in VPE size of this sort?

The answer can be found in recent work by Ahn (2011, In Progress), who argues that the distributional similarity of strict identity and VMM follows if reflexivity is itself a variety of grammatical voice, encoded (in part) as a feature on the head we are calling VoiceSyn<sup>0</sup> just like passive and active.<sup>19</sup> Thus, strict identity of this sort arises when the VPE site is below the ellipsis clause's VoiceSyn<sup>0</sup>, allowing the latter's features to differ from the antecedent's VoiceSyn<sup>0</sup> without violating the identity condition on VPE. In the case of strict identity, then, the E-clause's VoiceSyn feature is [-REFL] (perhaps simply [ACTIVE]), whereas the A-clause's VoiceSyn feature is [REFL].<sup>20</sup> A sketch of this is below:

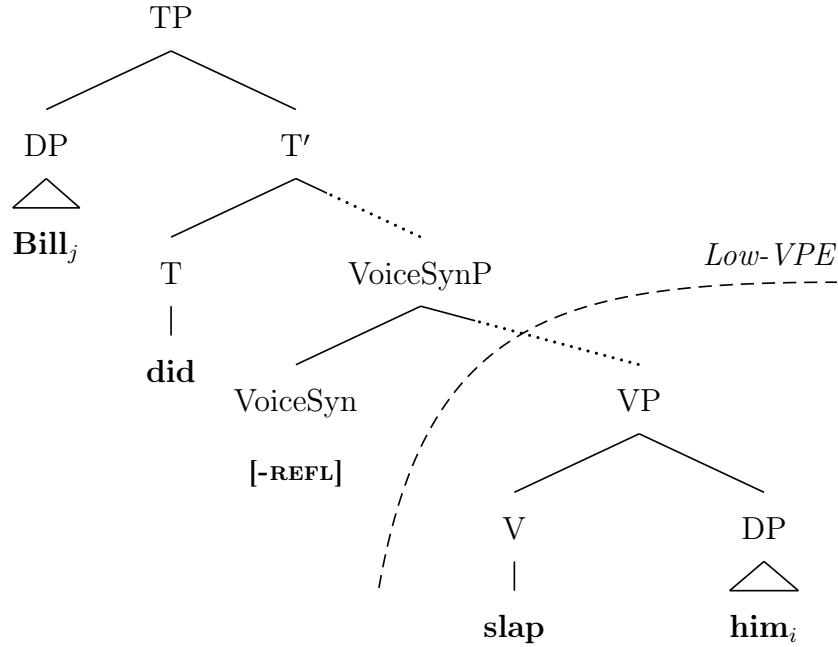
---

<sup>19</sup>See *op. cit.* for the distinction between *clausal reflexivity* (which involves VoiceSyn<sup>0</sup>) and *non-clausal reflexivity* (which does not).

<sup>20</sup>In principle, sloppy identity could arise regardless of whether VoiceSyn<sup>0</sup> is inside the ellipsis site; however, for independent reasons, Ahn (2011) argues that it in fact only arises when VoiceSyn<sup>0</sup> is elided. I leave this aside, as it is not relevant for us (but see fn. 25).

*Strict identity in low-VPE*

(19) John<sub>i</sub> [<sub>VoiceSynP</sub> **[REFL]**] slapped himself<sub>i</sub> because...



Thus, if Ahn’s analysis is correct, strict identity arises from the same configuration that gives us VMM of a more familiar sort (e.g. passive-active).

Importantly, even if it turns out that reflexivity does not involve a feature on VoiceSyn<sup>0</sup>, the fact that strict identity distributes like VMM is strongly suggestive of a structural similarity between the two with respect to VPE size (although its source would then be a mystery). In support of the claim that the similarity in question is one of VPE size, I will return to the distribution of strict identity shortly.

A summary of our findings so far is in (20):

*Summary of VPE size diagnostics (1 of 3)*

(20)

VPE size	Exemplar configuration	VMM possible?	Strict identity possible?
Low	subordination	yes	yes
High	coordination	no	no

### 1.4.3 Preferential (non-)recovery of antecedent VP-modifiers

Moulton (2008) observes that when a VPE antecedent clause contains a preverbal manner adverb,<sup>21</sup> this modifier is sometimes preferentially recovered as part of the interpretation of the elided material, and other times it is not recovered. Once again, this asymmetry tracks the coordination/subordination distinction we have already seen with respect to the E-A configuration.

Presenting results from grammaticality judgment experiments, Moulton (2008) shows that when the ellipsis clause is simply coordinated with the antecedent clause, speakers exhibit a clear preference for interpreting the elided material as modified,<sup>22</sup> that is, the interpretation of a VPE site will include any verbal modifiers from the antecedent if that VPE site arises inside a coordinated E-clause. However, when the E-clause is subordinated (or otherwise embedded, e.g. inside a relative clause), speakers exhibit a bias toward recovering unmodified elided material; that is, the interpretation of a subordinated/embedded VPE site preferentially excludes any verbal modifiers present in the antecedent (examples adapted from Moulton 2008):<sup>23</sup>

<sup>21</sup>I do not discuss post-VP adverbials in this section; see Moulton (2008) for some discussion. As Cinque (1999:§1.4) and others have suggested, post-VP adverbials (in head-initial languages) are derivationally more complex than their preverbal counterparts, involving predicate fronting, focus movement, etc. I also leave aside adverbs of other (non-manner) types.

<sup>22</sup>Moulton notes that this result, which comes from his experiment 1, is potentially influenced by the presence of *too* in the stimuli (cf. Arregui et al. 2006, Clifton and Frazier 2010:290). Ideally, this factor would be controlled for in a replication study, but this must be left to future work. However, even if we take Moulton’s (2008) experimental results tentatively, it is important to note that they directly align with native speaker intuitions (my own, as well as those of other speakers I have consulted).

<sup>23</sup>See Moulton (2008:§7) for discussion of some (apparent) exceptions. Moulton argues that pragmatic factors are ultimately responsible for patterns of adverb (non-)recovery in VPE, not the coordination vs. subordination distinction itself. Again, essentially nothing in the

*Coordinated E-clause: Modified verbal material preferentially recovered*

- (21) Jordy carefully reviewed the book, and then Kiley did.
- a. ...and then Kiley carefully reviewed the book.
  - b. #...and then Kiley reviewed the book. (Not necessarily carefully)

*Subordinated E-clause: Unmodified verbal material preferentially recovered*

- (22) Jordy carefully reviewed the book after Kiley did.
- a. #...after Kiley carefully reviewed the book.
  - b. ...after Kiley reviewed the book. (Not necessarily carefully)

This pair of configurations once again corresponds to our high- vs. low-VPE exemplar configurations: the coordination in (21) is a VMM-incompatible configuration (high-VPE), whereas the subordination in (22) is a VMM-compatible configuration (low-VPE).

I argue that the pattern in (21) vs. (22) arises directly from variation in VPE size. Indeed, this is just the sort of evidence we would expect to see if hypothesized high-VPE configurations (e.g. coordination) involved ellipsis of a larger constituent, and thus recovery of a larger constituent – in this case, one crucially including the adjunction site for adverbs such as *carefully*. Likewise, the fact that low-VPE configurations (e.g. subordination) correspond to those in which a smaller constituent is recovered – one below the aforementioned adjunction site – provides direct support for the claims made here. This phenomenon, then, not only lends itself as an additional diagnostic of VPE size, but it also grants additional insight into the fine structure of the clausal region immediately local to VoiceSyn<sup>0</sup>: this test reveals that verbal modifiers such as *carefully* must be merged

---

present proposal hinges on coordination vs. subordination being the “true” determining factor in the high- vs. low-VPE alternation; these environments were simply chosen to conveniently demonstrate the phenomenon. The conclusions here will still stand even if it turns out that pragmatic or other factors are the real source of the alternation.

at least as high as VoiceSynP. Later, we will see independent evidence that such verbal modifiers cannot adjoin lower than VoiceSyn<sup>0</sup>, directly consistent with these findings.

Additional evidence comes from an approach taken in Matsuo (2001), who argues that VPE obligatorily includes the adjunction site of manner adverbials (in examples involving coordination). She illustrates this by showing that contradictory manner adverbials cannot be present in the A-clause and the E-clause, suggesting that the A-clause adverbial is obligatorily recovered, and thus in competition with the adverbial in the E-clause. Matsuo goes on to note that this state of affairs does not apply in examples involving subordination (specifically, antecedent-contained deletion), lending additional evidence to the VPE size alternation argued for here:<sup>24</sup>

- (23) a. \*Jane carefully fixed the car, and Sue recklessly did too.  
b. Jane carefully fixed the car after Sue recklessly did.

Building on Matsuo's (2001) logic, we can conclude that this asymmetry arises because the subordinated example in (23b) involves a low-VPE site that does not include the attachment site for manner adverbs such as *carefully/recklessly*, allowing the E-clause to have its own adverb explicitly present in that position. On the other hand, high-VPE sites found in coordinated examples such as (23a) do include the adjunction site for such adverbials, meaning there is no way for *recklessly* to survive ellipsis in the E-clause (except by way of focus and/or movement, but see fn. 24).

Before moving on, an additional prediction arising from the adverbial recovery diagnostic must be tested: if e.g. preferential recovery of unmodified material is

---

<sup>24</sup>Note that such examples may improve with focal stress on the adverbials (and omission of *too*); however, such focus has been argued by Cinque (1999) and others to involve movement of the adverbials in question, thus undermining their diagnostic value for our purposes. See also fn. 21.



a valid diagnostic of the same low-VPE operation we discovered using VMM, then the two should pattern together in the relevant contexts. In other words, we predict that whenever VMM is possible with an A-clause containing a verbal modifier, that modifier should not be interpreted as part of the elided material: VMM requires low-VPE, but a low-VPE site is not big enough to include the locus of verbal adjunction. This prediction is confirmed:

*Unmodified verbal material preferentially recovered when VMM present*

- (24) The janitor should carefully clean the room whenever it is apparent that it should be.
- a. #...whenever it is apparent that it should be carefully cleaned.
  - b. ...whenever it is apparent that it should be cleaned. (Not necessarily carefully)

Likewise, if we continue the trend of good housekeeping in reconciling our VPE size diagnostics, then the presence of a strict identity reading, being contingent upon low-VPE, should thus entail recovery of an unmodified verbal antecedent in the relevant contexts (discussed in the last subsection). This is also confirmed:<sup>25</sup>

*Unmodified verbal material preferentially recovered when strict identity present*

- (25) Lea<sub>i</sub> gently slapped herself<sub>i</sub> after Jane<sub>j</sub> did.
- a. #...after Jane<sub>j</sub> gently slapped her<sub>i</sub>.
  - b. ...after Jane<sub>j</sub> slapped her<sub>i</sub>. (Not necessarily gently)

In sum, these patterns of (non-)recovery of verbal modifiers support the present

---

<sup>25</sup>Although sloppy readings are largely irrelevant here, I will mention that, to my ear, a sloppy reading of (25) seems to prefer recovery of the modifier, i.e. *Lea<sub>i</sub> gently slapped herself<sub>i</sub> after Jane<sub>j</sub> #(gently) slapped herself<sub>j</sub>*. While this does not follow from anything proposed here, it is consistent with Ahn (2011), who argues on other grounds that sloppy identity actually requires what I have identified as high-VPE. It seems that high-VPE is once again found to include the locus of adjunction for manner adverbs, consistent with the claim I make that this adjunction site is no lower than VoiceSynP.

proposal: differences in E-A configurations trigger principled variation in VPE size.  
 Updating our summary to reflect this:

*Summary of VPE size diagnostics (2 of 3)*

(26)

VPE size	Exemplar config.	VMM possible?	Strict ID possible?	Adverbs excluded?
Low	subord.	yes	yes	yes
High	coord.	no	no	no

I turn now to the final diagnostic I will discuss, which involves the (un)elidability of certain aspectual auxiliaries.

#### 1.4.4 Aspectual mismatch and auxiliary omission

In the previous subsection, we found that subtle variations in the size of VPE could be revealed by looking at interpretational patterns in ellipsis recovery. Below, I extend this diagnostic approach to the (non-)recovery of various auxiliary verbs from the A-clause. As we will see, high-VPE configurations allow omission and recovery of particular auxiliaries, while low-VPE configurations do not.

Recall from §1.3.2.1 that certain data involving the apparent omission of auxiliaries under VPE have led some to the position that VPE is inherently variable in its size. This was exemplified previously as (12), repeated below:

- (12) John should have been studying, but
- a. ...Mary shouldn't have been.
  - b. ...Mary shouldn't have.
  - c. ...Mary shouldn't.

We now return to these data, with a particular emphasis on the interpretations available to them. In the examples throughout this subsection, ~~struck-through~~ material indicates a particular interpretation for an ellipsis site. If an example containing such material is marked as infelicitous ( $\#$ ), this indicates that the interpretation corresponding to the struck-through material is unavailable for that E-clause. In some cases, a different interpretation may be available, i.e. one that does not correspond exactly to the struck-through material. For example, consider (12c). Looking closely, its ellipsis site (represented with the ellipse symbol,  $\circ$ ) can, with varying levels of acceptability subject to dialectal variation,<sup>26</sup> receive any of the following interpretations, which vary solely on their aspectual content:

- (27) John should have been studying, but Mary shouldn't  $\circ$ .
- a.  $\circ$  = study *Aspect mismatch (no perf, no prog)*
  - b.  $\circ$  = be studying *Aspect mismatch (no perf)*
  - c.  $\circ$  = have been studying *Matching aspect*

If we set aside interpretation for a moment and simply compare the E-clause surface string to the A-clause surface string in (27), then, superficially, we might deduce that the VPE site corresponds to everything below the A-clause's modal, i.e. *have been studying*. However, the available interpretations for this VPE site reveal that this is not the only possibility.

The contents of the interpretations in (27a) and (27b) do not match the aspectual profile of the A-clause, which is perfect progressive; instead, they each bear a different aspectual profile, meaning such interpretations involve an aspectual mismatch between E-clause and A-clause. Crucially for us, it stands to reason that

---

<sup>26</sup>The decidedly dialectal nature of this variability (and that in (12)) first came to my attention during my initial work on VPE in tag questions (Sailor 2009). I discuss this further in §3.4. I thank Tim Stowell and Carson Schütze, speakers of Canadian English, for their vocal disagreements with my own American English judgments during the early stages of that project, leading me to dig deeper into the matter.

if a particular aspect is not part of the interpretation of a VPE site, then all of the morphosyntactic exponents of that aspect – e.g. its auxiliary, its morphology – were never present inside the VPE site to begin with (Sailor 2009:29). In other words, although some of the A-clause auxiliaries are missing in the E-clause in (27), they have not necessarily been elided: in the case of an aspectual mismatch interpretation like those in (27a) and (27b), they were simply absent from the E-clause numeration from the very start.

This poses a potentially serious confound for the use of auxiliary omission as a diagnostic of VPE size – something I am openly promoting in this subsection. We therefore require a strategy for eliminating the possibility of aspectual mismatches in the data. Employing such a strategy would allow us to conclude with confidence that any A-clause auxiliaries missing from the E-clause are indeed part of the E-clause’s numeration (and thus its interpretation), but they have been eliminated from the surface representation by VPE.

Aspectually-fixed idioms provide just such a strategy. To rule out the possibility of an aspectual mismatch between the E- and A-clauses, I appeal to idioms that necessarily include a particular aspect as part of their complex lexical entry, such as *dying to X*, which is obligatorily progressive,<sup>27</sup> and *have been to X*, which is obligatorily perfect.<sup>28</sup> The following examples illustrate the aspectually-fixed nature of each expression. Note that they can appear embedded under various tenses and more than one aspect; the only requirement is that the aspect that has been lexicalized as part of the idiom (progressive for *dying to X*, and perfect for *have been to X*) be present in the clause.<sup>29</sup> If this lexicalized aspect is not present, then the (partial) idiom string is either ungrammatical or can only receive a literal

---

<sup>27</sup>Other obligatorily-progressive idioms include *cruisin’ for a bruisin’*, *champing/chomping at the bit*, *pushing up daisies*, etc.

<sup>28</sup>See Harwood (2013a) and Aelbrecht and Harwood (2013) for similar methodology involving idioms (but with different results). Other obligatorily-perfect idioms include *have been around the block*, *have had it*, etc.

<sup>29</sup>However, *have been to X* cannot be put into the progressive (e.g. *\*is having been to X*) for independent reasons: it is a stative/individual-level predicate.

interpretation (indicated with #):

*An obligatorily-progressive idiom*

- (28) a. John {was / has been / will be} dying to go to that new leather bar.  
b. #John died to go to that new leather bar.  
c. #John has died to go to that new leather bar.

*An obligatorily-perfect idiom*

- (29) a. John {had / will have} been to that new leather bar.  
b. \*John was to that new leather bar.

Since these idioms require particular aspects, using them as antecedents for VPE will allow us to control for the aspectual mismatch problem seen above in (27). If an A-clause involving one of these idioms contains some auxiliary that is missing from the E-clause, then we can be confident that auxiliary has been elided if the VPE site is licit (i.e., it is not \* or #). With this groundwork laid, we can undertake the matter of diagnosing VPE size using patterns of auxiliary omission.

First, consider omission of the variety of BE that takes as its complement the progressive participle (see fn. 7 on the notations BE,  $X^0_{BE}$ , etc.). Under the model of the inflectional system I have adopted here (essentially that of Bjorkman 2011), this instance of BE occupies  $Prog^0$ , which is just above  $VoiceInff^0$ . Constructing examples using the obligatorily-progressive idiom *dying to X* (and filling  $T^0$  with a modal to prevent raising of  $Prog^0_{BE}$ ), we see that  $Prog^0_{BE}$  can be omitted in a typical coordinate structure with VPE:

*Coordinated E-clause: omission of  $Prog^0_{BE}$  possible*

- (30) a. John will be dying to go, and Mary will be ~~{dying to go}~~ too.  
b. John will be dying to go, and Mary will ~~{be dying to go}~~ too.

In the (b) example above, the E-clause lacks an overt  $Prog^0_{BE}$  where one is present in the A-clause. Importantly, though, this E-clause retains the idiomatic inter-

pretation of *dying to X*, indicating that progressive aspect is indeed part of the E-clause’s semantic content (and  $\text{Prog}_{\text{BE}}^0$  part of its numeration), meaning there is no aspectual mismatch between the A- and E-clauses here. The only conclusion to draw is that VPE has elided  $\text{Prog}_{\text{BE}}^0$  along with the progressive participle.

However, the pattern changes when the E-A configuration changes from coordination to subordination. Subordinated E-clauses cannot retain the idiomatic interpretation for *dying to X* if  $\text{Prog}_{\text{BE}}^0$  is omitted:

*Subordinated E-clause: omission of  $\text{Prog}_{\text{BE}}^0$  illicit*

- (31) a. John will be dying to go for the same reason Mary will be ~~be dying to go~~.
- b. #John will be dying to go for the same reason Mary will ~~be dying to go~~.

While some interpretation for the E-clause in (31b) might be available (e.g. ...*for the same reason Mary will go*), crucially the idiomatic progressive interpretation *dying to go* is not available, indicating that  $\text{Prog}_{\text{BE}}^0$  cannot be recovered in this subordinated configuration. Recall that subordination configurations are generally compatible with VMM, and thus by hypothesis are low-VPE configurations. On these grounds, it is logical to conclude from the above pattern that subordination is not blocking the recovery of  $\text{Prog}_{\text{BE}}^0$  in (31b), *per se*; rather, it is blocking ellipsis of a structure as large as ProgP, meaning  $\text{Prog}_{\text{BE}}^0$  cannot be recovered simply because it cannot be elided in the first place in this configuration. Thus, patterns in the (un)elidability of  $\text{Prog}_{\text{BE}}^0$  track the VPE size distinction we have seen evidence for already, indicating that this can be taken as yet another diagnostic of the phenomenon. Example (30b) also reveals that high-VPE can elide a structure at least as large as ProgP, which was not evident from our earlier diagnostics. We will return to this later.

Similarly, an epistemic reading of *must* can be made to force a progressive interpretation, since such a reading requires that *must* takes a stative complement.<sup>30</sup> Predicates which are otherwise eventive nevertheless get a stative interpretation when they appear in the progressive aspect (or habitual, etc.). Without the stative interpretation, though, the epistemic reading for *must* is unavailable:

*Epistemic ‘must’ requires a stative complement (e.g. a progressive)*

- (32) a. John must be getting a coffee. ✓ *Epistemic*  
 b. John must get a coffee ~~#~~(every morning). ~~#~~ *Epistemic*

This property can be exploited to diagnose whether the progressive is present in the E-clause when  $\text{Prog}_{\text{BE}}^0$  is omitted under VPE: if it is, then *must* will allow an epistemic reading; if it is not, then an epistemic reading for *must* will be unavailable. With this in mind, when we turn to the behavior of epistemic *must* in the proximity of VPE, we see the same coordination/subordination split arise:

*Coordinated E-clause: omission of  $\text{Prog}_{\text{BE}}^0$  possible*

- (33) [Context: “John and Mary aren’t in their offices. Given that their coffee cups are gone...”]  
 a. John must be getting a coffee, and Mary must be ~~{getting a coffee}~~ as well.  
 b. John must be getting a coffee, and Mary must ~~{be getting a coffee}~~ as well.

*Subordinated E-clause: omission of  $\text{Prog}_{\text{BE}}^0$  illicit*

- (34) [Context: “John and Mary aren’t in their offices. Given that Mary’s coffee cup is gone, and given that they do everything together...”]  
 a. John must be getting a coffee, because (it appears that) Mary must

---

<sup>30</sup>Thanks to Keir Moulton for bringing this to my attention as a potential diagnostic.



be ~~{getting a coffee}~~ as well.

- b. #John must be getting a coffee, because (it appears that) Mary must ~~{be getting a coffee}~~ as well.

Thus, epistemic readings of *must* are on par with idioms such as *dying to X* in their ability to diagnose the (un)availability of a progressive reading in an E-clause.

The differential behavior of auxiliary recovery in coordinated vs. subordinated E-clauses continues in examples involving omission of the perfect auxiliary *have*. Use of the obligatorily-perfect idiom *have been to X* ensures that an aspectual mismatch will not confound the results. It should be noted from the outset that omission of perfect *have* in VPE has been reported to be ungrammatical in various places in the literature, including early on in Akmajian and Wasow (1975) and Sag (1976). This sentiment has been repeated recently in my own work (Sailor 2012a) and, in particular, in Aelbrecht and Harwood (2013:§3.2), although they acknowledge the presence of dialectal variation in this domain, with some speakers accepting omission of perfect *have* under VPE even when aspectual mismatch is controlled for using the aforementioned idiom.<sup>31</sup> I discuss this variation in more detail in §3.4. For now, I will indicate the variable judgments with the appropriate grammaticality mark, %. Crucially for us, though, the only examples involving omission of *have* that have been reported as acceptable in the literature (by Aelbrecht and Harwood 2013, Thoms 2010, and my own consultants) are those involving coordination of the E-clause and A-clause, as in the following:<sup>32</sup>

---

<sup>31</sup>Bošković (2013: fn. 47), citing a manuscript *in preparation* by Susi Wurmbrand, even acknowledges the analytical challenge presented by aspectual mismatches, suggesting that once such mismatches are controlled for, omission of perfect *have* is consistently ungrammatical. While I of course agree that aspectual mismatch must be factored out, it seems clear that the judgments reported by Bošković (via Wurmbrand) do not reflect the full spectrum of possibilities available to different varieties of English. It is critical that we probe more than one variety of English in this domain; see §3.4 for an initial attempt. As I show here, though, omission of perfect *have* under VPE is absolutely possible for many speakers.

<sup>32</sup>These examples happen to involve the so-called epistemic necessity reading of *will*, which is widely available in varieties of British English, but is more constrained in American varieties.

*Coordinated E-clause: omission of Perf<sup>0</sup><sub>HAVE</sub> possible (dialectally)*

- (35) a. (I'm betting that) John will have been to Paris, and his wife will have ~~been to Paris~~ as well.
- b. %John will have been to Paris, and his wife will ~~have been to Paris~~ as well.

Importantly, subordinating the E-clause renders omission of *have* impossible, even for those who accept example (35b), above. This is shown below in (36b):

*Subordinated E-clause: omission of Perf<sup>0</sup><sub>HAVE</sub> illicit*

- (36) a. (I'm betting that) John will have been to Paris as many times as his wife will have ~~been to Paris~~.

- b. \*John will have been to Paris as many times as his wife will ~~have been to Paris~~.

Similar patterns arise using other methods for controlling aspectual interpretation. For example, as noted in Sailor (2012a:20), certain contexts involving temporal adjuncts can force a perfect reading:

*Contexts involving temporal adjuncts requiring perfect aspect*

- (37) a. I could have studied harder for the exam before taking it yesterday.  
 b. #I could study harder for the exam before taking it yesterday.

In the familiar way, we can exploit this phenomenon for its diagnostic power in VPE. Doing so, the differential behavior of coordinated E-clauses as compared to their subordinated counterparts emerges once again:

*Coordinated E-clause: omission of Perf<sup>0</sup><sub>HAVE</sub> possible (dialectally)*

- (38) a. Mary could have studied harder for the exam before taking it yesterday, and Bill could have ~~studied harder~~ as well.  
 b. %Mary could have studied harder for the exam before taking it yesterday, and Bill could ~~have studied harder~~ as well.

*Subordinated E-clause: omission of Perf<sup>0</sup><sub>HAVE</sub> illicit*

- (39) a. Mary could have studied harder for the exam before taking it yesterday, just like Bill could have ~~studied harder~~.  
 b. \*Mary could have studied harder for the exam before taking it yesterday, just like Bill could ~~have studied harder~~.

The preceding data and discussion reveal that non-finite forms of *be* selecting progressive participles, as well as non-finite forms of *have*, are able to be omitted

in some VPE contexts, but not others. This tracks the distinction in ellipsis-antecedent configurations correlating to high-VPE vs. low-VPE. Like the adverb recovery data from the previous subsection, these findings allow us to further refine the size of high-VPE: it is able to include at least the projection headed by *be* that selects progressive participles (see also Aelbrecht and Harwood 2013), and for some speakers, it can even include the projection headed by *have* in the perfect. In other words, we can conclude that the maximal domain for high-VPE is the complement of  $T^0$ , i.e. the VPE domain originally assumed by Lobeck (1995) and others under a much less fine-grained view of the verbal-inflectional domain. We can conclude less about the size of low-VPE from these examples, except that it cannot even elide the lower of the two projections discussed here. We return to the size of low-VPE shortly.

For completeness, as we did before, we must reconcile these results with those from earlier in this section, e.g. VMM. In brief, the prediction is this: if VMM is present, then we are dealing with a low-VPE structure, meaning auxiliary omission of the sort we have just seen should be unavailable, requiring high-VPE. Unfortunately, neither of the fixed-aspect idioms introduced in this subsection tolerate passivization,<sup>33</sup> meaning it will be difficult to rule out the possibility of an aspectual mismatch; similar problems with VMM arise with the epistemic *must* test, as well. However, we can use the temporal adverb contexts seen above to show that at least  $\text{Perf}_{\text{HAVE}}^0$  cannot go missing when VMM is present.<sup>34</sup>

---

<sup>33</sup>This is surely no accident: grammatical voice is encoded on  $\text{VoiceSyn}^0$ , which is merged lower in the inflectional hierarchy than the various aspectual heads. If idioms represent stretches of lexicalized structure, as is widely held, then it follows that an idiom with fixed features on a given aspectual head (e.g.  $\text{Prog}^0$ ) will also have fixed features on inflectional heads lower down. See Sailor and Ahn (2010:§4) for discussion.

<sup>34</sup>For reasons that will become clear shortly, VMM of the passive-active type is necessary here. Similar problems arise in attempts to blend this diagnostic with e.g. the strict identity diagnostic; however, it should be noted that omission of auxiliaries (e.g.  $\text{Perf}_{\text{HAVE}}^0$ ) whose surface position in the antecedent is above a verbal modifier correlates with preferential recovery of such modifiers, e.g.:

- (i) %[(I'm betting that) John will have secretly been to Paris, but his wife won't.]
  - a. ...his wife won't have secretly been to Paris.
  - b. #...his wife won't have been to Paris. (Not necessarily secretly)

*Omission of Perf<sub>HAVE</sub> unavailable when VMM present*

- (40) a. The trash shouldn't have been emptied then, even though John thinks by that time someone should have ~~emptied it~~.  
 b. #The trash shouldn't have been emptied then, even though John thinks by that time someone should ~~have emptied it~~.

A summary of our findings from this subsection is in (41):

*Summary of VPE size diagnostics (3 of 3)*

(41)

VPE size	Exemplar config.	VMM possible?	Strict ID possible?	Adverbs excluded?	have/BE <sub>prog</sub> intact?
Low	subord.	yes	yes	yes	yes
High	coord.	no	no	no	no

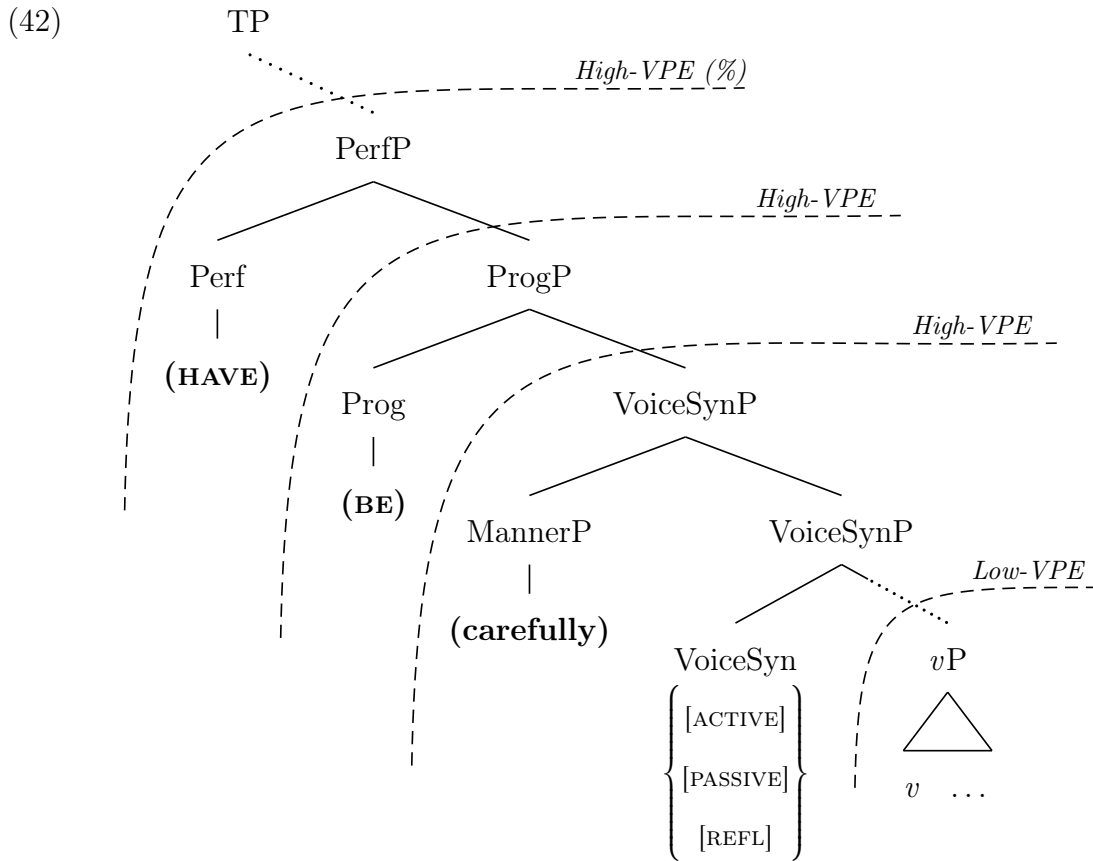
#### 1.4.5 Summary and consequences

Merchant (2013b) convincingly argues that the patterned (un)availability of VMM in sluicing versus VPE should be taken as the result of differences in ellipsis size between the two operations. In this section, I pushed this logic further: rather than just diagnosing differences in ellipsis size between TPE vs. VPE, I argued that VMM can be used to diagnose differences in ellipsis size within VPE itself.

To VMM, I added strict identity of reflexives, adverbial recovery, and ellipsis of auxiliaries as phenomena that double as diagnostics of VPE size. These led us to a novel distinction within VPE: one class of ellipsis-antecedent configurations forces ellipsis to be low, while another allows high-VPE. A summary of these results is below, with the low-VPE ellipsis size corresponding to the subordination contexts discussed, and the various high-VPE possibilities corresponding to the coordination contexts:<sup>35</sup>

<sup>35</sup>As before, dotted branches indicate additional structure that has been omitted for clarity. In

Summary of VPE sizes by diagnostic results



Despite these findings, the preceding discussion offered only limited insight into the category of the XP(s) elided in low-VPE. In the next chapter, I probe the structure of low-VPE in greater detail in order to determine precisely what category is targeted by this type of VPE. We will see that low-VPE does not fluctuate in size; instead, it consistently targets the same maximal projection (despite superficial evidence to the contrary). I return to high-VPE in chapter 3.

---

particular, we will see evidence in chapter 2 for another voice-related projection in between VoiceSynP and *v*P (namely VoiceInflP). Here I depict the lowest position of manner adverbials as occupying a “MannerP” adjoined to VoiceSynP; however, nothing crucially relies on this traditional adjunction structure over an approach like that of Cinque (1999).

## CHAPTER 2

### On the Fine Structure of Low-VPE

#### 2.1 Introduction

In the previous chapter, we saw evidence that VPE could affect varying sizes of structure depending on the nature of the configuration relating the E-clause to the A-clause. This variability was presented as fundamentally binary in nature—low-VPE vs. high-VPE, a purely structural distinction that hinges entirely on the status of the syntactic head bearing the clause’s grammatical voice features. We saw that this distinction could be brought out using evidence independent of matters concerning grammatical voice (i.e. independent of VMM); nevertheless, the deciding factor driving the distinction is still the head responsible for the syntax of voice: namely, whether  $\text{VoiceSyn}^0$  is included or excluded from the VPE site. This implicates that any application of VPE that includes  $\text{VoiceSynP}$  constitutes high-VPE, even if the ellipsis site is actually higher than  $\text{VoiceSynP}$  (e.g.  $\text{ProgP}$ ). We will return to this later.

Within this other evidence we saw circumstances in which high-VPE elided a structure large enough to contain manner adverbs and auxiliaries associated with progressive and perfect aspect. Crucially, ellipsis of such large structures was shown to be inherently incompatible with VMM, which follows if Merchant’s (2013b) proposal is correct – that VMM can be ruled out by the syntactic identity requirement on VPE – and the syntactic head whose identity is so critical,  $\text{VoiceSyn}^0$ , is lower in the structure than the positions associated with the aspects

and modifiers in question (as Cinque 1999 and others have shown independently).

Having seen that VPE can be smaller or larger depending on context, and that this variation is principled, we ought to ask what precise categories are targeted in VPE of each size. That is, what are VPE's structural upper and lower bounds?

The focus of the present chapter is on determining its lower bound. Given that VMM is possible in low-VPE environments, the maximal constituent targeted by low-VPE must always be below the head responsible for the syntax of grammatical voice. I will discuss this in detail below, showing that close scrutiny of low-VPE allows us to diagnose the precise constituent that is elided, which in turn will lead us to a more detailed picture of the structure that makes up this region of the clause.

## 2.2 *Being* being stranded where it shouldn't be (\*being)

Recall from §1.3.2 that some analyses of VPE have assumed that its domain of application is constrained to the projection headed by the surface position of the main verb, i.e. *vP*. However, a simple empirical point poses a challenge for this assumption.

If VPE always targets nothing higher than the projection hosting the main verb, namely *vP*, then the inflectional auxiliary *being* that arises in progressive passive clauses is incorrectly predicted to survive VPE in all circumstances. On the contrary, this auxiliary cannot ever be stranded outside of a VPE site, in either low- or high-VPE – it is obligatorily included among the elided material:<sup>1</sup>

---

<sup>1</sup>This observation can be traced back at least as early as Akmajian and Wasow (1975:222); see Johnson (2001: fn. 6) for a list of relevant references. It has been reported that *being* can be stranded adjacent to an ellipsis site if it is immediately preceded by a contrastively focused element. Such examples are robustly ungrammatical for both the American and British English speakers I have consulted (reflecting my own judgments), so I will not consider them here. See Aelbrecht and Harwood (2013:§7) for some discussion.



*Obligatory ellipsis of 'being' in both low- and high-VPE*

- (1) a. \*He should be being criticized whenever she should be being ~~criticized~~.  
b. He should be being criticized whenever she should be ~~being criticized~~.  
c. \*He should be being criticized, and she should also be being ~~criticized~~.  
d. He should be being criticized, and she should also be ~~being criticized~~.

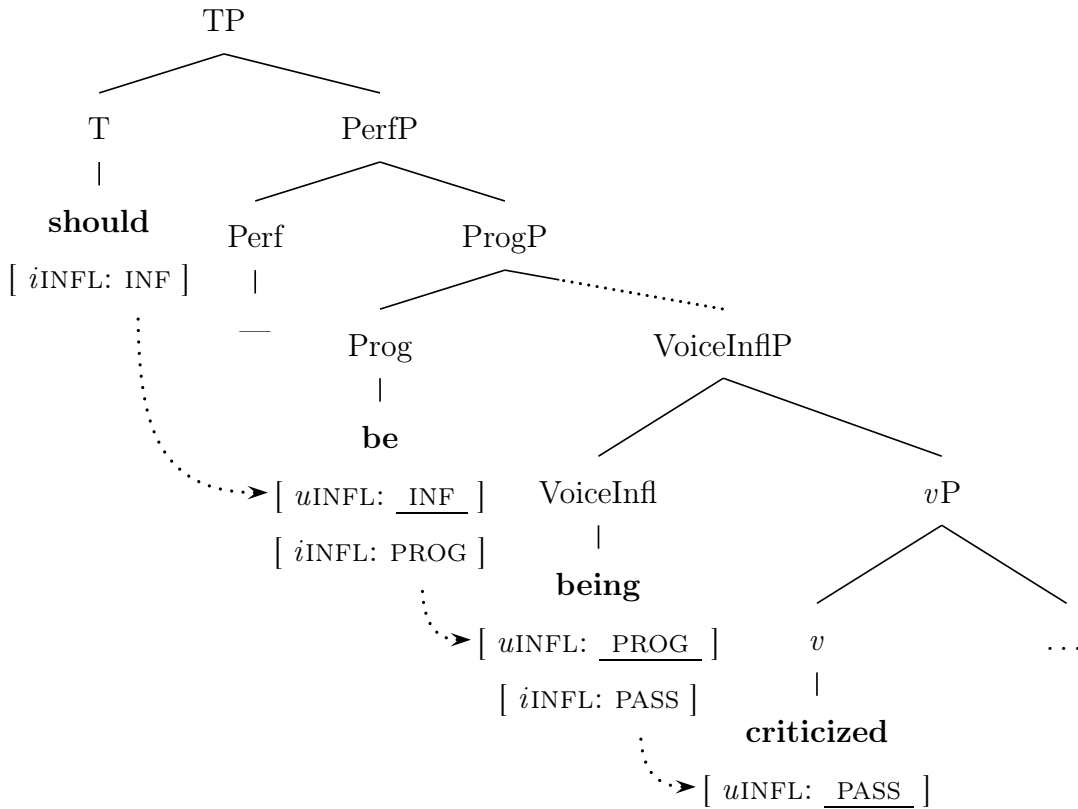
Based on this observation alone, it appears that the constituent elided by VPE is necessarily bigger than VP / *v*P, under the standard analysis in which the English main verb always comes to occupy *v*<sup>0</sup> by the derivation's end (with *being* located somewhere higher). By this reasoning, it must be that VPE elides a constituent that is at least as high in the structure as the surface position of *being* in progressive passive clauses.<sup>2</sup> Recall from §1.3.1 the structure that we are assuming for the English INFL-domain (repeated below):

---

<sup>2</sup>While this is a simple conclusion to draw, it is not at all trivial: to the extent that they discuss *being* at all, prior analyses of VPE have simply had to stipulate that it cannot be stranded adjacent to the ellipsis site (cf. Sag 1976:27, who broadens the constraint to *-ing* forms more generally; but, see also Thoms 2010 and Aelbrecht and Harwood 2013).

*Inflection via Agree*

(2) ...should be being criticized



In this sentence, the main verb occupies  $v^0$ , as is standard; *being* occupies a voice-related inflectional head we have been calling  $\text{VoiceInfl}^0$ ; and, *be* (in progressive passive clauses such as (1)) occupies  $\text{Prog}^0$ , the next highest inflectional head.<sup>3</sup> Since *being* is obligatorily elided but *be* is left intact, this suggests that the elided constituent in (1b,d) is VoiceInflP. However, recall the lesson from the previous chapter: in low-VPE configurations such as (1b), the head bearing the clause's grammatical voice features is crucially outside the ellipsis site. Putting these conclusions together motivates a novel refinement of the verbal-inflectional domain involving two discrete heads associated with voice, as we will now see.

<sup>3</sup>As we will see, the order of heads in this domain is actually  $\text{Prog}^0 > \text{VoiceSyn}^0 > \text{VoiceInfl}^0$ ; but, crucially,  $\text{VoiceSyn}^0$  is not an inflectional head. See §2.3.

## 2.3 Choices of Voice<sup>0</sup>s

In the literature, a syntactic head labeled “Voice” has been employed for two distinct (but related) purposes. One usage – that of Kratzer (1996), Sailor and Ahn (2010), Merchant (2013b), and many others – has seen Voice moderating the syntax and semantics of the predicate’s external argument: it is the external theta-role assigner, and argument-structure alternations implicating the external argument, e.g. passive and middle voice, are encoded in Voice’s featural matrix. The other usage of Voice in the literature is rather different: in e.g. (Cinque 1999, Bjorkman 2011), and others, Voice is used to introduce or host inflectional morphology (or inflectional features) associated with grammatical voice. No attempt has been made in the literature to unite these distinct implementations of Voice, nor is it obvious that these properties can (or should) be reduced to the influence of a single syntactic head.

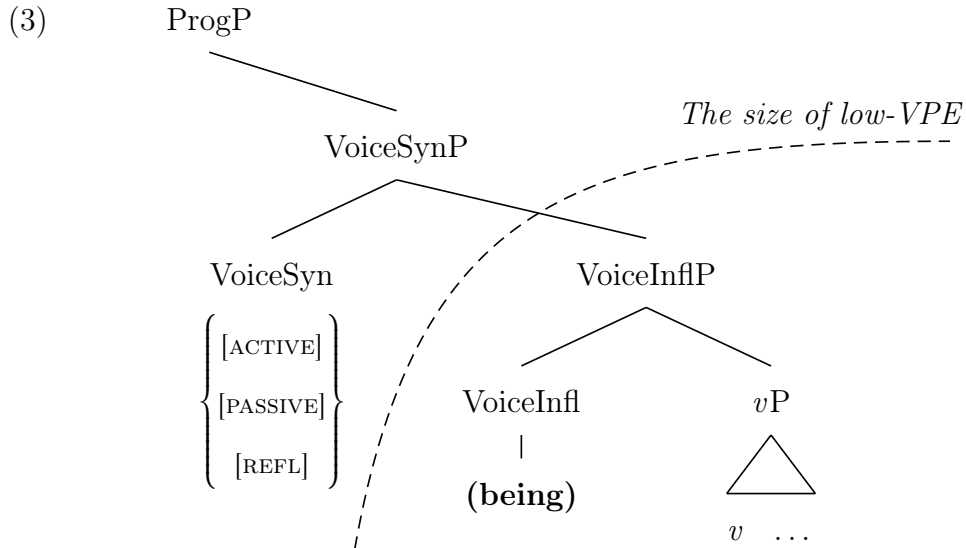
Assuming a model of the inflectional domain that involves a hierarchy of functional projections associated with the relevant morphology (and/or the attendant auxiliaries in a language like English), as in Tenny (1987) and Cinque (1999), the head associated with voice-related inflectional morphology behaves very much like the higher inflectional heads associated with e.g. aspect; that is, in English, passive morphology undergoes the same “affix-hopping” effect as other inflectional morphemes in the language (regardless of whether lowering is actually the proper analysis for this phenomenon): see Chomsky (1957). To my knowledge, none of these higher inflectional heads is capable of introducing arguments or otherwise affecting (or effecting) argument structure. At the same time, I am not aware of any attempts to explicitly collapse these distinct voice-related functions – one inflectional, one syntactic – into a single head.

The conclusion from the previous subsection has shown that this matter can be resolved on empirical grounds: what the data demand are distinct (but apparently

adjacent) functional projections associated with voice. The choice of labels for these heads from the preceding discussion has anticipated this conclusion, but now we have seen direct empirical evidence for it: the head controlling the syntax of voice (VoiceSyn<sup>0</sup>) must be distinct from, and just higher than, the head associated with the inflectional morphology of voice (i.e. the one hosting *being*: VoiceInfl<sup>0</sup>). In other words, we derive the order ProgP > VoiceSynP > VoiceInflP > vP. Stating the conclusion differently, we saw previously that the projection associated with voice syntax, VoiceSynP, must be properly excluded from low-VPE in order to capture the availability of VMM; now, we see that a voice-related projection associated with inflection, VoiceInflP, must be properly included in low-VPE in order to capture the obligatory ellipsis of *being*.

This finer-grained picture of low-VPE within the verbal-inflectional domain is schematized below. The VoiceSyn<sup>0</sup> head selects VoiceInflP, and low-VPE cuts right in between them:

*Voice shells and the size of low-VPE*



To summarize, the evidence suggests that the distinct functions of voice – one inflectional, one structural – must correspond to distinct (albeit adjacent) heads,

reminiscent of the shell-structure assumed for the derivation of lexical categories. I propose that the head Bjorkman (2011) identifies as Voice<sup>0</sup> – the one occupied by *being* in passive progressives – is not the same head that Merchant (2013, following others) refers to as Voice<sup>0</sup>. Bjorkman’s (2011) Voice<sup>0</sup> is responsible for the inflection associated with grammatical voice (i.e. passive morphology, the other voices lacking morphological exponence in English), whereas Merchant’s (2013b) Voice<sup>0</sup> is responsible for the syntax associated with grammatical voice (Sailor and Ahn 2010).<sup>4</sup>

I explore some consequences of this analysis in the next subsection.

### 2.3.1 Raising your Voice: On apparent variation in low-VPE size

The present proposal predicts that any auxiliary in VoiceInfl<sup>0</sup> must undergo ellipsis. This holds absolutely when VoiceInfl<sup>0</sup> is realized as *being* (i.e. in progressive passives); however, *being* is not the only morphological realization of VoiceInfl<sup>0</sup>: it can also be realized as *be* (in modal passives) and *been* (in perfect passives): see §1.3.1. We therefore expect low-VPE to elide these forms of VoiceInfl<sup>0</sup>, as well. Going to the data, it appears that VPE can elide these other forms (4), but that it need not (5), which poses a problem:

*Predicted: low-VPE can elide non-‘being’ forms of VoiceInfl<sup>0</sup>*

- (4) a. You think this problem can’t be solved, even though everyone else thinks it can [~~be solved~~].
- b. This problem hasn’t been solved because the other one has [~~been solved~~].

*Not predicted: low-VPE need not elide non-‘being’ forms of VoiceInfl<sup>0</sup>*

- (5) a. You think this problem can’t be solved, even though everyone else

---

<sup>4</sup>It is unclear whether VoiceSyn<sup>0</sup> ever has morphological exponence in English, though its specifier is presumably implicated in the derivation of e.g. passive syntax: see Sailor and Ahn (2010).

thinks it can be ~~solved~~.

- b. This problem hasn't been solved because the other one has been ~~solved~~.

Observing the asymmetry in elliptical behavior between *being* (as in (1)) and other forms of BE arising in passives (as here), Akmajian and Wasow (1975:223) note that "...passive BE is part of the VP [i.e., it elides with the predicate, -CS] just in case progressive BE is present". Another way of thinking about the above examples is that they appear to show that VPE can optionally target a constituent below VoiceInflP (say, *vP*), contrary to the generalization seen up to this point. Framed as the optional ellipsis of auxiliaries, this is a known problem in the literature, and has long resisted analysis (see e.g. Akmajian and Wasow 1975, Johnson 2001, Thoms 2010, and Aelbrecht and Harwood 2013 for some attempts). For completeness, below I present a full paradigm of BE-omission patterns in passive clauses under low-VPE (see §1.4.4 on omission of HAVE):<sup>5</sup>

*Auxiliary omission in passive clauses with low-VPE*

- (6) Modal + passive: VoiceInfl<sup>0</sup><sub>be</sub> = [ uINFL: INF ]
- a. Sue will be promoted whenever John will be ~~promoted~~.
- b. Sue will be promoted whenever John will ~~be promoted~~.
- (7) Perfect + passive: VoiceInfl<sup>0</sup><sub>been</sub> = [ uINFL: PERF ]
- a. Sue has been promoted because John has been ~~promoted~~.
- b. Sue has been promoted because John has ~~been promoted~~.
- (8) Modal + perfect + passive: VoiceInfl<sup>0</sup><sub>been</sub> = [ uINFL: PERF ]

---

<sup>5</sup>As I note in various places in this dissertation, but in particular in §3.4, there seems to be subtle dialectal preferences for larger or smaller VPE sizes in certain cases. The judgments here reflect standard American English. Importantly, there is no detectable difference in meaning between omission vs. non-omission of VoiceInfl<sup>0</sup><sub>BE</sub> in any of these examples. Recall that certain aspectually-mismatched readings may intrude (indicated with #), giving the illusion of acceptability of a fully-matched reading (i.e. of the struck-through interpretation) where none actually exists; see §1.4.4 for discussion and controls.

- a. Sue could have been promoted if John could have been ~~promoted~~.
  - b. Sue could have been promoted if John could have ~~been promoted~~.
- (9) Progressive + passive: VoiceInfl<sup>0</sup><sub>being</sub> = [ uINFL: PROG ]
- a. \*Sue was being criticized before John was being ~~criticized~~.
  - b. Sue was being criticized before John was ~~being criticized~~.

- (10) Perfect + progressive + passive:  $\text{VoiceInfl}^0_{\text{being}} = [ \text{uINFL: } \underline{\text{PROG}} ]$
- a. \*Sue had been being criticized before John had been being ~~criticized~~.
  - b. Sue had been being criticized before John had been ~~being criticized~~.
  - c. #Sue had been being criticized before John had ~~been being criticized~~.
- (11) Modal + perf. + prog. + passive:  $\text{VoiceInfl}^0_{\text{being}} = [ \text{uINFL: } \underline{\text{PROG}} ]$
- a. \*Sue should have been being criticized before John should have been being ~~criticized~~.
  - b. Sue should have been being criticized before John should have been ~~being criticized~~.
  - c. #Sue should have been being criticized before John should have ~~been being criticized~~.

The data in (5) – along with all of the (a) examples in (6)-(8) – represent a pattern that poses a potential problem for the present analysis, since they superficially seem to involve ellipsis below  $\text{VoiceInflP}$ , contrary to the analysis put forth here. However, we need not abandon the parsimonious position that low-VPE always targets  $\text{VoiceInflP}$ : these exceptions, which involve a surface realization of  $\text{VoiceInfl}^0_{\text{BE}}$  in the E-clause, follow a pattern that is regular and predictable under the model of the inflectional domain we have been assuming thus far (based heavily on that of Bjorkman 2011). Consider the full set of morphosyntactic realizations of  $\text{VoiceInfl}^0$  according to this model:

*Morphosyntactic realizations of  $\text{VoiceInfl}^0$*

- (12) a.  $\text{VoiceInfl}^0_{\text{be}} = [ \text{uINFL: } \underline{\text{INF}} ]$   
 You think this problem can't be solved.
- b.  $\text{VoiceInfl}^0_{\text{been}} = [ \text{uINFL: } \underline{\text{PERF}} ]$   
 This problem hasn't been solved.
- c.  $\text{VoiceInfl}^0_{\text{being}} = [ \text{uINFL: } \underline{\text{PROG}} ]$



This problem should be being solved.

Comparing (12) to the INFL hierarchy illustrated earlier in §1.3.1, recreated below, we see that the only instance of  $\text{VoiceInfl}^0$  that cannot be stranded ( $\text{VoiceInfl}^0_{\text{being}}$ ) corresponds to the one whose [ *uINFL*: \_\_\_ ] feature is valued by the next inflectional head immediately above it, namely  $\text{Prog}^0$ .<sup>6</sup>

(13) TP > PerfP > ProgP > (VoiceSynP) > VoiceInflP > vP

Crucially, the other heads capable of being stranded,  $\text{VoiceInfl}^0_{\text{be}}$  and  $\text{VoiceInfl}^0_{\text{been}}$ , are valued by heads higher up in the tree, above  $\text{Prog}^0$ . Given that every INFL head with visible features is going to be realized overtly, an analysis for this empirical pattern starts to emerge.

I suggest (following an earlier proposal in Sailor 2012a) that  $\text{VoiceInfl}^0$  is able to undergo relativized head movement (Roberts 2010) to the next highest inflectional head,  $\text{Prog}^0$ . This movement is relativized in that it skips over  $\text{VoiceSyn}^0$ , which is not an inflectional head. This movement is possible if and only if the landing site is an available inflectional position; i.e., if  $\text{Prog}^0$  lacks visible INFL features. If it undergoes this movement,  $\text{VoiceInfl}^0$  escapes the ellipsis site, and is stranded next to it in  $\text{Prog}^0$ . The advantage of this is that it accurately predicts why  $\text{VoiceInfl}^0_{\text{be}}$  and  $\text{VoiceInfl}^0_{\text{been}}$  can move: they are always locally c-commanded by an empty (inactive)  $\text{Prog}^0$  head, meaning they will always have an “escape hatch” for avoiding ellipsis. Note that this analysis crucially assumes (as Bjorkman 2011 does) that inactive INFL heads are nevertheless projected in the structure; see below for discussion. Likewise, it accurately predicts why  $\text{VoiceInfl}^0_{\text{being}}$  can never undergo movement: by definition, it can only appear underneath a filled  $\text{Prog}^0$ , which blocks movement.

---

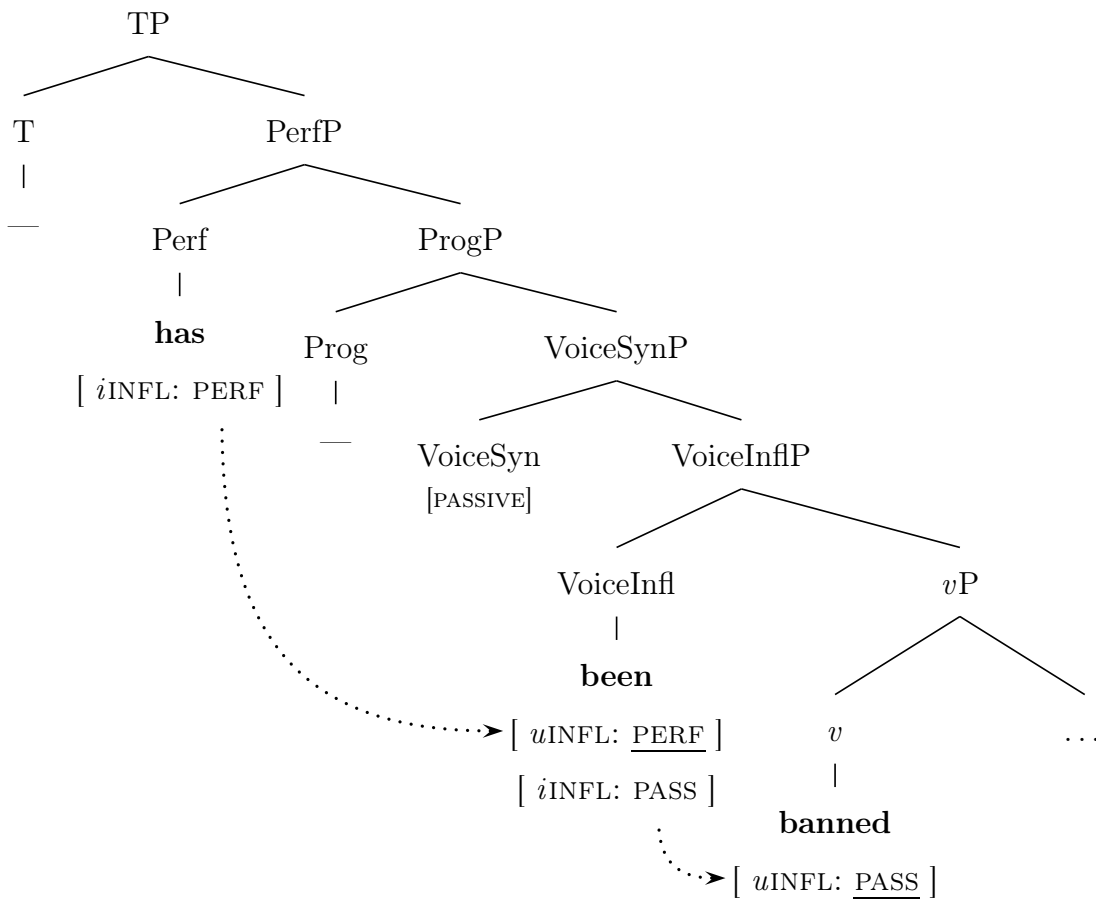
<sup>6</sup>As before,  $\text{VoiceSynP}$  is included here for completeness, but it is not an inflectional head, and is therefore irrelevant to the INFL-feature valuation process.

The examples and accompanying derivations below illustrate the analysis.

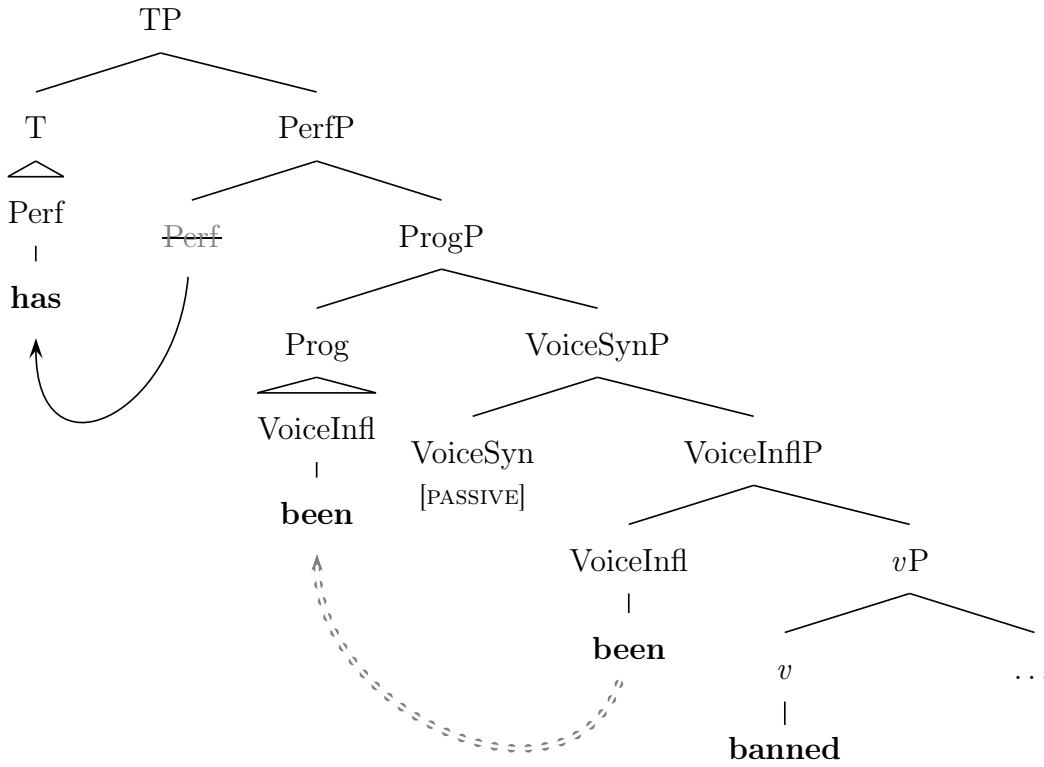
*Optional ellipsis of VoiceInft<sup>0</sup> = optional instance of VoiceInft<sup>0</sup>-to-Prog<sup>0</sup>*

- (14) a. The Queen should not be frisked, although the Prime Minister probably should (be).  
 b. Your mother had been banned from the premises before mine had (been).  
 c. This problem is being solved, even though nobody thinks it should be (\*being).

(15) **Step 1: Check INFL features**

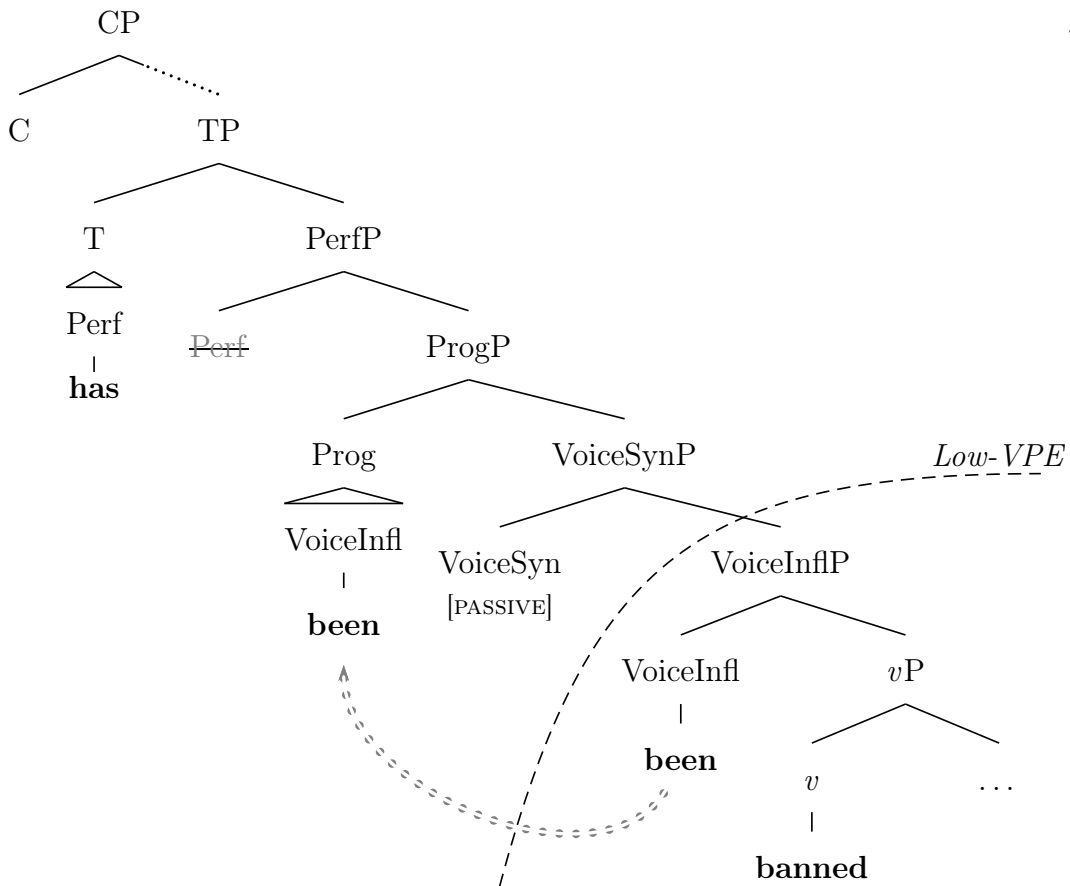


- (16) **Step 2:** *Has* raises to  $T^0$ ; *been* optionally raises to  $\text{Prog}^0$  (indicated with a dotted double-line)<sup>7</sup>



<sup>7</sup>I assume head movement proceeds via head adjunction, as is standard; I obscure this here for simplicity.

(17) **Step 3:** VoiceInflP sent to PF to be elided

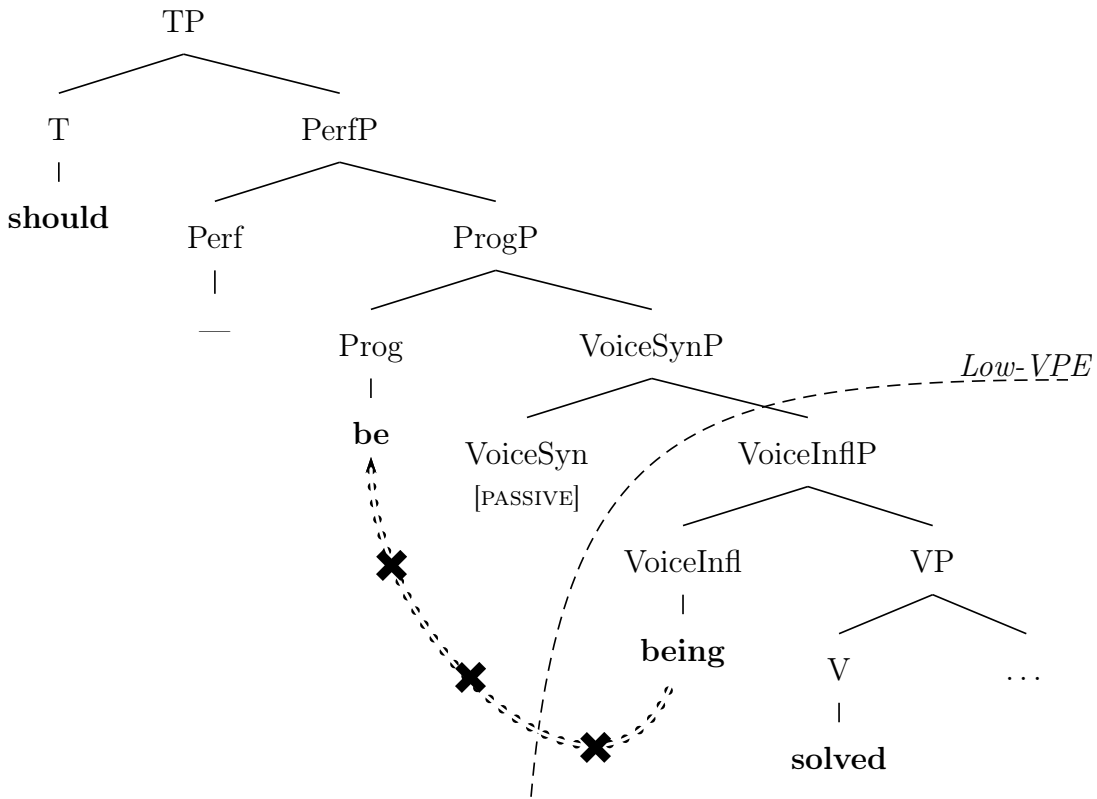


If  $\text{VoiceInfl}^0$  raises to the open  $\text{Prog}^0$  head in Step 2, then it survives ellipsis of  $\text{VoiceInflP}$  in Step 3. On the other hand, if  $\text{VoiceInfl}^0$  stays in-situ in Step 2, then it is elided with the rest of  $\text{VoiceInflP}$  in Step 3. Thus, the optional ellipsis of *be* and *been* in (14) is not due to a variability in the size of VPE, but due to optional head movement of  $\text{VoiceInfl}^0$  to a position outside of the elided  $\text{VoiceInflP}$ . In a sense, this is a much smaller-scale instance of an ellipsis pattern seen in the various verb-raising languages with VPE: namely, the so-called “V-stranding VPE” pattern (Goldberg 2005 and references therein), only with an auxiliary verb rather than a main verb. I return to this shortly.

Now consider *being*, which cannot undergo this movement. The reasoning is

straightforward: in order to arrive at the morphological form *being*, VoiceInfl<sup>0</sup> must bear the feature value [ uINFL: PROG ], but this is only possible if Prog<sup>0</sup> has visible features, and is thereby filled. As a consequence, head movement of *being* into Prog<sup>0</sup> will always be blocked, forcing it to remain in VoiceInflP to ultimately be elided at Spell-Out.<sup>8</sup> This is represented below (at Step 3 in the prior derivational sequence):

(18) **Step 3:** *Being* cannot raise to Prog<sup>0</sup>; VoiceInflP sent to PF to be elided



In sum, the unique structural configuration that licenses *being* correlates with its inability to escape from low-VPE. I take this as strong support for the movement analysis proposed here. (See Sailor 2012a:§4.5.3 for additional discussion of

<sup>8</sup>Note that *being* remains in-situ even if Prog<sup>0</sup> is able to move to T<sup>0</sup>, consistent with known constraints on head movement, i.e. that head movement out of a particular position does not avail (feed) head movement into that position.

this movement.)

Note, however, that I have not attempted to provide an explanation for why such movement to  $\text{Prog}^0$  should take place to begin with; indeed, there is no obvious independent semantic or syntactic motivation for it, as Sailor (2012a) and Harwood (2013b) note: in particular, there is no interpretational difference whatsoever between eliding  $\text{VoiceInfl}^0_{be}$  (4) or pronouncing it (5). The only generalization to be drawn is simply that the two heads involved are both inflectional heads (i.e. they bear  $i\text{INFL}$  features), but this does not get us much, since there is no clear evidence of movement into any of the other inflectional heads (excluding  $T^0$ ). One possible direction to take would be to assume that  $\text{Prog}^0$  has a special surface-realization requirement, i.e. that there is generalized auxiliary raising to  $\text{Prog}^0$  in the absence of a base-generated auxiliary there, à la  $T^0$  in English. However, this is too restrictive: avoidance of VPE by  $\text{VoiceInfl}^0_{BE}$  is optional, not obligatory (but see below); thus, if  $\text{Prog}^0$  were required to have a surface realization in all clauses, we would be left to explain why this requirement can be overlooked in the context of ellipsis (as would be the case for (4) and the (b) examples of (6)-(8)). Although the question posed here must remain open, I present additional evidence in support of this  $\text{VoiceInfl}^0$ -to- $\text{Prog}^0$  raising approach in the next section.

This state of affairs in head movement is not without precedent. The literature contains many instances of head movement whose motivations are ill-understood, to head positions whose presence would, *sans* movement, apparently be unnecessary in the derivation: for example, verb and/or auxiliary movement to a left-peripheral head in various languages, either to derive neutral V-initial word order or to satisfy a verb-second requirement (including related phenomena such as inversion in English questions). Thus, the movement I assume here needs to be better understood, to be sure, but it is in good company among various other

instances of head movement in that regard.<sup>9</sup>

The fact that this movement has no obvious semantic motivations makes it all the more essential that all INFL heads are projected in the syntax in all clauses, even vacuously (as Bjorkman 2011 also assumes). This is of course crucial for facilitating VoiceInfl<sup>0</sup>-to-Prog<sup>0</sup> movement: if INFL heads were not merged when they lacked visible features, then VoiceInfl<sup>0</sup> would always be stuck in-situ (unless called up to T<sup>0</sup>), and thus predicted to be elided obligatorily by VPE, contrary to fact. Indeed, the optionality of this movement (such that VoiceInfl<sup>0</sup><sub>BE</sub> either survives VPE or is deleted) would seem to provide indirect evidence for the presence of Prog<sup>0</sup> in the structure even in non-progressive clauses, given the independent arguments supporting the general inflectional architecture assumed here (see Bjorkman 2011). Nevertheless, the assumption that apparently-featureless inflectional heads are projected may be seen as undesirable to some, depending on one's theoretical commitments. In particular, some might object that merging such heads (and movement into such heads in particular) would violate various principles of Minimalism relating to economy. However, Cinque (1999:§6.2) suggests that projecting the complete functional hierarchy in all clauses – even when various heads in the array would be projected vacuously – may actually be the most Minimalist approach of all, since it would require the fewest assumptions (regarding selection among heads in the array in particular). Thus, I adopt it here, since doing so avails us of a straightforward explanation for the asymmetry between VoiceInfl<sup>0</sup><sub>be(en)</sub> versus VoiceInfl<sup>0</sup><sub>being</sub> with respect to VPE.

---

<sup>9</sup>Some of these questions could perhaps be reduced to pre-existing questions in the literature. For example, a possibility I will not explore in depth here is that VoiceInfl<sup>0</sup>-to-Prog<sup>0</sup> movement is a side-effect of independent (but optional) movement of *v*P to Spec-VoiceInflP. If this movement of *v*P occurs, it would yield a doubly-filled COMP configuration; however, this illicit configuration could be resolved (in the spirit of Koopman 1996) by moving VoiceInfl<sup>0</sup> to a position higher in the tree, namely Prog<sup>0</sup> (meaning movement of *v*P would be blocked when Prog<sup>0</sup> is filled, since the doubly-filled VoiceInfl configuration could not be resolved). The question then becomes why *v*P should (optionally) move, but this question is much more manageable, and could be approached several different ways, perhaps building off prior work arguing that a projection of the verb undergoes short movement within the verbal-inflectional domain, e.g. Baltin (2002).

In the next section, I consider some additional supporting evidence from GET-passives for this inflectional state of affairs.

## 2.4 Supporting and extending the analysis: GET-passives

### 2.4.1 Support: GET can't get stranded, but BE can (\*get)

So far in this chapter, I have been assuming that the projection we now know as VoiceInflP is always elided in VPE, based on the unstrandable nature of *being*. However, *being* is the only exponent of VoiceInfl<sup>0</sup> we have seen so far that is obligatorily elided: as I argued above, VoiceInfl<sup>0</sup><sub>*be/been*</sub> can escape VPE via head movement (an option unavailable to VoiceInfl<sup>0</sup><sub>*being*</sub>). At this point, the skeptic might assert that the source of these patterns is not VoiceInflP generally, but rather *being* in particular, perhaps by way of some special lexical property. In other words, perhaps VPE truly does just target *v*P and nothing higher, and the obligatory omission of *being* in the relevant contexts follows from some additional factor unique to *being*. Before continuing with the present analysis, I will first argue against this alternative approach on empirical grounds.

Under the present analysis, the inability of *being* to raise out of the ellipsis site dooms it to deletion. The other exponents of VoiceInfl<sup>0</sup>, *be* and *been*, always have raising as an option, meaning they always have ellipsis-avoidance as an option. However, if the alternative analysis is correct, then raising is irrelevant, and indeed need not even occur: *being* is special, the reasoning would go, not VoiceInfl<sup>0</sup>. To decide the matter on empirical grounds, we require a different, non-BE realization of VoiceInfl<sup>0</sup> – one that never raises, regardless of its inflectional form (i.e., regardless of whether Prog<sup>0</sup> is an open position). If such a flavor of VoiceInfl<sup>0</sup> exists, it could be used to investigate whether short movement to Prog<sup>0</sup> is indeed the phenomenon responsible for rescuing VoiceInfl<sup>0</sup> heads from VPE. I claim that this flavor of VoiceInfl<sup>0</sup> is in fact embodied in the GET that arises in the English



GET-passive.<sup>10</sup> First, I will argue that this use of GET occupies VoiceInfl<sup>0</sup> (putting it in competition with BE), then I will show that it fails to raise to Prog<sup>0</sup>, revealing its diagnostic value.

Although VoiceInfl<sup>0</sup><sub>being</sub> plays host to progressive morphology (*-ing*), it also bears [ *i*INFL: PASS ], meaning it is associated with the interpretation of passive in the clause. English has a second type of passive – the so-called GET-passive – which involves the use of GET, but which never co-occurs with *being*. In fact, this use of GET behaves as though it also occupies VoiceInfl<sup>0</sup>, putting it in complementary distribution with VoiceInfl<sup>0</sup><sub>BE</sub>:<sup>11</sup>

*Complementary distribution of English BE-passives and GET-passives*

- (19) a. John could {be/get} arrested for wearing that.  
 b. \*John could be gotten arrest(ed) for wearing that.  
 c. \*John could get been arrest(ed) for wearing that.
- (20) a. The children should have been {being/getting} yelled at.  
 b. \*The children should have been being gotten yell(ed) at.  
 c. \*The children should have been getting been yell(ed) at.

This particular instance of GET is associated with a passive interpretation and triggers the appearance of passive morphology on its complement, which, in combination with the distributional facts above, suggests that it occupies VoiceInfl<sup>0</sup>. Additional evidence for this can be found by comparing the behavior of passive GET with that of VoiceInfl<sup>0</sup><sub>being</sub> in a variety of contexts in which the latter exhibits special behavior as compared to the higher inflectional auxiliaries, e.g. in VP topicalization (see chapter 4). Across the board, passive GET behaves the same as

---

<sup>10</sup>Using the notation introduced earlier for forms of BE (see ch. 1: fn. 7), I use small caps (e.g. GET) to abstract away from surface inflectional forms.

<sup>11</sup>Crucially, I am referring only to the variety of GET in examples whose surface subjects lack what Orfitelli (2012) calls the *responsibility reading*. See *op cit.* and references therein for many complex details concerning GET that I leave aside here.

VoiceInfl<sup>0</sup><sub>being</sub>, strongly suggesting that the former is also located in VoiceInfl<sup>0</sup>.

That being said, VoiceInfl<sup>0</sup><sub>BE</sub> differs from VoiceInfl<sup>0</sup><sub>GET</sub> in (at least) one significant way: whereas the former undergoes head movement to T<sup>0</sup> in various circumstances, as is typical of finite auxiliaries in English (Akmajian and Wasow 1975), the latter apparently never does. Consider the following:

*VoiceInfl<sup>0</sup><sub>GET</sub> does not raise to T<sup>0</sup>*

- (21) a. John was supposedly criticized today.  
 b. \*John got supposedly criticized today. (cf. ✓ *John supposedly got...*)
- (22) a. Bill wasn't promoted this year.  
 b. \*Bill gotn't promoted this year. (cf. ✓ *Bill didn't get...*)
- (23) a. Was Mary arrested for her performance?  
 b. \*Got Mary arrested for her performance? (cf. ✓ *Did Mary get...*)
- (24) a. John was arrested before Mary was.  
 b. \*John got arrested before Mary got. (cf. ✓ *...before Mary did.*)

Thus it seems that VoiceInfl<sup>0</sup><sub>GET</sub> does not undergo movement to T<sup>0</sup> in the relevant environments, unlike VoiceInfl<sup>0</sup><sub>BE</sub>.

Returning now to VPE, the diagnostic value of VoiceInfl<sup>0</sup><sub>GET</sub> becomes clear. Recall that VoiceInfl<sup>0</sup><sub>be</sub> and VoiceInfl<sup>0</sup><sub>been</sub> are able to survive VPE, but VoiceInfl<sup>0</sup><sub>being</sub> never can. I argued that this asymmetry followed from the ability or inability of the auxiliary to raise to Prog<sup>0</sup>. If VoiceInfl<sup>0</sup><sub>GET</sub> is immobile, as the data above in (21)-(24) suggest, then we expect it to pattern like VoiceInfl<sup>0</sup><sub>being</sub> in all environments, regardless of its surface inflectional form (i.e. regardless of the status of Prog<sup>0</sup>). In other words, we expect passive GET to be obligatorily elided by VPE in all environments, and this is exactly what we find:

*Optional ellipsis of VoiceInfl<sup>0</sup><sub>BE</sub>, but not VoiceInfl<sup>0</sup><sub>GET</sub>*

- (25) a. My application will be rejected before yours will be.

- (26) b. \*My application will get rejected before yours will get.
- a. Doug should have been sterilized for the same reason Nick should have been.
- b. \*Doug should have gotten sterilized for the same reason Nick should have gotten.

Thus, a variety of VoiceInfl<sup>0</sup> that is independently immobile – one realized by forms of GET – cannot ever be stranded by VPE. This provides clear (albeit indirect) evidence for the argument from the preceding subsection that material in VoiceInfl<sup>0</sup> can only survive VPE by way of raising to Prog<sup>0</sup>,<sup>12</sup> militating against an approach that dismisses the unstrandability of *being* as e.g. a quirk of its lexical entry.<sup>13</sup>

#### 2.4.2 Extensions: GET-passives and the identity condition on ellipsis

Having established the category and behavior of (a variety of) passive GET, we are in a position to ask more complex questions about its behavior in VPE contexts. I would like to focus in particular on what it might be able to tell us about the identity condition on ellipsis, which has been the subject of a lengthy and ongoing debate in the literature (see Merchant 2013a, van Craenenbroeck and Merchant 2013, and van Craenenbroeck *forthcoming* for recent surveys of the discussion).<sup>14</sup> To approach this question, we will look at the behavior of VoiceInfl<sup>0</sup><sub>BE</sub> when it is in a position to antecede ellipsis of VoiceInfl<sup>0</sup><sub>GET</sub>, and vice-versa.

<sup>12</sup>Harwood (2013b:§5) argues against the analysis from the previous subsection that short head movement is responsible for the (non-)strandable nature of varieties of VoiceInfl<sup>0</sup><sub>BE</sub>; however, it is unclear how his approach could account for the GET-passive data discussed here.

<sup>13</sup>Of course, the fact that VoiceInfl<sup>0</sup><sub>GET</sub> is unable to raise (and thus cannot be stranded by VPE) requires some discussion, but here I believe it makes sense to invoke the lexicon: if we regard such instances of GET as “auxiliaries” (given their distribution in the inflectional array), then their immobility is a noteworthy property, but it is not unique among heads in the INFL domain: see, for example, possessive HAVE in American English (and other varieties).

<sup>14</sup>Briefly, two major positions have been taken regarding the relationship between the ellipsis site and its equivalent in the antecedent: one claims the two must be syntactically identical, and the other claims that they must be semantically identical.

First,  $\text{VoiceInfl}_{\text{BE}}^0$  cannot antecede ellipsis of  $\text{VoiceInfl}_{\text{GET}}^0$ , which we can test by exploiting the latter's inability to move, as established above around (21):

*$\text{VoiceInfl}_{\text{BE}}^0$  cannot antecede ellipsis of  $\text{VoiceInfl}_{\text{GET}}^0$*

- (27) a. John was criticized before Mary was ~~criticized~~.  
b. \*John was criticized before Mary did ~~get criticized~~.

- (28) a. John was evicted for the same reason Mary was ~~evicted~~.  
 b. \*John was evicted for the same reason Mary did ~~get evicted~~.

As we saw in the previous subsection in (25),  $\text{VoiceInfl}^0_{\text{GET}}$  must be elided with the main verb for grammaticality (i.e., it cannot be stranded), but this principle is not being violated in (27). What, then, is the source of its ill-formedness? It would appear that these examples are ruled out by the identity condition on ellipsis, since the equivalents of the examples in (27) involving E-clauses with BE-passives are completely grammatical. This tells us relatively little about the nature of that condition, though, since the mismatch between  $\text{VoiceInfl}^0_{\text{BE}}$  and  $\text{VoiceInfl}^0_{\text{GET}}$  is morphological in nature, given that the heads they appear on would seem to have the same interpretable and uninterpretable INFL features (and, to the extent that it matters, the two  $\text{VoiceSyn}^0$  heads selecting  $\text{VoiceInfl}^0$  in these two types of passives are presumably identical as well, given the similarities in their passive syntax).<sup>15</sup> Such a mismatch would not fall entirely within the domain of a syntactic (featural) identity requirement, nor a semantic one either. I return to this momentarily.

Consider the opposite configuration, where the E-clause contains  $\text{VoiceInfl}^0_{\text{BE}}$ , and the A-clause contains  $\text{VoiceInfl}^0_{\text{GET}}$ . Here, we need  $\text{VoiceInfl}^0_{\text{BE}}$  to raise out of the ellipsis site in order to confirm the mismatch in the surface representation (since, if it were elided, there would be no reason to assume a mismatch was present). The result is again ungrammatical:

*$\text{VoiceInfl}^0_{\text{GET}}$  cannot antecede ellipsis of  $\text{VoiceInfl}^0_{\text{BE}}$*

- (29) a. \*You always get cited in the same papers I am  ~~$\text{VoiceInfl}^0_{\text{GET}}$   $t_{\text{am}}$  cited in~~  
 b. \*John will get rewarded if Mary will be  ~~$\text{VoiceInfl}^0_{\text{GET}}$   $t_{\text{be}}$  rewarded~~

---

<sup>15</sup>It should be noted that GET-passives are not completely synonymous with regular passives: for example, the surface subject of the former carries an interpretation of “affectedness” that the latter need not (Orfitelli 2012). Perhaps this difference is sufficient to constitute non-identity, but whether it is semantic or syntactic in nature (or some combination thereof) is unclear.

We can view these data as indirectly supporting the present analysis that VPE always targets VoiceInflP: if it elided just *v*P, then the reason for the ill-formedness in (29) would be mysterious, given that the two elided *v*Ps are wholly identical to the A-clause *v*Ps. If, on the other hand, VPE elides VoiceInflP, then these examples are presumably unacceptable for the same reason as those in (27): the elided VoiceInflP is non-identical to its antecedent in some relevant way, which now must be determined.

If this is on-track, then example (29) has potentially intriguing consequences for the theory of VPE. Since, as mentioned earlier, such examples would not obviously run afoul of a syntactic (featural) identity requirement, nor a semantic one, the question of what rules it out remains. To address this, we must consider data from an unexpected source; namely, VPE in languages with verb-raising.

In such languages – Hebrew, Irish Gaelic, etc. – the main verb moves out of the VPE site prior to deletion, deriving the V-stranding VPE pattern exemplified in (30):<sup>16</sup>

*V-Stranding VPE* (Irish Gaelic example adapted from McCloskey 2005:6)

- (30) Níor **cheannaigh**<sub>v</sub> [ *t*<sub>v</sub> mé ariamh teach ], ach **cheannóinn**<sub>v</sub>  
 NEG.PAST buy.PAST I ever house but buy.COND  
 {~~*t*<sub>v</sub> mé teach~~}  
 ‘I never bought a house, but I would’ (lit. ‘...but would buy’)

Goldberg (2005:§4.1) observes that such VPE configurations are subject to an additional identity requirement beyond the featural or semantic identity requirements typically discussed – one that is either not relevant or whose effects cannot be observed in non-V-stranding VPE configurations. Specifically, she shows that V-stranding VPE configurations are subject to what she calls the *Verbal Identity Requirement*, which demands that the A-clause’s verb and the E-clause’s verb be

---

<sup>16</sup>See Goldberg (2005) for a series of diagnostics ruling out object-drop as the source of the pattern in data like (30).

completely identical in both their root and derivational morphology (but inflectional mismatches are permissible, as seen in the above example). The following example, which involves a mismatch in verb root morphology, illustrates the point. Typically, such a mismatch might be ruled out by e.g. a semantic identity requirement, since verbs differing in their roots would presumably also differ in their lexical semantics, meaning they would fail to mutually entail one another in an antecedent-ellipsis configuration (see Merchant 2001 on e-GIVENness). However, McCloskey (2005) observes that relevant examples avoiding this confound can be constructed in Irish Gaelic by exploiting a morpheme (*-áil*) that turns English borrowings into licit Irish Gaelic verbs. This allows for the construction of an example in which e.g. an Irish Gaelic verb is in the E-clause, and a synonymous verb borrowed from English is present in the A-clause (or vice versa). Even though such examples presumably would not violate a semantic identity requirement, they are nevertheless ungrammatical (McCloskey 2005:7):

*Verbal Identity Requirement: No mismatching roots*

- (31) Q: a-r            **mhiss-eáil**<sub>v</sub> [ *t<sub>v</sub>* tú é ]?  
           INT-PAST *miss-ÁIL*                    you him  
           ‘Did you miss him?’
- A: \***chrothnaigh**<sub>v</sub> [ ~~*t<sub>v</sub>* mé é~~ ]  
           miss.PAST  
           Intended: ‘I did.’

The mismatch in root morphology between the A-clause’s verb *mhiss*, borrowed from English, and the E-clause’s native Irish Gaelic verb *chrothnaigh* (‘miss.PAST’) constitutes a violation of the Verbal Identity Requirement, and thus the sentence is ruled out.<sup>17</sup> Crucially, this requirement applies even though both verbs have undergone movement outside of the relevant constituent (by hypothesis, VoiceInflP).

<sup>17</sup>McCloskey (2005:7) shows that examples such as (31) are grammatical if the verbs are made to match in their root morphology, e.g. by replacing *mhiss-eáil* with *chrothnaigh*, or vice versa. For additional examples illustrating the Verbal Identity Requirement, see Goldberg (2005:§4.1).

Thus, for the purposes of satisfying the preconditions on ellipsis, it is as though the main verbs occupy their base positions rather than their surface positions, a property which Goldberg (2005:180) characterizes as obligatory reconstruction of the verbs (which I will assume as well, though nothing hinges on this).

We return now to the English example involving GET in (29). The ellipsis identity condition (whether featural, semantic, or some combination thereof) compares an E-clause VoiceInflP to some A-clause VoiceInflP. For ease of reference, call these  $VIP_E$  and  $VIP_A$ . On the surface,  $VIP_A$  in (29) is [ *get mentioned* ], and  $VIP_E$  is [  $t_{\text{VoiceInfl}^0}$  *mentioned* ]. The latter contains a trace that the former lacks; however, this does not seem to be the problem, given that examples of the following sort are good:

- (32) This problem is [VoiceInflP *being solved* ], even though nobody thinks it should be ~~[VoiceInflP  $t_{\text{be}}$  *solved* ]~~

This example also involves a  $VIP_E$  containing a trace that the  $VIP_A$  lacks (since *being* is immobile, as we saw). It seems we must conclude, then, that even these auxiliaries obligatorily reconstruct (in the sense of Goldberg 2005:180) for the purposes of satisfying the identity condition, otherwise example (29) could not properly be ruled out. In other words, this is precisely the pattern we would expect to see if head movement out of the ellipsis site were possible in English, and if such movement were subject to the Verbal Identity Requirement.

This novel evidence for the Verbal Identity Requirement in English suggests that this requirement is relevant for any type of verb movement out of an ellipsis site, whether it is of a main verb or an auxiliary.<sup>18</sup> Note that this requirement has only been claimed to hold in V-raising languages. Goldberg (2005:167) even states

---

<sup>18</sup>The Verbal Identity Requirement may be even more general than that: it might require identity of any head that moves out of an ellipsis site (making the name a misnomer), rather than holding just for verbs. I leave investigation of this question to future work.



directly that English cannot provide evidence for or against this requirement, since it lacks V-raising, and although the auxiliaries *have* and *be* differ in their root morphology and do undergo raising, the former never occupies a position within the ellipsis site.<sup>19</sup> Presently, we have overcome this by appealing to data that Goldberg does not consider, namely data from the GET-passive: the key to the above result in (29) was finding two elements in the INFL-domain that could occupy the same base position but differ in their root morphology (since, again, inflectional differences are not relevant to the Verbal Identity Requirement). Such examples can only be constructed in English by using the GET-passive, since GET and BE can each be generated within the ellipsis site, and yet the two do not differ in their semantic and featural properties.

### 2.4.3 Summary

Combining Bjorkman's (2011) proposal with the view that VPE only ever elides VoiceInflP yields significant advantages, as we have just seen. It neatly resolves two longstanding bugbears in the ellipsis literature: namely, the obligatory deletion of *being*, and the optional deletion of VoiceInfl<sup>0</sup><sub>be(en)</sub>. Scrutiny of the latter problem reveals the existence of a head movement operation from VoiceInfl<sup>0</sup> to Prog<sup>0</sup>, sometimes optionally. Although we do not currently understand why this movement occurs, we have seen evidence that it nevertheless does occur, and only in a constrained, predictable set of environments (namely, when Prog<sup>0</sup> is an open position). Although certain issues remain to be explained, the analysis for VPE that I develop here allows us to explore questions that we could not previously. Specifically, evidence from GET-passives suggests that the Verbal Identity Require-

---

<sup>19</sup>I argued in chapter 1 that, for some speakers, ellipsis of perfect *have* is in fact possible with VPE. However, relevant examples exploiting this property to test the Verbal Identity Requirement still could not be constructed, since doing so would require an aspectual mismatch of a sort that would necessarily involve a featural mismatch in addition to the verbal root mismatch, thereby invalidating the test. See Goldberg (2005:§4.1.2) for reasons that the relevant examples cannot be constructed by using possessive *have* in British English.

ment is active in English, which is a significant result for the VPE typology.

## 2.5 Low-VPE vs. high-VPE: The Principle of Antecedent-Containment Avoidance

Now that we have a clearer picture of the region of the clause involved in VPE – and in low-VPE in particular – we can return to the bigger picture, which is the basic distinction between low-VPE and high-VPE. To this point, I have only attempted to establish that this distinction exists, not why it exists; that is, I have not said anything about what this distinction might arise from, or why the grammar should be burdened with variability in this domain at all. Below, I develop an initial attempt at a theory that addresses these questions.

I will start with an idealization of the data such that all ellipsis clauses are related to their antecedents by either coordination or subordination,<sup>20</sup> and all subordinate clause types – temporals, reasons, purposes, etc. – occupy the same position in the clause (to be made clear below).<sup>21</sup> We now know that VPE within a clause that is subordinated beneath its antecedent is always low-VPE. We have also seen that this is not true for coordinated ellipsis clauses: those allow high-VPE. Below, I argue that the existence of this preference for particular sizes of VPE in particu-

---

<sup>20</sup>It may be that this is not an idealization at all. Ellipsis is an anaphoric process and thus tightly constrained by discourse. It is plausible that all ellipsis clauses are either coordinated with or subordinated under their antecedents at the level of discourse, if not in the syntax explicitly. It is further possible that even discourse-level coordination and subordination have syntactic reflexes. See below.

<sup>21</sup>Despite the wealth of recent cartographic work mapping the structure of various adjunct-like elements (adverbs, PPs, adjectives, etc.), almost no treatment of this sort has been given to subordinate adverbial clauses. Such an undertaking far exceeds the scope of this dissertation, but it would seem that at least two basic possibilities exist: either these clauses are generated in distinct positions throughout the extended verbal/inflectional domain – presumably those identified by Cinque (1999:§4) for various sub-clausal adverbial elements of similar meaning – or they are generated in more or less one position, perhaps without a rigid order, akin to the varied set of “circumstantial adverbials” of Cinque (1999:§1.5). The data below seem to favor the latter option, but I leave the matter open. See Valmala (2009) and Haegeman (2012:§4.5) for discussion.

lar environments should be understood as a side-effect of *antecedent-containment avoidance*, a general principle of grammar. Specifically, I argue that high-VPE in subordination structures is marked because recovery of such a large ellipsis site would require taking an antecedent that contains the locus of subordination, thus yielding an antecedent-contained deletion (ACD) configuration. This problem is avoided if low-VPE occurs instead. High-VPE is free to apply in coordination contexts, however, since such structures could never yield ACD.

Since this argument will crucially rely on the attachment site for subordinate clauses within the verbal-inflectional domain, I address this matter first. Then, I discuss some background on ACD, and I address why and how the grammar avoids it. If sound, this line of reasoning can derive the low-VPE vs. high-VPE distinction we have seen throughout this chapter.

### 2.5.1 On the locus of subordination

The argument to come requires that we ascertain the position occupied by subordinate clauses in English. I discuss two diagnostics for addressing this below.<sup>22</sup>

First, there is evidence that subordinate clauses attach within the extended projection of the predicate in English, but properly below the surface position of subjects in Spec-TP. Consider first an argument from Hornstein (1998), who shows that subordinate clauses attach above the subject's base position, but below its surface position. Hornstein presents the following pair of examples – his (12) – and discusses the available scope interpretations for each:

---

<sup>22</sup>As mentioned above, I treat all subordinate clauses (central adverbial clauses in the terminology of Haegeman 2012:§4.5) uniformly, as the VPE data suggest they should be. There may be other diagnostics that draw distinctions within this set (on this see discussion in Valmala 2009), but as these could conceivably have non-structural explanations, I leave them aside. It may be that particular subordinate clause types bear a direct relation to the clause's event variable, suggesting a structural relation between them; however, I leave this question (and any discussion of events) aside for future work.

- (33) a. Someone serenaded every woman.  
b. Someone<sub>i</sub> serenaded every woman before he<sub>i</sub> left the party.

Hornstein notes that although (33a) famously allows a wide-scope interpretation for the quantified object, this reading disappears in (33b), where the subject obligatorily takes wide scope when it binds a pronoun inside a subordinate clause. In order for the object to take wide scope, he argues, the subject must be interpreted in its base position within the predicate; however, being in such a low position would preclude the subject from binding into the subordinate clause at the relevant level of representation. Assuming Hornstein’s reasoning is on track, this tells us that subordinate clauses such as the *before* clause in (33b) attach above the first-merge position of the subject, but below the subject’s surface position in Spec-TP.

However, we now know that there is a highly articulated structure between those two positions (setting aside the precise location of the subject’s base position for the moment); so, although Hornstein’s argument narrows the possibilities for the position of subordination within the array, it does not pinpoint it. While the above argument is sufficient to make the point to come regarding ACD avoidance, we would gain more insight if we could somehow force the potential binder to surface lower in the array—this would allow us to shrink the set of possible attachment sites for subordination depending on the resulting interpretations.

Existential constructions provide such a configuration. As Harwood (2012) argues at length, the associate of a *there*-existential in English occurs in a specific position within the inflectional array. In sentences of the sort we have seen throughout this chapter involving maximal exponence of inflectional heads, existential associates always surface to the immediate left of *being*:

*Associates immediately precede ‘being’* (adapted from Harwood 2012: ex. (23))

- (34) a. \*There could have been being a man punished for his crimes  
 b. There could have been a man being punished for his crimes  
 c. \*There could have a man been being punished for his crimes  
 d. \*There could a man have been being punished for his crimes  
 e. \*There a man could have been being punished for his crimes

This tells us that associates occupy a position lower than Prog<sup>0</sup> (hosting *been* here) but higher than VoiceInfl<sup>0</sup> (hosting *being*).

The set of possible positions for existential associates can be narrowed further if we look back at VPE. Even in low-VPE contexts, existential associates cannot surface, suggesting that the position hosting the associate falls within VoiceInflP (but, again, above its head). In the data below, the # mark indicates that only a non-ellipsis reading is possible in the subordinate clause:

*Existential associates do not survive low-VPE*

- (35) a. #There will be several women chosen for that job before there will be several men.  
 b. Mary thought there should've been a Republican being considered even though there clearly shouldn't have been (#a Republican).

Thus, we can conclude that existential associates occur in Spec-VoiceInflP, since they must precede *being* in VoiceInfl<sup>0</sup> and yet they must not survive ellipsis of VoiceInflP.<sup>23</sup> See Harwood (2012) for additional discussion of the position of associates.

With this in place, we can now investigate the binding properties of these low

---

<sup>23</sup>Note that e.g. “VP” topicalization yields the same result:

- (i) If Mary said there should have been a man being interrogated, then...  
 a. ...a man being interrogated there should have been.  
 b. \*...being interrogated there should have been a man.

See chapter 4.

subjects with respect to pronouns inside subordinate clauses.<sup>24</sup> If low existential associates can bind into subordinate clauses, then those clauses must attach below VoiceInflP (and thus be properly included in even low-VPE). However, if they cannot bind into such clauses, then those clauses presumably attach higher in the inflectional array, and require a higher subject – one in Spec-TP – for the bound reading to arise. The data below suggest the latter conclusion:

*Existential associates cannot bind into subordinate clauses*

- (36) a. \*There should have been [no student]<sub>i</sub> sleeping in the library before he<sub>i</sub> was finished with exams.
- b. \*There will be [at least one student]<sub>i</sub> being praised because his<sub>i</sub> parents are present.
- c. \*There has been [no officer]<sub>i</sub> investigated after he<sub>i</sub> used the phrase “resisting arrest”.
- d. \*There will be [no officer]<sub>i</sub> being promoted if he<sub>i</sub> hasn’t met the brutality quota.

Since these associates in Spec-VoiceInflP are unable to bind pronouns inside accompanying subordinate clauses, we must conclude that these clauses are attached outside the c-command domain of existential associates, i.e. they are attached above Spec-VoiceInflP. However, they are apparently all attached within the c-

---

<sup>24</sup>The facts regarding variable binding in subordinate clauses are not entirely straightforward. For example, as Lasnik and Stowell (1991) and others note, there are instances where even direct objects appear to be capable of binding into predicate-level adverbials, as in (i). However, a fully clausal analogue of (i) is predictably degraded, as in (ii), raising questions about the status of the adverbial in (i).

- (i) Paul Masson will sell [no wine]<sub>i</sub> before its<sub>i</sub> time.
- (ii) \*Paul Masson will sell [no wine]<sub>i</sub> before it<sub>i</sub> should be sold.

It may be that (i) does not involve variable binding at all, and that the adverbial is actually being interpreted as though it contains a (contracted) cleft structure, i.e. ...*before it’s time (for us to sell)*. Still, a small set of examples similar to (i) have been put forward elsewhere that resist alternative explanation (see e.g. Valmala 2009). While the existence of such cases potentially weakens the diagnostic I use here, their existence also poses a fundamental problem for theories of c-command and constituency more generally. Since they are apparently a problem for everyone, I leave them aside.

command domain of subjects in canonical subject position, Spec-TP:

*Subjects in canonical subject position can bind into subordinate clauses*

- (37)
- a. [No student]<sub>i</sub> should have been sleeping in the library before he<sub>i</sub> was finished with exams.
  - b. [At least one student]<sub>i</sub> will be being praised because his<sub>i</sub> parents are present.
  - c. [No officer]<sub>i</sub> has been investigated after he<sub>i</sub> used the phrase “resisting arrest”.
  - d. [No officer]<sub>i</sub> will be being promoted if he<sub>i</sub> hasn’t met the brutality quota.

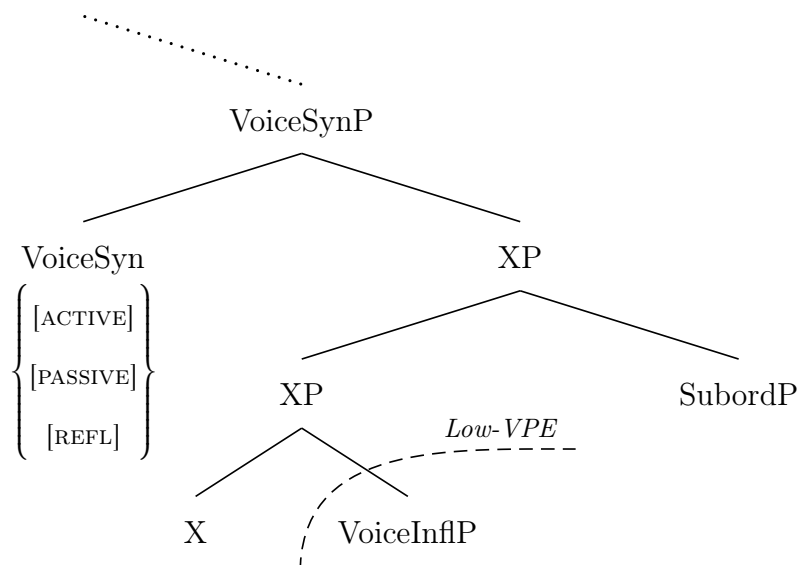
The conclusion is that subordinate clauses occupy a position below Spec-TP, but above Spec-VoiceInflP.

For concreteness, I assume that subordinate clauses are of category SubordP, and are adjoined in the superordinate clause to the right of an XP immediately dominating VoiceInflP.<sup>25</sup> As I discuss in detail shortly, this XP is outside the domain of low-VPE, but obviously within the domain of high-VPE:

---

<sup>25</sup>It may be that they adjoin directly to VoiceInflP instead; see below.

(38)



Next we move to the details of ACD, and consider a possible origin of the low-VPE vs. high-VPE distinction.

### 2.5.2 Background on antecedent-contained deletion

Antecedent-contained deletion (ACD) refers to ellipsis configurations in which the antecedent for some instance of ellipsis properly includes the ellipsis site itself (Bouton 1970, Sag 1976, May 1985, among others). This sort of structure would seem to give way to a problem of interpretational recursion, dubbed the “infinite regress” problem, such that recovery of the ellipsis site would yield yet another ellipsis site to be recovered, ad infinitum. A canonical example is below, accompanied by a simple sketch of the problem:



*Antecedent-Contained Deletion*

- (39) John will speak to everyone Mary will ○.  
○ = [ speak to everyone Mary will ○ ]

Such examples are fully grammatical and interpretable; however, compositional approaches to interpretation cannot accommodate such sentences straightforwardly, so various proposals have been put forth to solve this infinite regress problem.

The most widely-adopted proposal is that of May (1985), who argues that the quantified object (including its relative clause containing the ellipsis site) undergoes an instance of Quantifier Raising (QR) to a position outside the antecedent VP, thereby eliminating the source of the regression. A QR-based approach has broad applicability, since the prototypical ACD examples cited in the literature involve quantified objects of the sort that have been thought to undergo QR independently. However, as May (1985) and Harley (2002) point out, ACD configurations can be built around non-quantified objects as well:

- (40) I read the book that John did ○.  
○ = [ read the book that John did ○ ]

Harley argues that maintaining a QR analysis for resolving infinite regress in ACD requires saying that even non-quantified objects must undergo QR in the context of ACD (see also May 1985).<sup>26</sup>

In other words, for QR-based approaches, ACD configurations require a special grammatical mechanism (or a special application of one); their resolution cannot be reduced entirely to independently-needed processes. This analytically-marked status of ACD configurations will play an important role in the discussion to come.

---

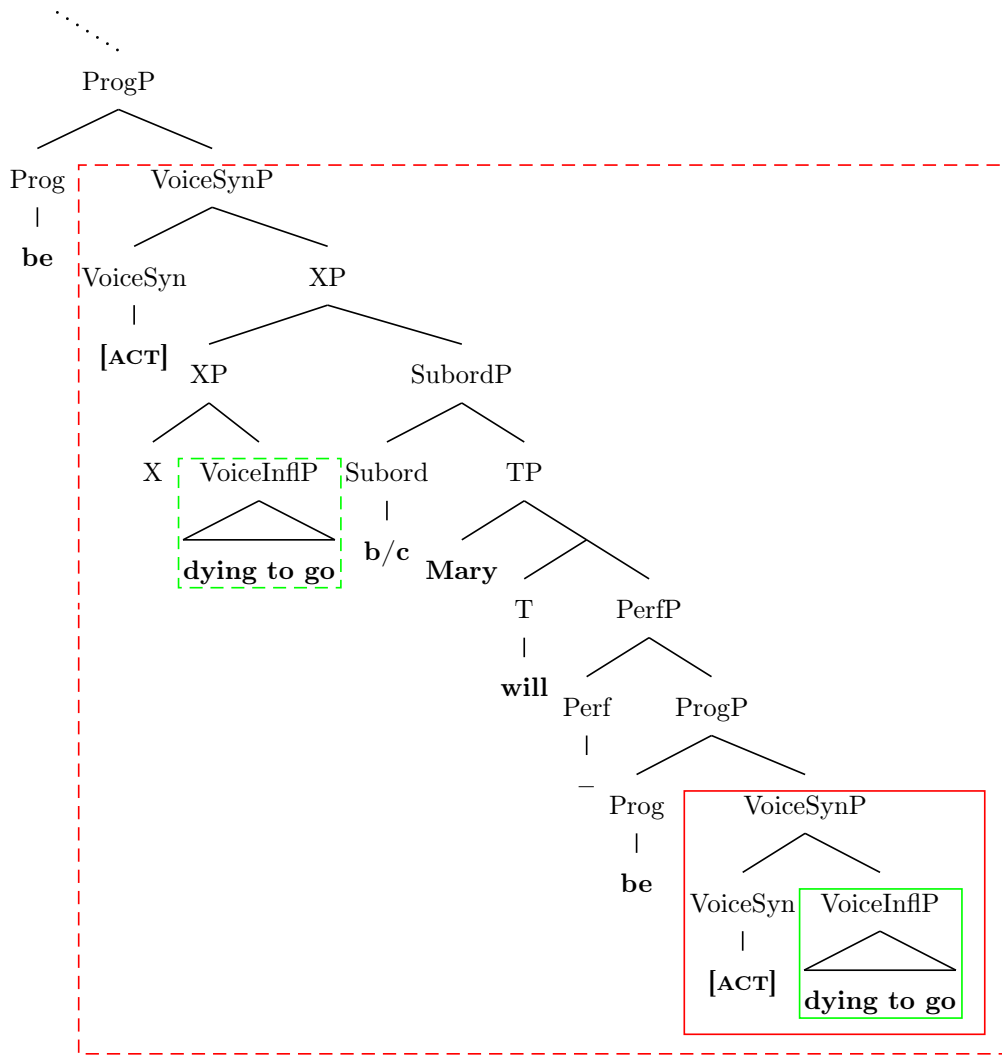
<sup>26</sup>Specifically, Harley (2002:662) argues that “DPs containing ACD structures are quantificational, but DPs without such structures are not quantificational”; that is, she argues that there are fundamentally two different (competing) types of definite DP in the grammar, and only a derive-and-crash approach to structure building can explain where each converges.

### 2.5.3 Subordination can create ACD configurations

Examples of ACD in the literature have always involved ellipsis sites embedded within a higher  $vP$ , e.g. within a relative clause attached to a matrix  $vP$ -internal object DP, as in (39)-(40) above (but see Yoshida 2010). However, if the size of the VPE site is larger than  $vP$ , as I have shown that it is throughout the dissertation so far, then a novel additional source for ACD presents itself: one in which the ellipsis site occurs within a subordinate clause occupying a position within the extended verbal domain of the antecedent. All that is necessary for this to occur is for the subordinate clause in question to be located low enough in this domain that it falls within the span of structure that a VPE site would take as its antecedent.

As discussed in §2.5.1, subordinate clauses attach higher than Spec-VoiceInflP, but lower than Spec-TP. This means that a sufficiently large instance of VPE will indeed lead to an ACD configuration. This is illustrated in the tree diagram below. The solid green box indicates an instance of low-VPE, and the dashed green box indicates its corresponding low-VPE antecedent. The pair of red boxes represent the outcome of eliding anything higher: if we apply even the smallest instance of high-VPE, corresponding to the solid red box in the tree, then the dashed red box reflects the resulting antecedent, and it constitutes an ACD configuration:

(41) John will be dying to go because Mary will be ~~[dying to go]~~.



Thus, for sentences involving subordinated VPE sites such as those we have been considering in this section, an application of high-VPE would yield an ACD configuration, whereas low-VPE avoids this problem.

However, we know that low-VPE is always preferred in subordinated structures such as these—in other words, the depiction of high-VPE above is seemingly irrelevant since high-VPE never occurs when it would be subordinated beneath

its antecedent (as opposed to being coordinated with its antecedent). At the same time, to this point we have had no independent explanation for why this might be the case. In the next subsection, I argue that it is precisely this asymmetry with respect to ACD – high-VPE induces it in subordinated E-clauses whereas low-VPE does not – that leads to the grammar’s preference for low-VPE in these environments.

#### 2.5.4 The costs of antecedent-containment

For ease of reference, let us assign the two ellipsis sizes depicted in the tree in (41) their own example numbers in the form of the bracketed structures below:

*Low-VPE in subordinated E-clause: ACD configuration avoided*

- (42) John will [<sub>XP</sub> [<sub>YP</sub> leave] [ before Mary will ○ ]].  
 ○ = [<sub>YP</sub> leave ]

*High-VPE in subordinated E-clause: ACD configuration present*

- (43) John will [<sub>XP</sub> leave [ before Mary will ○ ]].  
 ○ = [<sub>XP</sub> leave [ before Mary will ○ ]]

Note once again that the ellipsis representation in (43) never actually surfaces in the grammar; all the evidence seen so far has suggested that low-VPE is the only VPE size tolerated in E-clauses subordinated beneath their A-clause. The central question we can now address is simply this: why should this be the case, given that (43) is syntactically well-formed – high-VPE is widespread – and given the existence of prototypical ACD configurations tolerated elsewhere?

I claim that (43) is ruled out precisely because it contains an ACD structure when another viable representation – that in (42) – does not. In other words, when faced with two possible ellipsis representations, one involving ACD and one not,

the grammar will always prefer the ACD-free representation (*mutatis mutandis*).

I state this in the form of a principle to be accounted for:

(44) **Principle of Antecedent-Containment Avoidance**

Prefer ellipsis representations involving antecedent-exclusion over those involving antecedent-containment.

This principle does not privilege low-VPE over high-VPE universally – after all, high-VPE is widespread, as we will see in particular in the next chapter – rather, high-VPE is dispreferred just when it would introduce an infinite regress problem that low-VPE would not.

Recalling the discussion in §2.5.2, an explanation for the principle in (44) emerges. We saw there that ACD configurations are inherently costly: they introduce an infinite regress problem that must be resolved before the structure in question can be properly interpreted. The resolution is an application of QR, regardless of whether the element in question would undergo QR independent of ACD or not. In other words, ACD configurations require a special operation that non-ACD configurations do not.

Therefore, I argue that the principle in (44) arises as a consequence of economy. The representation in (42) is more economical than (43), since the former can be interpreted without application of any special rule like QR, while the latter cannot. This economy-based explanation for (44) therefore rules out high-VPE in E-clauses subordinated beneath their A-clause.<sup>27</sup> At the same time, this explanation has nothing to say about low-VPE vs. high-VPE in prototypical ACD configurations,

---

<sup>27</sup>It is worth considering whether QR would even be a viable solution to the infinite regress problem introduced by subordination in a hypothetical scenario where high-VPE (and thus ACD) were somehow the only option. In such a case, would there be any application of QR that could avoid infinite regress? Recall that the analysis of prototypical ACD configurations involves QR of a structure containing the VPE site to a position outside of the antecedent. The analogous state of affairs here would presumably involve QR of the entire subordinate clause to a position at least as high as TP (higher than the upper-bound of high-VPE) in the superordinate clause. We might ask whether particular subordinate clause types independently require QR—those with decidedly quantificational or comparative qualities, for example, as might be the case with temporals like *before/after*. However, QR is generally thought to arise strictly for type-theoretic concerns, i.e. to resolve a type mismatch, and it is not clear what the nature of this mismatch would be with subordinate clauses.

since the two are equally costly (both introduce infinite regress); likewise, it has nothing to say about low-VPE vs. high-VPE when ACD would not be involved at all (e.g. when E-clause and A-clause are simply coordinated). Given that high-VPE is preferred with coordination (recall that VMM is impossible there), we must assume that high-VPE is the preferred VPE representation by default, and is only ruled out when a less costly alternative representation is provided by low-VPE. This explains why low-VPE always, and apparently only, arises in subordinated contexts.

Recall that this section began with an idealization of the data such that all E-clauses were related to their A-clause by either coordination or subordination. We know of course that this need not be the case on the surface, since VPE can freely apply across sentence and speaker boundaries. I mentioned briefly in fn. 20 that such cases might conceivably be reducible to the same basic coordination/subordination distinction we have discussed throughout: in a structured approach to modeling discourse, it may be that all E-clauses are either coordinated with or subordinated underneath their A-clause, given the inherently anaphoric nature of ellipsis. Given the structural account given above for the principle in (44), a uniform treatment of these phenomena would require positing syntactic reality to what might otherwise be thought of as discourse-level coordination and subordination. In other words, it may be that instances of low-VPE applying across speaker boundaries arise to avoid an ACD configuration arising from a silent superordinate A-clause, under which the surface E-clause is syntactically subordinated. This sort of approach raises many intriguing questions, but I will not attempt to address any of them in depth here. I leave this to future research.

## 2.6 Summary: On the uniform size of low-VPE

In the preceding discussion, I attempted to diagnose the precise size of the constituent elided in low-VPE, focusing on patterns of (non-)elidable auxiliaries. These patterns led to the claim that low-VPE targets a single projection just below the head containing the clause’s grammatical voice features. Since the projection being targeted is also related to grammatical voice, this necessitated a novel view of this part of the inflectional domain: one that involves Voice shells. These Voice shells comprise a higher projection associated with the clausal syntax of grammatical voice (VoiceSynP), and a lower, novel projection associated with the morphology of voice (VoiceInflP). In describing the empirical profile of low-VPE, I paid particular attention to the apparent variability of VPE size with respect to auxiliaries generated in VoiceInfl<sup>0</sup>. I argued that this variability in fact did not reflect differences in VPE sizes (unlike the low- vs. high-VPE distinction established in chapter 1); instead, it reflected the presence or absence of short head movement out of the ellipsis site. This approach allows us to analyze low-VPE as having a uniform domain of application: VoiceInflP. With this, we are able to resolve two longstanding bugbears in the theory of English VPE: the obligatory ellipsis of *being*, and the optional ellipsis of other auxiliaries. Finally, while somewhat surprising, the existence of a fundamental distinction in VPE size was seen to follow directly from the environments they are found in: low-VPE arises in environments where ellipsis of anything higher would lead to an ACD configuration. Since resolving the infinite regress problem borne by ACD structures is inherently costly, concerns of economy demand low-VPE in such cases.

Beyond the conceptual appeal of a uniform ellipsis site in low-VPE – it is parsimonious – this approach is consistent with an observation about this region of the clause that goes well beyond what VPE alone can reveal: namely, that the constituent I have identified here as VoiceInflP has a “privileged” status in



the syntax with respect to a large number of different phenomena, including VP topicalization, participle preposing, pseudoclefting, etc. I explore this in depth in chapter 4.

At this point it is worth taking a step back to review the findings reached so far. In particular, it will be worthwhile to abstract away from the analytical details of the inflectional model used here. After all, it may be that a lowering-like approach to the INFL domain such as the one I appeal to here (based on Bjorkman 2011) is the wrong choice; perhaps a model based on generalized auxiliary-raising, as in Harwood (2013b), is better suited to the task. If that turns out to be true, what would remain of the findings put forth here?

A great deal would remain, it turns out. The guiding force behind the discussion so far has been the collection of empirical facts laid out in the first chapter which diagnose a fundamental, binary distinction in VPE size along different configurations. The labels given to the projections involved are not crucial to the analysis; what is crucial is that the analysis accounts for the decidedly structural nature of the distinction. The details of low-VPE – the subject of the present chapter – should follow similarly: even if the INFL domain is properly modeled in a very different way than what is assumed here, the facts discussed above reveal that VPE has a very particular set of properties when it has the capacity (by way of its small size) to allow VMM. This set of properties lends itself, I believe, to a uniform ellipsis size in low-VPE, and I have provided arguments implicating short head movement as the factor responsible for apparent exceptions to this uniform low-VPE size. In the end, the labels and individual syntactic projections involved are secondary to the core observations described here.

In the next chapter, I turn to the other fundamental VPE size described in chapter 1, namely high-VPE, which apparently resists the uniform structural approach taken to low-VPE here. Later, though, in chapter 4, I return to the constituent targeted by low-VPE, and lay out a great deal of evidence independent

of VPE showing that it has a privileged status in the grammar. This will provide some explanation for why low-VPE should be apparently constrained to eliding the same XP each time, despite other derivational variables.

## CHAPTER 3

# Case Studies in High-VPE: Retorts and Tag Questions

### 3.1 Introduction

The discussion in chapter 1 revealed that VPE was subject to principled variation in its size. It was hypothesized that low-VPE always targeted a particular constituent in the inflectional domain, namely VoiceInflP; the results of the previous chapter support that position. However, we also saw in chapter 1 that its counterpart, high-VPE, is not so tightly constrained: it appears to be capable of eliding VoiceSynP or any phrasal category dominating it, up to but not including TP ( $T^0$  being the licenser of VPE: Lobeck 1995). In this chapter, we will take a closer look at high-VPE, with an eye toward gaining a better understanding of its properties and those of the environments that license it.

The forthcoming investigation of high-VPE will take the form of a case study detailing the behavior of two types of utterances, namely retorts (polarity-reversing assertions) and tag questions. We will see that these seemingly-distinct phenomena are united in demonstrating the hallmark properties of high-VPE with respect to the diagnostics discussed in the previous chapters. Interestingly, we will also see that these environments share properties in common beyond VPE. Although some questions will remain about the nature of high-VPE, I suggest that these shared properties perhaps identify some natural class common to all the E-A configurations that license high-VPE.

First, I consider retorts in detail; then, I compare them to the behavior of tag questions. Their similarities will lead us to a consideration of how these two types of utterances bear on broader matters concerning high-VPE.

## 3.2 Retorts

### 3.2.1 Introduction

Suppose an English speaker asserts either (1a) or (1b), depending on what she perceives to be the facts about the situation she is describing:

- (1) a. John hasn't left.
- b. John has left.

Now imagine that her interlocutor does not accept the truth of this assertion. Unwilling to admit the content of (1) into the common ground, he might choose to respond with an assertion that reverses its polarity. Farkas and Bruce (2010:100,105) define such *reversing assertions* as anaphoric conversational moves that “commit their author to the complement of the denotation of a declarative sentence” in the immediately preceding discourse.<sup>1</sup> Such assertions can take several forms; among these, we can identify at least two types.

The first class of reversing assertions is systematically insensitive to the morphosyntactic form of the assertion it reverses, as in (2). That is, any of the following would be a felicitous follow-up to (1), regardless of whether (a) or (b) were uttered as the initial assertion:

#### *Polarity-insensitive reversing assertions*

---

<sup>1</sup>Also see Farkas and Bruce (2010) on assertions that only reverse a subpart of the denotation of their antecedent (e.g. *A: John ate chicken. B: No, it was beef.*), which they refer to as *partial reversals/denials*. I leave such assertions aside entirely.

- (2) You're wrong! / That's not true! / I don't believe you! / etc.

By contrast, the other class of reversing assertion exhibits a strict sensitivity to the morphosyntactic form of the assertion it denies, including, but not limited to, its polarity: such reversing assertions are *morphosyntactically dependent*:

- (3) **Definition:** *Morphosyntactic Dependence* – A dependence in morphosyntactic form between an anaphoric utterance type and its antecedent.

This is to be contrasted with utterance types which are anaphoric, but whose forms do not covary with their antecedents, as in (2).

By their nature, morphosyntactically-dependent reversing assertions involve contrastive polarity. In English, this involves a pitch accent (indicated with SMALL CAPS) expressing contrastive focus associated with some expression of polarity—one that is the reverse of the antecedent assertion's polarity:<sup>2</sup>

- (4) a. He actually HAS(N'T) (left).  
b. But he HAS(N'T) (left)!  
c. On the contrary, he HAS(N'T) (left).  
...etc.

Importantly, what counts as an “expression of polarity” is not necessarily straightforward in a language like English. Unlike (emphatic) negation, which has its own morphology in English, emphatic affirmation is often realized simply as contrastive focus on material occupying T<sup>0</sup> – an observation going back at least as early as Chomsky (1957:65) – and can be a trigger for *do*-support in that capacity (see §3.2.3 and Laka 1990:ch. 2).

---

<sup>2</sup>This description of reversing assertions does not extend to instances of metalinguistic negation (Horn 1989:ch. 6), which have different properties. See below for discussion.

However, in special circumstances (to be made clear), affirmative contrastive polarity in English has its own special morphology: namely, the affirmative polarity particles *too/so* (Klima 1964:257).<sup>3</sup> In other affirmative reversing assertions, the left-edge polarity particle *yes* is used. What these reversing assertions have in common, along with their negative counterparts, is their use of polarity particles—special morphemes expressing polarity in responsive (often contrastive) contexts (see below on the status of *not*). As we will see, the class of morphosyntactically-dependent reversing assertions that make use of polarity particles behave differently than those in (4) which do not. I refer to the class involving polarity particles as *retorts*, and draw a distinction within this class on the basis of the surface position of the polarity particle (in bold):<sup>4</sup>

*Clause-Internal Polarity (CIP) retorts*

- (5) a. He has {**TOO/so**}!  
 b. He has **NOT**!

*Left-Edge Polarity (LEP) retorts*

- (6) a. **YES** he has!  
 b. **NO** he hasn't!

I refer to the reversing strategy exemplified in (5) as the *Clause-Internal Polarity* (CIP) retort strategy, reflecting the surface position of the polarity particles

---

<sup>3</sup>Klima only discusses *so*, but, in present-day American English, the contrastive polarity use of *too* has exactly the same distribution as its counterpart *so*. In fact, my own impression is that *too* is significantly more frequent than *so* in these contexts, at least among younger speakers of American English (though this should be examined empirically). I refer to them collectively as *too/so* throughout, and any claims made about one should be taken to apply to the other; that is, I regard them as synonymous in this context.

<sup>4</sup>Here I have illustrated LEP retorts as involving a pitch accent on the polarity particle, but this is only one of two possible prosodic realizations of LEP retorts. The other realization involves a pitch accent on the clause-internal expression of polarity. I discuss the intonational properties of retorts at length in §3.2.2.4; until then, though, I will simply represent LEP retorts with the pattern seen here.

TOO/SO and NOT.<sup>5</sup> This contrasts with the distribution of the polarity particles *yes* and *no* in (6), whose clause-initial position leads me to refer to this strategy as the *Left-Edge Polarity* (LEP) retort strategy. (For the remainder of this chapter, assume that either (1a) or (1b) is in the immediate discourse context of any exemplified retort.)

Shortly, we will look at some noteworthy properties of retorts, both of the CIP variety and the LEP variety. First, though, some background on the analysis of polarity is called for.

### 3.2.1.1 On $\Sigma$

Laka (1990) is the earliest and most thorough attempt to provide a formal syntax of polarity (and not simply negation, for which there is an ancient formal tradition: see Horn 1989 for an exhaustive discussion of this tradition). Using data from English, Spanish, and Basque, Laka argues that the morphosyntactic locus of polarity – both negative and affirmative – lies within a dedicated syntactic projection she calls  $\Sigma$ P. This projection, which is immediately adjacent to TP,<sup>6</sup> is therefore the base position for not only sentential negation, but also for its affirmative counterpart, which typically lacks morphological exponence crosslinguistically.

As we have already seen, though, there are circumstances in which affirmative polarity has exponence, as Laka (1990:ch. 2) discusses at length. Specifically, Laka identifies two distinct “flavors” of affirmative  $\Sigma$ , which are both emphatic, but licensed in different environments. The first variety is manifested by Basque

---

<sup>5</sup>See below for arguments that this occurrence of ‘NOT’ is a genuine negative analogue of TOO/SO: that is, it is a specialized marker of negative polarity in reversing contexts. It is homophonous with, but lexically distinct from, the sentential negation morpheme in English. See Laka (1990:§2.5.1) for a brief discussion of affirmative CIP retorts in both English and Basque.

<sup>6</sup>Laka argues that the relative position of  $\Sigma$ P with respect to TP is subject to parametric variation, with  $\Sigma$ P below TP in English, but above TP in Basque. This is not critical to the present discussion, so I leave it aside. See also Haddican (2004).

*ba* and English *too/so*, which are only licensed in reversals of negative assertions (specifically, in CIP retorts: (5)). The other variety of affirmative  $\Sigma$  Laka discusses is what she calls “[<sub>aff</sub>]” (or occasionally just “aff”), whose only phonological content is stress (i.e. a floating stress morpheme; see below). In English, this stress morpheme is realized on (i.e., it *docks to*) material in  $T^0$ , triggering *do*-support in the absence of a modal or raised auxiliary there (*op. cit.*:§2.2, following early observations in Chomsky 1957:65 and Klima 1964:257).<sup>7</sup> This variety of  $\Sigma$  is licensed in all non-retort contexts involving contrastive affirmative polarity. A summary of Laka’s taxonomy for  $\Sigma$  is below (adapted from *op. cit.*:106):<sup>8</sup>

*Laka’s (1990) lexical entries for  $\Sigma$*

(7)	<b>Polarity</b>	<b>Basque</b>	<b>English</b>
	Negative	<i>ez</i>	<i>not (-n’t)</i>
	Affirmative (non-retort)	[ <sub>aff</sub> ]	[ <sub>aff</sub> ]
	Affirmative (retort)	<i>ba</i>	<i>too/so</i>

For Laka, then,  $\Sigma$  is sensitive to the retort/non-retort distinction in the affirmative, but apparently not in the negative.

As the discussion of retorts progresses, I will argue for a different state of affairs. First, I will argue that most, perhaps all, expressions of polarity in these languages are generated in the Specifier of  $\Sigma P$ , not in  $\Sigma^0$ . Second, I will argue that the floating stress morpheme Laka identifies as [<sub>aff</sub>] is not specified for polarity; it simply arises whenever  $\Sigma$  is focused (represented here as [ ' ]; cf. fn. 35), and it docks onto whatever is to its left (at a particular stage in the derivation), regardless of whether that host is itself specified negative or affirmative. Third, I

<sup>7</sup>The realization of [<sub>aff</sub>] in Basque is more complicated: see Laka (1990:105, 140), where it is suggested that [<sub>aff</sub>]’s floating stress morpheme is realized on the subject in Basque. I leave this aside. See also Hualde and Ortiz de Urbina (2003:§4.4.3.2.2).

<sup>8</sup>Laka (1990:§2.7.4) claims that the responsive particles *yes* and *no* are not manifestations of  $\Sigma$ , but rather are complementizers specified for polarity. Thus, they are not included in this table. See below for arguments that such particles are generated in  $\Sigma(P)$ .



will argue that negative polarity is sensitive to the retort/non-retort distinction, but that this distinction is superficially obscured by homophony with the standard negative marker; to disambiguate this homophony, I will represent the *not* used in CIP retorts as ‘*not<sub>R</sub>*’. Finally, en route to a unified analysis of retorts, I will argue that *yes* and *no* (in both LEP retorts and in responses to polar questions, perhaps along with embedded *so* and *not*) are also generated in Spec- $\Sigma$ P, before moving to a position at the edge of the clause. Thus, a preview of my analysis for elements in  $\Sigma$ P is below (leaving aside Basque):

*Elements argued here to originate in  $\Sigma P$*

(8) Polarity type	Affirmative	Negative
Neutral/default	$\emptyset$	<i>not</i> (- <i>n't</i> )
Focused	[ ' ]	[ ' ]
Reversing (CIP)	<i>too/so</i>	<i>not<sub>R</sub></i> (*- <i>n't</i> )
Reversing (LEP)	<i>yes</i>	<i>no</i>
Responsive (matrix)	<i>yes</i>	<i>no</i>
Responsive (embedded)	<i>so</i>	<i>not</i>

These arguments are scattered throughout the discussion below; see §3.2.3 for their formalization. For a recent summary of various authors' analytical invocation of  $\Sigma P$  (and equivalent polarity projections), see Wood (2008:§2.1).

### 3.2.1.2 On *yes* vs. *yeah* and *no* vs. *nah*

Polar response particles such as *yes* and *no* have garnered much recent interest in the literature. For example, the present consensus regarding their syntax seems to lie with the proposal put forth in Kramer and Rawlins (2011), who argue that bare *yes* and *no* utterances reflect full clausal structures that have undergone ellipsis, leaving behind only these particles, which survive ellipsis by way of occupying a high left-edge position in the clausal structure. On the interpretive side, recent work by Farkas and Roelofsen (To appear) (building on earlier work by Farkas and Bruce 2010) develops a typology of polar responses by decomposing their semantic and pragmatic contributions into atomic features ([+] vs. [-], and [SAME] vs. [REVERSE]).

While these (and other) recent contributions have rightfully considered polar response particles in different contexts – e.g. as responses to neutral questions and biased questions (qua negative questions and rising declaratives), agreeing with

or rejecting prior assertions, etc. – surprisingly little has been said about their colloquial counterparts (crosslinguistic analogues of which appear to be common).<sup>9</sup> Here, I will say a few brief words about colloquial polar response particles in American English. A non-exhaustive list is below:<sup>10</sup>

(9) **Affirmative:** *yeah*, *yep*, *yup*, *mm-hmm* (see fn. 10), etc.

**Negative:** *nah* [nã] (also [nã]), *nope*, *mm-mm* [ʔm,ʔm] (cf. fn. 10), etc.

In particular, I will focus on *yeah* [jæ] and *nah* [nã]. We will quickly see that these are not simply non-standard pronunciations of *yes* and *no*.

Consider their distribution as answers to polar questions. Whereas *yeah* seems to be available in any answering context that would normally license *yes*, this does not hold for *nah* with respect to *no*:

(10) Q: Is today Tuesday?

A1: Yes. / Yeah.

A2: No. / ??Nah.

This is not to say that *nah* can never be used as a response to a polar question. *Nah* seems to carry some additional meaning beyond what is normally conveyed by responsive *no* (see Farkas and Roelofsen to appear on the meaning of *no*). Impressionistically, this additional meaning involves speaker attitude (and is thus

---

<sup>9</sup>For discussion of colloquial polar response particles from a (socio-)phonetic perspective, see Bolinger (1946), Ward (2006), a.o. From a syntactic and/or semantic perspective, I am aware of only passing remarks, mostly about *yeah*: see Pope (1972:196), Farkas and Bruce (2010:115), Holmberg (2012: fn. 9), and Farkas and Roelofsen (To appear: fn. 23).

<sup>10</sup>A particularly interesting pair is the affirmative *uh-huh* [ʔã'hã] and a negative counterpart not listed in (9), namely *huh-uh* [hã,ʔã] (Bolinger 1946:92; Ward 2006). They comprise the same two syllables in inverse order, but with the same stress assignment (i.e., primary stress on the syllable with the voiceless onset, [hã]), and perhaps the same pitch assignments as well (with [hã] receiving higher pitch than [ʔã], though this may simply be a correlate of stress). A non-oral counterpart to this pair is also attested: affirmative *mm-hmm* [ʔm'mm] and negative *hmm-mm* [mm,ʔm]. I leave a decompositional analysis of these pairs to a braver linguist.

not-at-issue). It is difficult to characterize the nature of this attitude precisely, but *nah* responses seem to arise most naturally in environments where the speaker is providing information about his or her own mental state or participation in an event (something the question in (10) clearly does not facilitate). Consider the following minimal set:

- (11) a. Q: Do you think John is cute?  
A: Nah.
- b. Q: Does Mary think John is cute?  
A: ??Nah.
- (12) a. Q: Are you coming to the party tonight?  
A: Nah.
- b. Q: Is John coming to the party tonight?  
A: ??Nah.

When the question concerns a third party's actions or mental state, *nah* is degraded: intuitively, it feels as though the speaker is answering on behalf of the third party – something that, in most situations, the speaker does not have the authority to do – rather than simply reporting what he or she believes to be true of that third party. In circumstances where the speaker does have such authority, the result is much improved:

- (13) Q: Is your husband coming to the party tonight?  
A: Nah.

Consistent with the claim that *nah* carries attitudinal meaning absent from *yeah*, consider their differential behavior in reported/indirect speech contexts:<sup>11</sup>

---

<sup>11</sup>*Nah* is of course fine in quoted/direct speech, e.g. *She said, "nah, I'm not"*.

- (14) I asked Mary if she was hungry, and...
- a. ...she said yeah, she was.
  - b. ??...she said nah, she wasn't.

Thus, *nah* exhibits some properties in common with speaker-oriented adverbs (of particular note given its negative semantics, since speaker-oriented adverbs (e.g. *luckily*) have been argued to be positive polarity items: see Ernst 2009, a.o.). This might indicate a syntactic connection, e.g. an interaction between *nah* and the left-peripheral head(s) responsible for this variety of meaning. I leave this open.

This general pattern continues in other contexts that license *yes* and *no*, e.g. in interjection/exclamation-like utterances, and utterances expressing confirmation of or agreement with an antecedent assertion of the same polarity<sup>12</sup> (see Farkas and Roelofsen to appear). While *yeah* can appear freely in such contexts, *nah* cannot:

- (15) a. [*Context: Speaker's favorite team scores a goal*]  
 Yes!! / Yeah!!
- b. [*Context: Speaker's favorite team gets scored on*]  
 No!! / #Nah!!
- (16) A: The door is locked.  
 B1: Yes, it is.  
 B2: Yeah, it is.
- (17) A: The door isn't locked.

---

<sup>12</sup>In utterances confirming/agreeing with assertions of opposite polarity from the particle, the pattern changes. As Farkas and Roelofsen (To appear) note, *no* cannot be used with an affirmative antecedent in that context (*A: Peter called. B: \*No* (intended: he called)). The same holds for *nah*. However, when the antecedent assertion is negative, I find *yes* to be highly degraded as a confirmation, contra Brasoveanu et al. (2011) (*A: Peter didn't call. B: \*Yes* (intended: he didn't call)), while *yeah* is perfectly acceptable in this context (*A: Peter didn't call. B: Yeah* (intended: he didn't call)). I also greatly prefer *yeah* to *yes* as an affirmative (reversing) answer to a negative question; see Thoms (2012b), Holmberg (2013), Farkas and Roelofsen (To appear), a.o.

B1: No, it isn't.

B2: ??Nah, it isn't.

Importantly, though, *nah* is completely impossible in LEP retorts (regardless of whether the intonational contour typical of retorts is realized peripherally or internally: see §3.2.2.4). The same is not true of *yeah*, which, to my ear, is only slightly degraded in LEP retorts (most prominently so with internal pitch accent placement; again, see §3.2.2.4):

(18) A: The door isn't locked.

B1: (?)YEAH it is!

B2: ?Yeah it IS!

(19) A: The door is locked.

B: \*Nah it isn't!

For all of the above reasons, it seems clear that *yeah* and, in particular, *nah* have importantly different properties than their standard counterparts; likewise, *yeah* and *nah* differ from each other, too, in more than just their polarity. I leave formal analysis of these differences to future work.

### 3.2.2 Characteristic properties of retorts

CIP and LEP retorts are unified in their pragmatics: they are licensed in the same sorts of contexts, and their discourse functions are identical. They are also unified in making use of a polarity particle, with LEP retorts taking *yes* or *no*, and CIP retorts taking *too/so* or *not<sub>R</sub>*.<sup>13</sup> In this subsection, I show that the two strategies are also unified in other ways, including aspects of their morphosyntax

---

<sup>13</sup>Although *yes* and *no* are not unique to (LEP) retorts – they can of course be used in responses to polar questions, as well – the morphemes *too/so* and *not<sub>R</sub>*, in their polarity reversing capacity, are only licensed in (CIP) retorts. For a possible exception to this claim, see Wood's (2008) discussion of inverted *so*.

(despite appearances) and their prosody. I take these shared properties to suggest a common syntactic source for retorts, leading in the next section to an analysis that relates the polarity particles of each type, i.e. *yes* and *too/so* on the one hand, and *no* and *not<sub>R</sub>* on the other.

### 3.2.2.1 Neophobia

The content of both CIP and LEP retorts must be entirely discourse-old, with the sole exception of the polarity particle. Put differently, retorts cannot introduce discourse-new (non-polar) information,<sup>14</sup> a property I refer to as *neophobia*. I discuss some effects of neophobia in retorts here, but retorts are hardly the only neophobic phenomenon: see, for example, tag questions (§3.3).

Attempting to contribute new information to the discourse by, for instance, coordinating the subject in a retort is strictly prohibited, though it can be done in non-retort reversing assertions, i.e. those not involving a polarity particle, as in (4) above (henceforth *non-retort RAs*):

- (20) A: John hasn't gotten back from vacation yet.  
 B1: \*YES (both) he and Mary have! *LEP*  
 B2: \*[(Both) he and Mary have TOO/SO!] *CIP*  
 B3: (Both) he and Mary HAVE, actually. *Non-retort RA*

Adverbials are also impossible in retorts of both types, even including those that might otherwise serve to strengthen the denial (e.g. by expressing counter-expectational mood (*actually*) or a strong evidential source (*definitely*), etc.).

---

<sup>14</sup>“Discourse-new information” may require some additional qualification, e.g. “discourse-new at-issue meaning”, since, for example, whatever not-at-issue meaning is conveyed by *fucking* in the following example does not run afoul of the neophobia condition (see also §3.2.2.2):

- (i) A: You didn't do the dishes.  
 B1: Yes I fucking did! *LEP*  
 B2: I <fucking> did <fucking> TOO/SO! *CIP*

These are fine in non-retort RAs, however:<sup>15</sup>

- (21) a. \*YES he {actually/definitely/really/etc.} has! *LEP*  
b. \*He {actually/definitely/really/etc.} has TOO/SO! *CIP*  
c. He {actually/definitely/really/etc.} HAS. *Non-retort RA*

Even especially high adverbials that would normally be initial but within the same Intonation Phrase as the rest of the clause are ill-formed in retorts:

- (22) a. \*Apparently YES he has! *LEP*  
b. \*Apparently he has TOO/SO! *CIP*  
c. Apparently he HAS. *Non-retort RA*

However, in clause-final position, retorts with *apparently*, *actually*, etc. improve:

- (23) a. YES he has, apparently. *LEP*  
b. He has TOO/SO, apparently. *CIP*  
c. He HAS, apparently. *Non-retort RA*

Crucially, though, such adverbials can only appear clause-finally when they are separated from the rest of the clause by a prosodic break, even in unmarked declarative contexts (see §3.2.2.4). This directly correlates with their inability to take nuclear stress in final position, a fact which Stowell (in progress), following the framework for adverbial syntax laid out in Cinque (1999), takes to diagnose the presence of an independent elliptical clause containing the adverbial. The data in (23) would seem to support that conclusion, since the neophobia constraint that otherwise characterizes retorts is apparently overcome only when such adverbials

---

<sup>15</sup>The string in (21a) is less degraded if a strong prosodic break is present following the polarity particle; however, manipulating the prosody in this way has several other non-trivial consequences that ultimately disqualify such utterances as retorts. For more, see §3.2.2.4.



are construed clause-finally, in their own Intonation Phrase. Thus, it seems more likely that neophobia is in fact respected in such cases, since the new material (the adverb) is contained in a distinct clause from the retort.

### 3.2.2.2 Preferential pronominalization

Retorts also exhibit a strong preference for their subjects to be pronominalized, with non-pronominal DPs judged highly redundant to the point of aberrancy.<sup>16</sup>

This is apparently not the case for at least some non-retort RAs:

- (24) A: John {will / won't} adopt a baby wombat.  
 B1: \*YES John will! / \*NO John won't! *LEP*  
 B2: \*John will TOO/SO! / \*John will NOT! *CIP*  
 B3: On the contrary, John actually {WON'T / WILL}. *Non-retort RA*

In my discussion of neophobia in the previous subsection, I equivocated between describing it as requiring discourse-old material on the one hand versus prohibiting discourse-new material on the other. The preference for pronominalization seen here suggests that the former characterization is more accurate,<sup>17</sup> and it potentially allows for a deeper connection to be made between neophobia and the pattern exemplified in (24). That is, if neophobia is properly stated as enforcing maximal givenness, then it is at least compatible with, and perhaps even includes, a provision that given material be encoded anaphorically whenever possible (and perhaps as anaphorically as possible on some relevant scale: see Rezac 2013, following Sauerland 2005 and Wagner 2006).

Another intersection between preferential pronominalization and neophobia involves epithet subjects in retorts. Their acceptability varies across speakers (repre-

<sup>16</sup>Whether this intuition would be better represented with a judgment mark of #, rather than the \* used here, is a matter I leave aside.

<sup>17</sup>This would perhaps make *retrophilia* a more apt name for the phenomenon.

sented with the judgment mark %), with reports ranging from ‘slightly degraded’ to ‘unacceptable’:

- (25) A: John never gossips.  
B1: %YES that idiot does! *LEP*  
B2: %The son of a bitch does TOO/SO! *CIP*

This variability may reflect cross-idiolectal differences in the extent to which epithets carry some additional meaning beyond their reference (e.g. an evaluative implicature). Perhaps, for some speakers, this extra meaning is sufficient to constitute a discourse-new contribution, running afoul of the neophobia constraint and leading to deviance. Alternatively, perhaps there is speaker variability with respect to the nature of the additional meaning carried by epithets: for some, this meaning might be entirely not-at-issue, and thus somehow irrelevant to the neophobia constraint, while for others it might contain some at-issue content, violating the constraint. See fn. 14 for an example of not-at-issue content that apparently does not violate neophobia in retorts. See §3.3 for further discussion.

### 3.2.2.3 Preferential ellipsis

The careful reader will have noticed that, to this point, all examples of retorts presented so far are missing a main predicate. Similar to the preference for pronominalization seen in (24), retorts preferentially involve VPE as well.<sup>18</sup> In particular, LEP retorts with unelided predicates border on ungrammaticality, while non-retort RAs are acceptable.<sup>19</sup>

---

<sup>18</sup>See van Craenenbroeck (2010:§11.2) for (indirect) arguments that at least English LEP retorts involve VPE (rather than a verbal proform). CIP retorts pattern the same with respect to his diagnostics.

<sup>19</sup>This contradicts judgments reported earlier in Sailor (2011, 2012c), where I assumed that VPE was optional in both retort types. Further introspection and consultation with other native speakers has revealed the error in this assumption. Note that canonical retort intonation is crucial in bringing this judgment out: see §3.2.2.4. Also see §3.2.4 for some discussion of the

- (26) A: John left yesterday.  
 B1: ??No he didn't leave {yesterday / then}! *LEP*  
 B2: He actually DIDN'T leave yesterday. *Non-retort RA*

This preference relates directly to the discussion in the previous subsection regarding (24). Given the well-known similarities between pronominals and VPE (see Lobeck 1995 for a detailed summary), presumably whatever discourse conditions enforce a preference for pronominalization in retorts have a similar effect with respect to VPE. If given material in retorts is preferentially realized anaphorically, as suggested above, and VPE is seen as falling on the same anaphoric spectrum as e.g. pronominalization, then the pattern in (26) falls in line with earlier generalizations: see Kehler (2002:45), and §3.3 below for further discussion.

### 3.2.2.4 The intonation of retorts

Besides the morphosyntactic properties described here, retorts are also characterized by their intonation.

Foremost, retorts of both types involve only one Intonation Phrase (IntP). The string in (6) corresponding to the LEP strategy can be pronounced with a prosodic break following the left-edge polarity particle (indicated by a comma); however, the result differs in several ways from its single-IntP retort counterpart:

- (27) Yes, he has!

First, (27) is degraded as a denial of (1a),<sup>20</sup> so it is not a retort. Laka (1990:§2.7.4) argues that polarity particles such as *yes* and *no*, when contained inside their own

---

difference between CIP and LEP retorts with respect to VPE.

<sup>20</sup>In their discussion of reversing assertions, Farkas and Bruce (2010:§4.1) consistently place commas after the polarity particle in their examples. Since they do not discuss their prosody (cf. fn. 22), I assume this is simply an orthographic convention they have chosen to follow, and does not reflect pronunciation.

IntP, are the only surviving vestiges of full clausal structures that have undergone ellipsis. (If any material follows a *yes* or *no* contained in its own IntP – i.e., if there is a comma separating the polarity particle from some other clause – then that other clause is simply asyndetically coordinated with the polarity particle’s clause.) If this is correct, as I believe it is, then the infelicity of examples such as (27) as retorts follows from the fact that bare polarity particles are infelicitous as reversing assertions (Holmberg 2013:37). Coordinating a bare polarity particle with additional material should not, and does not, improve the result. This argument is even clearer in the negative form of (27) (stress placement is irrelevant here):

- (28) A: John hasn’t left.  
       B1: No, he has!  
       B2: \*No he has!

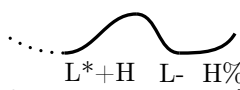
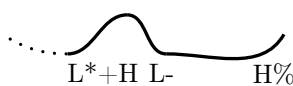
This pattern further suggests that a polarity particle contained within its own IntP, as in (28B1), has a different status than when it is incorporated into the same IntP as the rest of the utterance. When it occurs within its own IntP, it alone can express the [REVERSE] feature of Farkas and Bruce (2010), allowing the remaining material to convey the updated (reversed) polarity specification (which is affirmative in the case of (28B1)). This is clearly not possible when the polarity particle occurs within the same IntP as the remaining material, as in typical LEP retorts and (28B2). As we have seen, the *yes* particle is required for a well-formed LEP retort in such contexts.<sup>21</sup> This shows that intonation is crucial in this domain, and that the single-IntP utterances behave differently than those involving more

---

<sup>21</sup>Regarding the featural typology for polar responses in Farkas and Bruce (2010), a conclusion to be drawn from the above discussion is that retorts – single IntP reversing assertions of the sort I have been describing all along – cannot convey the [REVERSE,+] feature combination in English. A narrower statement of this conclusion is that retorts whose antecedents are negative assertions disallow the [REVERSE] use of *no*. Either *yes*, which conveys absolute polarity [+], must be used, or a non-retort RA involving two or more IntPs must be used.

than one. Therefore, the term *retort* here only refers to those reversing assertions involving a single IntP; others with more boundaries are set to the side as non-retort RAs.

As a second prosodic characteristic, observe that retorts of both types are realized with a [ L\*+H L- H% ] contour in ToBI formalism (Silverman et al. 1992; see fn. 26), with the pitch accent associating to an expression of polarity in the retort (with more than one option in LEP retorts, as we will see).<sup>22</sup> In the case of the CIP strategy, this is straightforward: there is only one marker of polarity in the retort.<sup>23</sup> To better illustrate the prosody in the examples below, I provide both short and long examples (the latter involving multiple post-focal stranded auxiliaries), with both affirmative and negative polarity. Since differing contexts would be necessary to license each member of these pairs, the following examples should not be treated as dialogues.<sup>24</sup>

- (29) a.  L\*+H L- H%  
He has TOO - O - O!
- b.  L\*+H L- H%  
He will NOT have been!

For the LEP strategy, though, there are two loci of polarity: the one expressed



<sup>22</sup>Though they do not offer a transcription, Farkas and Roelofsen (To appear: fn. 32) seem to be referring to this melody in their description of what they call the “smart Aleck” contour found with (what I call) LEP retorts (and, they claim, with certain reversing answers to polar questions, though I find their examples infelicitous with this contour). They go on to offer the intriguing suggestion that this contour realizes their [REVERSE] feature (*op. cit.*:§3.5), which, in other languages (e.g. Romanian), is realized as a discrete particle.

<sup>23</sup>The prosody of the CIP strategy likely also involves a pitch accent of some kind on T (e.g. *has* in (29a)), as there tends to be a slight rise and break immediately before the polarity particle in such retorts; I leave this aside.


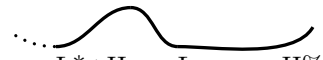
<sup>24</sup>In configurations such as (30a), in which the focused polarity marker occurs utterance-finally, the entire [ L\*+H L- H% ] melody is realized on that marker (i.e. *TOO*). In such cases, I elongate the orthographic representation of that marker purely to improve the legibility of the ToBI transcription and accompanying prosodic contour; this does not reflect any particular (morpho-)phonological claim (though such instances presumably do involve lengthening to accommodate the complex melody).

by the left-edge particle (*yes/no*) and the one expressed by the clause-internal  $\Sigma$  head (phonologically cliticized to T). Thus, the LEP retort strategy has two possible prosodic realizations, represented below:<sup>25</sup>

*LEP retort, pitch accent on left-edge polarity marker*

- (30) a.   
 L\*+H L- H%  
 NO he hasn't!
- b.   
 L\*+H L- H%  
 YES he should have been!

*LEP retort, pitch accent on clause-internal polarity (qua T<sup>0</sup>)*

- (31) a.   
 L\*+H L- H%  
 No he HAS - N'T!
- b.   
 L\*+H L- H%  
 Yes he SHOULD have been!

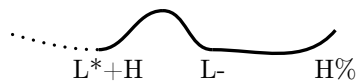
See §3.2.3 for further discussion of these two options for LEP pitch accent placement.

This intonational pattern is homophonous with the so-called “contradiction contour” originally identified in Liberman and Sag (1974), and further refined in Constant (2012:411). As the name indicates, this contour was originally described for examples of the following sort (adapted from Constant 2012:411):

<sup>25</sup>There may also be an additional (low) pitch accent on the T- $\Sigma$  complex in LEP retorts (see also fn. 23). If further investigation reveals an accent to be present there, readers are directed to an intriguing suggestion from van Craenenbroeck (2010:281 fn. 33) relating the presence of multiple focus pitch accents in similar Dutch examples to the presence of polarity concord of the sort I argue for (as does *ibid.*): see §3.2.3.

Note that complex tonal contours such as the L\*+H pitch accent can be (and often are) realized across more than one syllable or word. Thus, although L\*+H is associated with *No* in (30a), the H is typically realized on *he* (unsurprising, given the short duration of *No*).

(32) A: John finally managed to solve the problem.



B: He didn't "manage" to solve it – it was easy for him!

Retorts and utterances like (32B) share similar discourse functions. This might be taken to indicate that the intonational melody that characterizes retorts is in fact an instance of this contradiction contour.

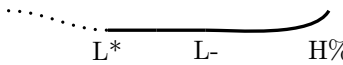
There may be reasons to doubt this, however. Constant (2012:§2.2) identifies some additional characteristics of this contradiction contour (in service of distinguishing it from what he calls the “rise-fall-rise” contour, which I leave aside), and not all of them apply to retorts. For example, Constant notes that the contradiction contour co-occurs with metalinguistic negation: in (32B), negation does not take its usual sentential scope; rather, it serves to signal objection to an implicature introduced by speaker A’s use of *manage* (Horn 1989:ch. 6). Put differently, speaker B quite deliberately does not commit himself to the proposition *He didn’t manage to solve it*; indeed, the contradiction contour helps to signal B’s rejection of a particular component of this utterance, but it does not add any semantic contribution independent of this.

This is not true of retorts: although they share with examples such as (32) a refusal to admit a prior assertion into the common ground, retorts also carry an independent semantic contribution – that is, as reversing assertions, they directly assert the rejected utterance’s polar opposite. This independent contribution distinguishes retorts from utterances bearing the contradiction contour such as (32). This either means that the semantic profile of the contradiction contour is broader than has been previously described, or that retorts bear a contour that is homophonous with, but ultimately distinct from, the contradiction contour that characterizes metalinguistic-negative utterances of the sort seen above.

There is some evidence in support of the second conclusion. Constant (2012:411)

goes on to note that the rising portion of the contradiction contour’s pitch accent is actually optional; i.e., the melody is more accurately characterized as [ L\*(+H) L- H% ]. Thus, example (32B) can also be felicitously rendered as (33B):

(33) B: He didn’t “manage” to solve it – it was easy for him!

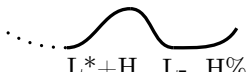


This holds across the board for utterances bearing the contradiction contour. However, this does not seem to hold for retorts: both CIP retorts (29) and LEP retorts (30)-(31) are degraded or infelicitous without the rising portion of the pitch accent (i.e., with just a L\* on the focused element) in American English.<sup>26</sup> Thus, there are valid empirical grounds for distinguishing Liberman and Sag’s (1974) contradiction contour and its associated utterances from retorts.

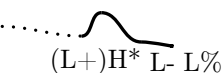
Finally, it is important to note that the [ L\*+H L- H% ] contour of retorts is largely infelicitous with non-retorts involving contrastive polarity, which instead must end in a falling contour:<sup>27</sup>

(34) John said he was planning to come, but in fact...

a. #...He was NO - O - OT!



b. ...he was NOT.



<sup>26</sup>At first glance, it seems Canadian and British varieties of English allow a [ L\* L- H% ] contour in what appear to be CIP retorts (e.g. (29)). Crucially, though, such examples fail to exhibit many of the other morphosyntactic properties of CIP retorts discussed above (and see below): for example, they allow the introduction of discourse-new material (underlined), contraction of the clause-internal polarity marker, etc.; likewise, they serve equally well as responses to yes/no questions (unlike canonical CIP retorts, as well as LEP retorts):

(i) A: The Queen has died. / Has the Queen died?  
 B: She most certainly HASN’T!  
 L\* L- H%

<sup>27</sup>For more on the distinction between contrastive polarity in non-retorts versus retorts, see e.g. van Craenenbroeck (2010:165-6).



Summarizing the preceding discussion, sentences with a particular morphosyntactic and pragmatic profile – what I am calling retorts – necessarily occur with a [ L\*+H L- H% ] intonational melody. While this melody is found elsewhere in English (see Constant 2012), it is not licensed simply in the presence of contrastive polarity; it requires a discourse of a particular sort – one which is closely related to the discourse contexts that license the contradiction contour, but without the necessity of metalinguistic negation. It seems likely that a deeper semantic-pragmatic connection could be established between these contours (perhaps accounting for the (non-)optionality of the rising portion of the pitch accent), but I will not explore this here. (See Constant 2012: fn. 7 for a similar sentiment.)

### 3.2.2.5 The root status of retorts

CIP and LEP retorts cannot be embedded, even under bridge verbs (note that intonation is crucial here):

*CIP and LEP retorts are root phenomena*

- (35) a. \*It's clear (that) he has TOO/SO!  
 b. #Mary thinks (that) he has NOT!
- (36) a. \*It's clear (that) YES he has!  
 b. \*Mary thinks (that) NO he hasn't!

As we have seen elsewhere, some of these strings are licit in other contexts (e.g. (36b)), but not in retort contexts (with retort intonation).

The root status of LEP retorts is perhaps unsurprising, given that *yes* and *no* are known independently to be restricted to main clauses in English (Laka 1990:162, Sailor 2012b, a.o.). The root status of CIP retorts is noteworthy, however. Trivially, emphatic negation can be embedded in non-retort contexts, as can emphatic affirmation (which, in non-retort contexts, is simply realized as a

segmentless stress morpheme, capable of triggering *do*-support; see §3.2.1.1):

- (37) a. John claims he mailed the package, but it's clear that he did NOT.  
b. John said he didn't eat all the pies, but Mary thinks he DID.

Thus, the root status of examples (35) and (36) is not a property of emphatic polarity in general; it appears to be specific to retorts. I therefore regard it as a core property of the phenomenon.

Certain types of emphatic polarity in other languages are similarly constrained to root contexts.<sup>28</sup> As Breitbarth et al. (2013:4) note, “the derivation of the [root emphatic polarity] phenomena in question has been argued to implicate (an operator in) a designated left-peripheral functional projection encoding focus on the polarity of the sentence in these constructions, possibly attracting a lower polarity projection [...] The restriction to root clauses and a subset of embedded clauses is accounted for by whatever account is invoked to account for the restricted distribution of other main clause phenomena” (see references in *op. cit.*). Elsewhere, Haegeman (2012) has argued for an “intervention” analysis of root phenomena. The premise for such an account is that non-root-like embedded contexts involve operators at their left edges (e.g. a factive operator), the presence of which blocks other left-edge operators that are an essential part of the derivation of what we know of as root phenomena (e.g. topicalization, inversion, etc.). That is, the embedded operator triggers an intervention effect for other left-edge operators, such as those involved in the derivation of root phenomena. Extending this sort of analysis to retorts may explain their root status, but it may also help explain more general properties of the phenomenon as well. I return to this shortly.

---

<sup>28</sup>The (non-)root status of emphatic polarity across languages is a topic that has recently garnered a great deal of attention, with a special issue of *Lingua* (2013, volume 128) contributing much to its empirical and analytical profile. See the introduction to that special issue, Breitbarth et al. (2013), for a summary.

### 3.2.3 The derivation of retorts

Before getting to a unified analysis of the two retort types described above,<sup>29</sup> one final observation is crucial: the polarity particles that characterize the CIP and LEP retort types are in complementary distribution. That is, a retort cannot involve both a left-edge polarity particle (e.g. *yes* and *no*) and a clause-internal polarity particle (e.g. *too/so* and *not<sub>R</sub>*):<sup>30</sup>

(38) \*Yes he has TOO/SO!

This is the puzzle to be addressed: what is the source of this co-occurrence restriction?

The only obvious difference between the two retort types is in the form and position of their polarity particles. It is therefore reasonable to interpret the complementary distribution of the two retort strategies as reflecting an underlying complementary distribution in these particles. In other words, *yes* and *too/so* are in competition (as are *no* and *not<sub>R</sub>*, despite appearances; see below), not the LEP and CIP retort strategies themselves.

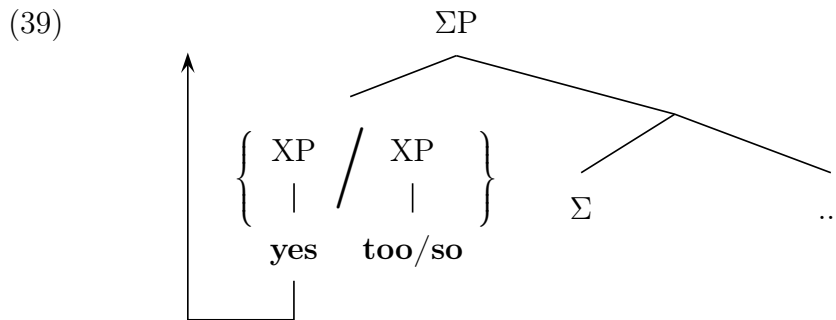
In keeping with classic syntactic reasoning, I take the complementary distribution of these polarity particles to indicate competition for a single syntactic position. Since only *yes* and *no* end up at the left edge (and since lowering and rightward movement are conceptually dispreferred), it follows that the contested position is the one occupied on the surface by *too/so* and *not<sub>R</sub>* in CIP retorts. I argue that left-edge polarity particles begin in this clause-internal position in LEP retorts (thus blocking merger of e.g. *too/so* and *not<sub>R</sub>*), eventually moving to

---

<sup>29</sup>My attempt to unify these retorts is similar in spirit to that of van Craenenbroeck (2010:ch. 14), who proposes a unified analysis of two seemingly-distinct denial strategies in (varieties of) Dutch. Beyond this (and a few technical details), though, the similarities end: the Dutch denials he discusses have very different syntactic properties than the retorts under discussion here; thus, I leave them aside.

<sup>30</sup>Superficially, it appears that the two can co-occur in the negative, given the (relative) acceptability of *?No he has NOT!*; but, below, I show that this is illusory: this *not* is not *not<sub>R</sub>*.

the left edge of the clause (contra Laka 1990:§2.7.4: see fn. 8). As all of these morphemes are specified for polarity, it further stands to reason that this contested clause-internal position is within  $\Sigma$ P.<sup>31</sup> The nature of the movement that at least *yes* and *no* undergo would seem to be phrasal movement rather than head movement, given its distance and insensitivity to intervening filled heads. I therefore assume that the polarity particles involved in retorts are all generated in Spec- $\Sigma$ P (see Zanuttini 1997, Haddican 2004, Kramer and Rawlins 2010, a.o. for relevant discussion).



Unlike the clause-internal polarity particles *too/so* and *not<sub>R</sub>*, the left-edge particles *yes* and *no* cannot remain in-situ. They have the status of operators,

<sup>31</sup>As an alternative to the movement-based account of *yes/no* that I adopt here, one could argue (as Kramer and Rawlins 2011 and Holmberg 2013 have) that these polarity particles are base-generated at the left edge of the clause, at or near their Spell-Out position, and they are associated with clause-internal  $\Sigma$  via Agree. The data in (38) would seem to work against such an approach, though it could perhaps be explained on a base-generation account as well.

Thoms (2012b) takes a mixed approach: he adopts the movement analysis argued for here (originally presented in Sailor 2011) for *yes*, but claims that *no* is base-generated in the left periphery. This difference is intended to capture Thoms' intuition, exemplified in his ex. (11), that while *yes* (esp. *yeah*: see §3.2.1.2) can reverse the polarity of an antecedent question/assertion (cf. Farkas and Roelofsen to appear: §3.6), apparently *no* cannot (see Thoms' ex. (27)). However, I suspect this asymmetry is illusory, owing to an inherent ambiguity with respect to negation in the antecedent itself (see also Holmberg 2013). If, for example, negation is entirely given in a discourse (and thus not likely to be construed as a marker of a biased question, nor available for focus-closure in an elliptical response such as those we are discussing), then Thoms' asymmetry disappears: *no* is able to reverse the polarity of its antecedent, e.g. *A: I know Mary didn't reply, and Bill didn't reply...what about John – did John not reply? B: No, (he did,) it just got lost in the shuffle.* Thus, the guiding intuition for distinguishing *yes* and *no* on syntactic grounds disappears, while the benefits of a movement approach argued for here and in Thoms (2012b) persist.

meaning they necessarily undergo movement to an initial position in the clause (Postma and van der Wurff 2007:228-9):

(40) \*He has yes!

I assume that this movement deposits *yes* and *no* into a high polarity projection in the CP domain – a projection which several authors have argued for independently in recent years: see Laka (1990), Zanuttini (1997), Holmberg (2001), Postma and van der Wurff (2007), Farkas and Bruce (2010), Kramer and Rawlins (2011), Breitbarth et al. (2013), among others. To distinguish this high polarity projection from the clause-internal  $\Sigma$ P position we have discussed so far, I will refer to the higher phrase as Pol(arity)P, which is also involved in *yes/no* answers to polar questions (which I leave aside here). On the motivation for this movement to the high PolP, as well as the difference between PolP and  $\Sigma$ P, I follow van Craenenbroeck (2010:165), who, in his analysis of retort-like utterances in dialects of Dutch, notes (following others): “[g]iven that contradictory sentential emphasis clearly has scope over the entire proposition, it seems natural to assume that it is focus marking on the high PolP that is used to express such emphasis”.<sup>32</sup> I leave the matter of focus-marking PolP aside here (see van Craenenbroeck 2010:168-9); relevant for us is only that the contrastive polarity particle is required to be in this high position.

A rough sketch for the derivation of LEP retorts is below (to be refined):<sup>33</sup>

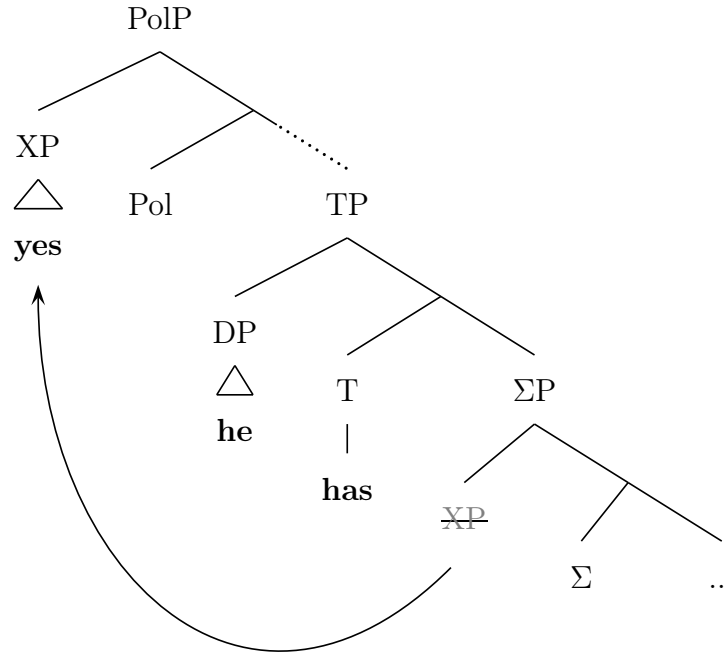
---

<sup>32</sup>See Rezac (2013) for the suggestion that, by selecting a propositional argument, Pol<sup>0</sup> can force the rest of the clause to be discourse-given (i.e., neophobic: §3.2.2.1). See §3.3, below, for further discussion.

<sup>33</sup>The exact location within the C-layer of what I call PolP remains to be determined. I am therefore not committed to Pol<sup>0</sup> selecting TP; I reflect this with a dotted branch indicating the possibility of additional (obscured) structure. I also omit movement of the subject and auxiliary for simplicity.

*The Left-Edge Polarity retort strategy*

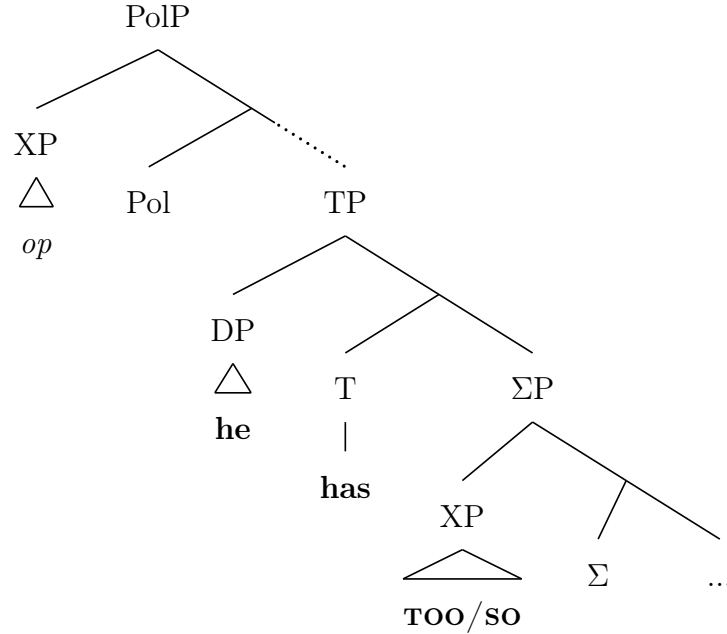
(41) Yes he has!



Turning now to CIP retorts, I suggest that they also involve a polarity operator (*op*), which is covert, and base-generated in (and not moved to) Spec-PolP:<sup>34</sup>

<sup>34</sup>Another possibility, which I do not explore here, is that *too/so* and *not<sub>R</sub>* are in fact generated in  $\Sigma^0$ , and the silent polarity operator described here is generated in Spec- $\Sigma$ P and moves to PolP, equivalent to *yes/no*. While unification of these polar operators is appealing, such an approach is at odds with an aspect of the analysis to be explored below: namely, that emphatic  $\Sigma^0$  is always realized as a segmentless stress morpheme (and therefore not as e.g. *too/so*).

(42) He has TOO/SO!



Whether covert or overt, the polarity operators in LEP and CIP retorts are incompatible with embedded environments, perhaps by way of introducing intervention effects with embedded operators (Haegeman 2012).

Interestingly, this difference in operator (non-)movement between LEP and CIP retorts may be responsible for a difference in their prosodic properties. As mentioned above in §3.2.2.4, the intonation of retorts uniformly involves a pitch accent on an expression of polarity. CIP retorts only have one polar expression, so pitch accent placement there is straightforward; LEP retorts, however, involve two expressions of polarity (the left-edge response particles *yes/no* and the regular clause-internal polarity morphology associated with  $\Sigma$ ), meaning there are two potential targets for the pitch accent associated with retort intonation: see the data in and discussion around (30) and (31), above. Assuming that prosody is fed by syntax – as is now common, following Bresnan (1971) and much subsequent

work – I would like to suggest here that the distinction in pitch accent placement between the retort types directly reflects a distinction in their derivation.

Building on Laka’s (1990) “[<sub>aff</sub>]”, I claim that emphatic polarity – both negative and affirmative, and both in retorts and elsewhere – is characterized by a segmentless stress morpheme in  $\Sigma$ ,<sup>35</sup> which I represent as [ ‘ ] (see also Schütze 2004:504). This morpheme needs a phonological host, and it docks to the nearest phonological word to its left (often in its own specifier, but not always, as we will see).<sup>36</sup> Regardless of the element to which it docks, this morpheme is prosodically realized as the L\*+H pitch accent described in §3.2.2.4. As the morpheme is associated with focus, it naturally corresponds to a pitch accent in the prosody; it therefore receives the pitch accent associated with the retort contour. The remaining component of the retort contour (the L- phrasal tone and the H% boundary tone) are mapped to the remaining post-focal material in the usual way (Silverman et al. 1992).

Independent support for this state of affairs – namely, that focusing a silent head causes the corresponding focal stress accent to be realized to the left of that head (e.g. on its specifier) – may come from the behavior of English reflexive anaphors in particular focus environments. Ahn (in progress:ch. 4) shows that reflexives in certain contexts necessarily receive focal stress even though, strictly speaking, they are not the semantic focus in the clause. This state of affairs arises

---

<sup>35</sup>I suspect that this may be a general phenomenon in English (and other languages: see below): namely, that focusing any silent head can, in principle, lead to a surface prosody in which the associated focus accent is realized to the left of that focused head. Thus, my characterization of this phenomenon as a floating stress morpheme [ ‘ ] here is perhaps too restrictive: it may very well be that this “morpheme” is not part of  $\Sigma$ ’s lexical entry, but is instead simply the canonical prosodic realization of focus in English (i.e., a particular pitch accent) construed in a non-canonical configuration (i.e., one in which it cannot surface on the intended morphosyntactic target). See below for a similar state of affairs that does not involve focused  $\Sigma$ .

<sup>36</sup>*Do*-support is generally thought to be the Last Resort insertion of a verbal host for “stray” tense morphology in  $T^0$ . While it might be intuitive to assume that *do*-support could also be triggered by the need to host a “stray” [ ‘ ], this does not seem to align with the facts: as I discuss below, the L\*+H pitch accent that characterizes retorts does not always surface on  $T^0$  (or the T- $\Sigma$  complex). Therefore, I assume that all instances of *do*-support in retorts arise for the usual reason, i.e. to support tense (made “stray” either by VPE or by the blocking of T-lowering, under standard assumptions). See Schütze (2004) for an alternative.



when reflexivity is the semantic focus: as Ahn has argued elsewhere, the locus of reflexive semantics is not the anaphor itself, but rather the silent head whose specifier the anaphor occupies (Ahn 2013); therefore, any environment in which reflexivity is focused necessarily involves focusing a silent head. As shown below, such environments surface with focal stress on the anaphor (e.g. *himself*), despite that its referent is already discourse-given in the question, and thus cannot be the semantic focus in the answer in this context (Ahn in progress:ch. 4):

- (43) A: Who hit Charles?  
 B1: Charles hit HIMSELF.  
 B2: #CHARLES hit himself.

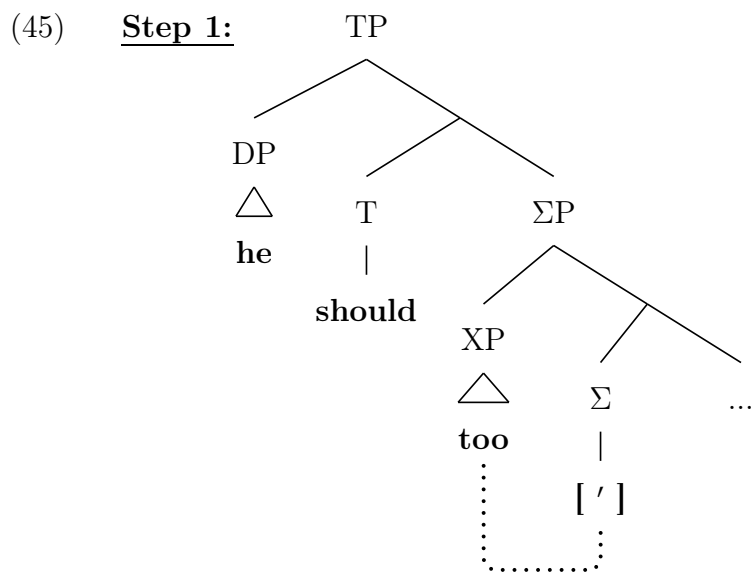
If Ahn is correct that the semantics of reflexivity lies with a silent functional head hosting the anaphor (rather than with the anaphor itself), then such examples pattern like the emphatic polarity cases discussed here: there is a subtle but noteworthy mismatch between information structure and surface prosody. Specifically, a focused head cannot support the accompanying focal stress because it is silent, so the focus accent is instead realized on the nearest phonological host to the left of that focused head (typically, perhaps always, on that head's specifier). For an additional possible example of this phenomenon, see Laka (1990:105, 140), where it is suggested that non-retort emphatic affirmation in Basque can surface as focal stress on the non-focused subject (and see Thoms 2012a: fn. 13 for a similar-looking phenomenon in Norwegian).

With this logic in place, I propose that the prosodic difference between (29) and (30), repeated in part below, is a side-effect of the derivational difference described above:

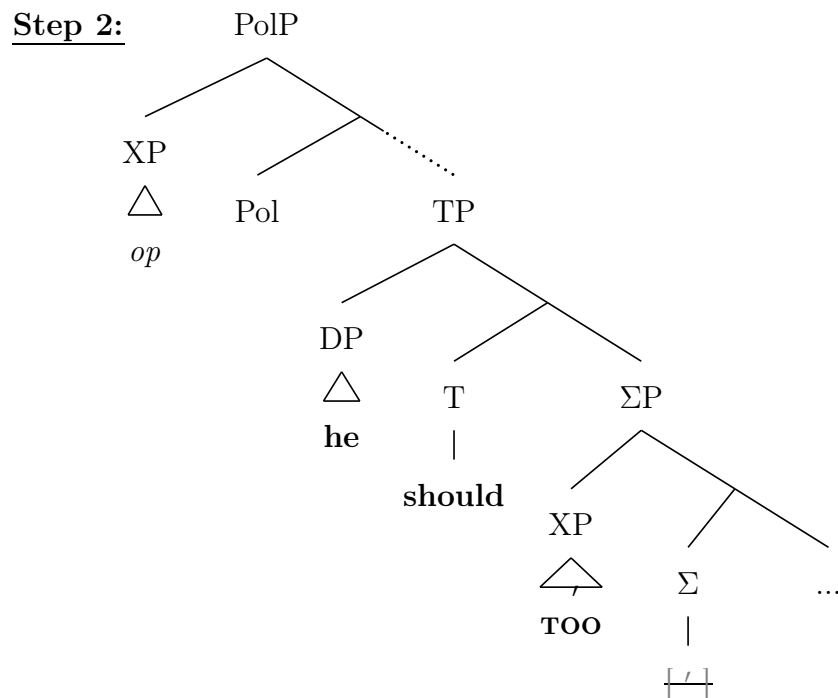
- (44) a. He should TOO have been! *CIP*  
 b. YES he should have been! *LEP (initial polarity stress)*

Recall that the analysis I advocate here for LEP retorts involves overt movement of the polarity operators *yes* and *no* from Spec- $\Sigma$ P to a left-peripheral PolP, supported by the observation that such polarity operators are in complementary distribution with lower CIP markers like *too/so* and *not<sub>R</sub>*. I claim that focused initial polarity particles in LEP retorts pick up their pitch accent the same way that CIP markers do: by being a suitable phonological host for [ ' ] by way of being in Spec- $\Sigma$ P at a relevant stage of the derivation (see below). Specifically, [ ' ] is able to dock to *yes/no* before they undergo movement to PolP. I depict the docking of [ ' ] in derivations for the two retort types below with a dotted line:<sup>37</sup>

*Derivational stress placement in a CIP retort*

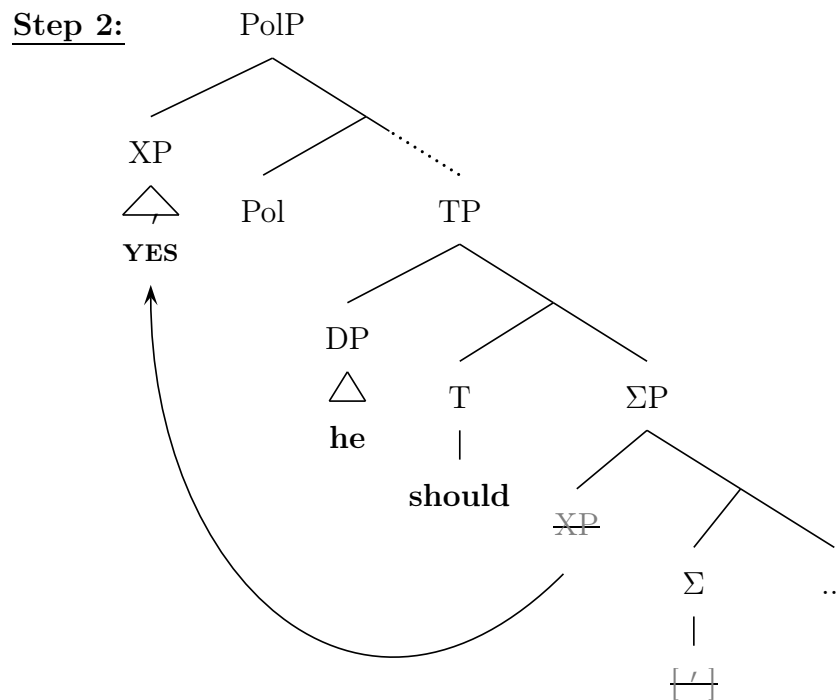
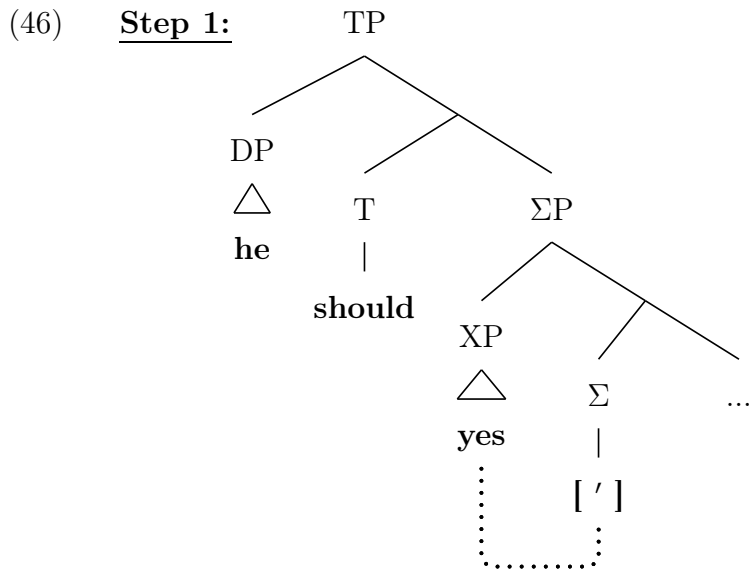


<sup>37</sup>Alternative approaches to the emphatic polarity-prosody connection described here involve either (i) a dedicated emphatic polarity projection (distinct from non-emphatic polarity projection(s)), as in Lipták (2003); or, (ii), a (morpho)syntactic interaction between polarity and a generalized focus projection (or feature), as in Holmberg (2001). See van Craenenbroeck (2010:§12.3) for discussion and deployment of each. It is unclear how (or whether) these alternatives can be distinguished from one another or from the option I suggest here on purely empirical grounds; thus, I leave the matter open.



In the CIP retort above, stress placement is straightforward: the polarity particle remains in situ, in a spec-head configuration with the focused  $\Sigma$ . Compare this with an LEP retort in which the initial *yes* or *no* bears the pitch accent in the retort contour:

*Derivational stress placement in an LEP retort (with stressed yes/no)*



In this derivation, [ ' ] docks to the polarity particle in Spec-ΣP prior to that XP's movement to the left edge. As a result, focal stress winds up getting pronounced on a morpheme that is quite far away, both linearly and structurally, from that stress's derivational origin: the information-structural focus,  $\Sigma^0$ .

This accounts for the prosodic difference between CIP retorts on the one hand and stress-initial LEP retorts on the other hand, as exemplified in (44). What remains, then, is to account for the other available stress pattern in LEP retorts, as exemplified earlier in (31): namely, the one in which focal stress is realized clause-internally, on  $\Sigma^0$ . Such examples look similar to the CIP stress pattern, repeated in (a) below:

- (47) a. He should TOO have been! *CIP*  
 b. Yes he SHOULD have been! *LEP (internal PE stress)*

There is no obvious semantic difference distinguishing the choice of pitch accent placement in LEP retorts: that is, (44b) and (47b) are semantically equivalent.<sup>38</sup> This suggests that the option of whether to dock the pitch accent to the polarity operators *yes* and *no* versus the internal expression of polarity is strictly governed by the syntax-phonology interface, i.e. the cycle, in the spirit of Bresnan (1971), Legate (2003), and many others.

With this in mind, I assume that the docking of [ ' ] in an LEP retort can optionally take place at a stage in the derivation that follows movement of *yes/no* to Spec-PolP. Looking leftward, [ ' ] will fail to find a suitable host in Spec- $\Sigma$ P (containing only a trace/deleted copy); instead, it will reliably find overt material in  $T^0$  onto which it can dock (see fn. 36).

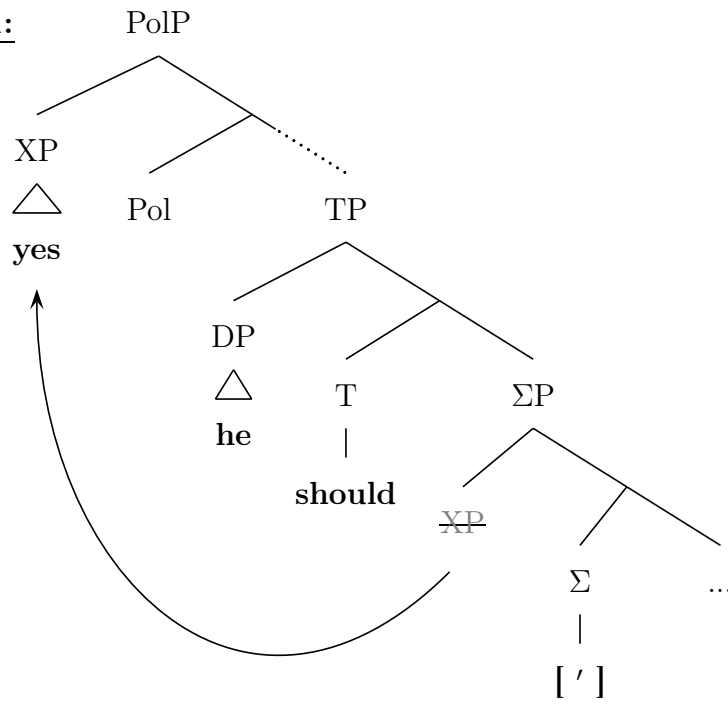
*Derivational accent placement in an LEP retort (with internal polarity stress)*

---

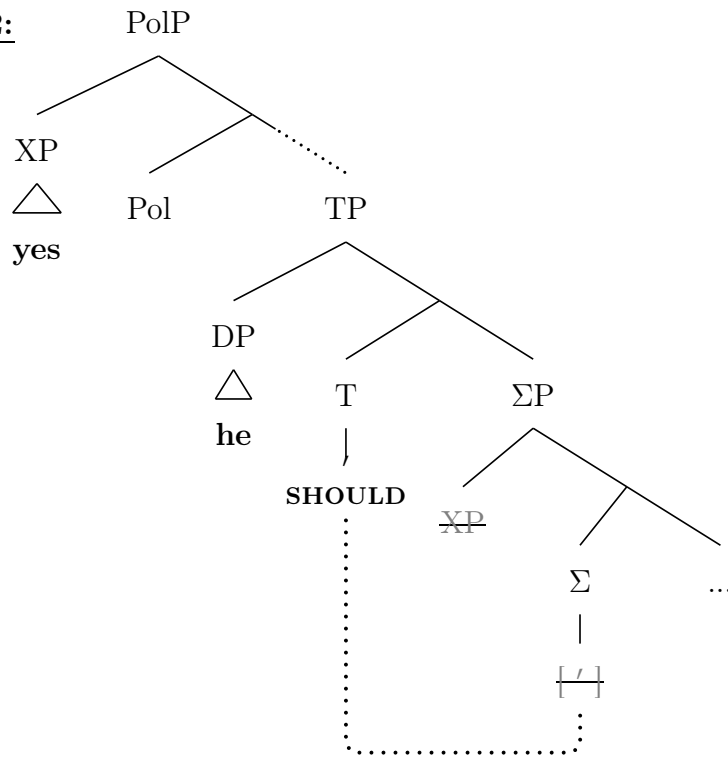
<sup>38</sup>I have very little to say about what governs this choice in LEP stress placement. My own intuition is that one speaker's choice can influence another's in a back-and-forth exchange of LEP retorts. That is, in an argument between speakers A and B, if A uses a LEP retort (e.g. with initial stress: NO *she isn't!*), then an LEP response by B is more likely to bear the opposite stress pattern (in this case, internal stress: *Yes she IS!*). In other words, an exchange of LEP retorts is likely to involve alternation between the two patterns, rather than repetition of one of the patterns. These intuitions await (dis)confirmation from naturally-occurring examples.

(48)

Step 1:



Step 2:



The choice between docking [ ' ] early (i.e. to *yes/no*) versus docking late (i.e. to the  $\Sigma$ -T complex) described here is reminiscent of independent interactions of cyclic A' movement and various phonological processes: see e.g. Legate (2003:512).

### 3.2.4 On some differences between CIP and LEP retorts

The similarities between CIP and LEP retorts described earlier recommended a unified analysis, which I provided in the previous subsection. However, there are properties that distinguish these retort strategies; I mention some here for descriptive completeness, but their source and analytical relevance remains open.

First, LEP retorts often appear with an initial particle, *oh*, within the same Intonation Phrase as the rest of the retort. However, this particle is entirely unavailable in the CIP strategy:<sup>39</sup>

- (49) a. **Oh** no he didn't!  
 b. **Oh** yes he should have been!
- (50) a. \***Oh** he did NOT!  
 b. \***Oh** he should TOO/SO have been!

The semantic contribution of *oh* is, like many discourse particles, difficult to characterize precisely; see Aijmer (2002:ch. 3) for extensive discussion.<sup>40</sup> Regardless,

---

<sup>39</sup>The appearance of *oh* in a LEP retort seems to prefer the “delayed” intonational contour described in (31), where the L\*+H pitch accent is realized on the clause-internal expression of polarity (rather than the clause-initial one). However, placement of the pitch accent on the initial polarity particle, as in (30), seems marginally possible as well.

In certain varieties of American English (e.g. AAVE), the strings in (50) are felicitous in contexts that do not license retorts (as defined here). For example, in such varieties, (50a) could be uttered by a speaker who accepts the antecedent assertion (i.e., does not move to block its admission to the common ground), but who wishes to express surprised disapproval (say) at said assertion. Given these different pragmatics, such utterances do not meet the criteria of “retort” offered here; but, their properties deserve to be worked out in more detail.

<sup>40</sup>In particular, see Aijmer (2002:§3.6.5) for instances of *oh* in reaction and objection contexts, including examples involving the particle sequence *oh but*. Note that denials such as *oh but he {IS/ISN'T}!* are attested; but, since they cannot co-occur with the polarity particles discussed here (among other reasons), they fall outside the domain of retorts (cf. ex. (4b)).

this asymmetry in the availability of *oh* in retorts is puzzling, given that the two retort strategies intuitively seem to share the same discourse function. Put differently, if *oh* is licensed in the same discourse contexts as LEP retorts, then we expect it to be licensed in the same contexts as CIP retorts as well, contrary to fact. The resulting inference holds that either (a) the discourse properties of the two retort strategies are in fact not identical, despite all other appearances; or, (b) the discourse particle *oh* is compatible with the discourse properties of CIP retorts, but it is rendered unavailable by some extra (perhaps syntactic) requirement that such retorts do not fulfill. I leave this matter open.

Second, there appears to be an asymmetry between the two retort types with respect to VPE. As noted earlier, VPE is strongly preferred (perhaps necessary) in LEP retorts; however, it is optional in CIP retorts:

- (51) A: John left yesterday.  
       B1: NO he didn't (??leave yesterday)! *LEP*  
       B2: He did NOT (leave yesterday)! *CIP*

Although CIP retorts share with LEP retorts a strong preference for pronominalized subjects (as noted earlier), the two differ with respect to VPE, suggesting the correlation between the two proposed earlier may be incomplete or only apparent.

Finally, consider two different types of morphological contraction – what I will call “T-contraction” and “ $\Sigma$ -contraction” – exemplified (in the negative) below:

- (52) He is not coming.  
       a. He’s not coming. *T-contraction*  
       b. He isn’t coming.  *$\Sigma$ -contraction*

As the terms indicate, T-contraction refers to a configuration in which material



occupying  $T^0$  (at some stage of the derivation) contracts to become part of the preceding phonological word,<sup>41</sup> while  $\Sigma$ -contraction involves a similar process for material occupying  $\Sigma^0$  (in this case, *not*  $\rightarrow$  *-n't*).

The two retort types can be distinguished in part by their behavior with respect to these types of contraction. Whereas LEP retorts freely allow either type of contraction, CIP retorts do not. Indeed, neither variety of contraction is possible in CIP retorts. Leaving aside affirmative cases for now (see below), consider the following:

*Negative LEP and CIP retorts with contraction*

- |      |                 |  |
|------|-----------------|--|
| (53) | a. NO he's not! | <i>T-contraction</i>                   |
|      | b. NO he isn't! | <i><math>\Sigma</math>-contraction</i> |
| (54) | a. #He's NOT!   | <i>T-contraction</i>                   |
|      | b. #He ISN'T!   | <i><math>\Sigma</math>-contraction</i> |

Though (54) contains two attested strings in English, neither is a retort as we have defined them: for example, they are infelicitous with retort intonation, they can be embedded, they are not neophobic, etc. This is a clear behavioral asymmetry distinguishing the two retort types, at least in the negative.

The pattern in (54) is unique to (CIP) retorts; it does not, for example, arise due to some general property of focused  $\Sigma$ . Specifically, in non-retort contexts with focused negation, T-contraction and  $\Sigma$ -contraction are both possible:<sup>42</sup>

---

<sup>41</sup>For reasons I do not understand, T-contraction in American English is more constrained in negative clauses than affirmative ones: in the negative, only forms of *be* can contract naturally. This is not true of British varieties, which seem to allow the full range of T-contraction in both affirmative and negative (e.g. *They've not said a word*, *They'd not do such a thing*, etc.).

<sup>42</sup>Note that we cannot test affirmative  $\Sigma$ -contraction, since focused affirmative polarity in non-retort contexts does not have segmental content. Instead, a non-segmental floating stress morpheme (which Laka 1990:§2.2 calls “[<sub>aff</sub>]”) is the only available strategy for expressing emphatic affirmative polarity in non-retort contexts. Thus, since [<sub>aff</sub>] has no segmental content, it clearly cannot undergo contraction in the usual sense.

We can test the availability of T-contraction in a clause containing [<sub>aff</sub>]. It is impossible:

(i) A: John is not coming.

*Contraction in non-retorts involving focused negation*

- (55) A: Is John coming?  
B1: I told you earlier that [he is NOT (coming)].  
B2: I told you earlier that [he's NOT (coming)]. *T-contraction*  
B3: I told you earlier that [he ISN'T (coming)]. *Σ-contraction*

Thus, the unavailability of contraction is not a general property of focused polarity; it is a property particular to CIP retorts. The same holds for affirmative CIP retorts, to the extent that relevant examples can be constructed:<sup>43</sup>

- (56) A: John isn't leaving.  
B1: He is TOO/SO!  
B2: \*He's TOO/SO! *T-contraction*  
B3: \*He ISTA! *Σ-contraction*

In addition to establishing a behavioral distinction between CIP and LEP retorts, the preceding discussion also provides evidence that the *not* in CIP retorts – which I have been calling ‘*not<sub>R</sub>*’ – is not simply focused sentential negation. It exhibits special properties unique to retort contexts, consistent with the hypothesis that this *not<sub>R</sub>* is in fact the negative analogue of the affirmative *too/so* (which are by definition retort-specific polarity particles). With respect to LEP retorts, the preceding suggests that negative instances of LEP retorts behave the same as affirmative instances: neither involves a special clause-internal polarity particle

- 
- B1: I told you earlier that he IS (coming).  
B2: \*I told you earlier that HE'S (coming).

This pattern stands in contrast to (55.B3), which shows that focused NOT in non-retort contexts is capable of undergoing contraction. The relevant generalization seems to be that material in  $T^0$  can ‘inherit’ the focus of a contracted element cliticized to it, whereas material in Spec-TP (i.e. subjects) cannot. I leave this aside; see Laka (1990:§2.5) and Zwicky and Pullum (1983: fn. 6) for some discussion.

<sup>43</sup>Since TOO/SO is only licensed in retorts, and contraction of clause-internal polarity particles in retorts is impossible, there is no licit contraction for TOO/SO in the grammar of English. Therefore, the ungrammatical case in (55.B3) is simply hypothetical.

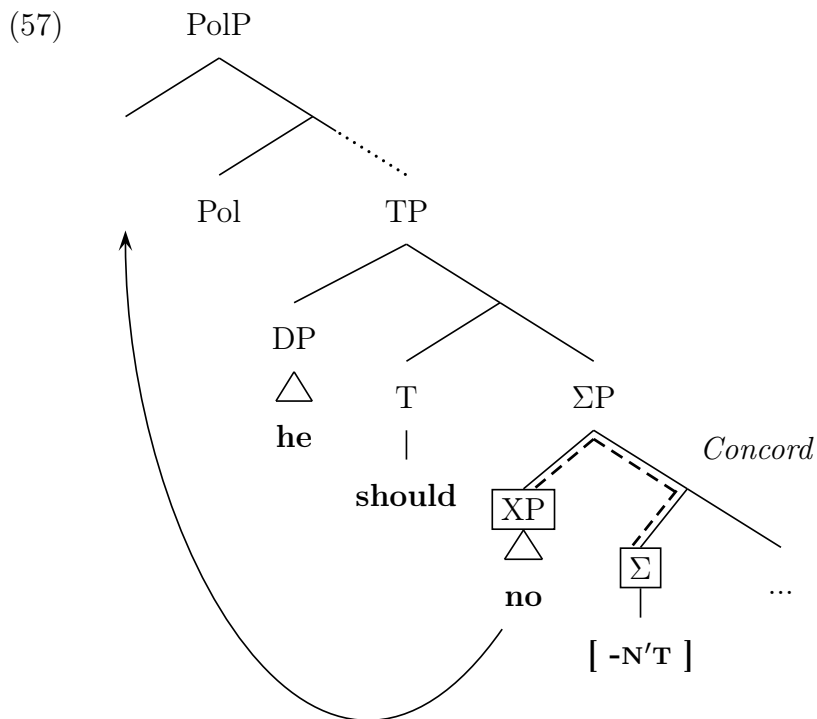
(e.g. *not<sub>R</sub>*), given the availability of contraction in retorts of that type.

An obvious question remains: what is the status of the clause-internal negative marker in negative LEP retorts? After all, it has the rather alarming property (for Standard American English) of sharing a clause with another negative marker, *no*, despite that similar negative co-occurrences within a clause lead to a canceling/polarity-flipping effect (Horn 1989:§§4.2, 5.1.3).<sup>44</sup> A negative LEP retort has but a single (emphatic) negative reading. Why does the double-negation effect not arise in such cases? Here, I follow Postma and van der Wurff (2007:228), van Craenenbroeck (2010:164), and Kramer and Rawlins (2011): the clause-internal negation that co-occurs with initial *no* is, in effect, an instance of negative concord. That is, the clause-internal negative marker in LEP retorts arises on  $\Sigma^0$  following spec-head agreement with the negative polarity operator *no* before the latter moves to PolP:<sup>45</sup>

---

<sup>44</sup>Note that negation in other non-retort RAs behaves similarly.

<sup>45</sup>In negative LEP retorts with the “delayed” pitch accent pattern, the L\*+H pitch accent always goes with the negative morpheme. Specifically, if negation is contracted, the pitch accent consistently surfaces on the T+ $\Sigma$  complex (e.g. SHOULDN’T); if negation is uncontracted (to the extent that such examples are acceptable), the pitch accent surfaces there (e.g. NOT). I take this to indicate that the floating stress morpheme [ ’ ] only looks leftward for a phonological host. I take no stance here on the derivation of  $\Sigma$ -contraction (i.e., on how *-n’t* gets pronounced on T<sup>0</sup>).



In sum, I have laid out a few noteworthy differences between CIP and LEP retorts in this subsection (some of which helped to establish or motivate important features of their analysis). While none of these differences directly challenge the present analysis, they deserve further treatment, which will not be provided here.

We now turn to some novel data involving English retorts.

### 3.2.5 On subjectless retorts

In arguments involving a back-and-forth exchange of CIP retorts (cf. fn. 38), an alternate form of CIP retort can arise – one first observed in Sailor (2011, 2012c), but which has otherwise gone unnoticed in the literature – whereby the subject is omitted:<sup>46</sup>

<sup>46</sup>See Holmberg (2001) for a similar-looking phenomenon in Finnish.

*Subjectless CIP retorts*

- (58) A: He has NOT!  
B: Ø has TOO/SO!  
A: Ø has NOT!  
...etc.

Retorts taking this form minimally comprise a filled  $T^0$ , followed by the polarity particle typical of CIP retorts; they maximally contain this material and the usual array of ‘strandable’ auxiliaries in verbal ellipsis contexts, and no more.

*Subjectless CIP retorts (multiple stranded auxiliaries)*

- (59) A: He should NOT have been!  
B: Ø should TOO/SO have been!  
A: Ø should NOT have been!  
...etc.

There are no restrictions on the phi-features of the omissible subject in (58). This is surprising, given that English uniformly prohibits *pro*-drop, even in non-retort reversing assertions (but see below):

*English subject-drop prohibited in similar contexts*

- (60) A: He<sub>i</sub> is leaving already.  
B: You’re wrong, \*(he<sub>i</sub>) is NOT.  
(61) A: He<sub>i</sub> isn’t leaving.  
B: [\*(He<sub>i</sub>) IS, actually.]

Crucially, though, omission of the subject in subjectless retorts such as (58) requires omission of the VP:

*Subject omission requires VP omission*

- (62) A: Is TOO/SO (\*leaving)!  
 B: Is NOT (\*leaving)!
- (63) A: Should TOO/SO have been (\*talking)!  
 B: Should NOT have been (\*talking)!

In other words, these missing-subject retorts would seem to require VPE. This is striking – as we saw above in (51), under normal circumstances (i.e. when the subject is present), VPE is optional in CIP retorts:

*CIP retorts: VPE not required*

- (64) He is TOO/SO (leaving)!

The fact that subject omission is tied to VP omission tells us that this phenomenon is manifestly a syntactic one: that is, the subject omission in (58) cannot be explained as, for example, the result of rapid speech or simple recoverability. As Haegeman (2013:90) says in her discussion of other exceptional cases of subject omission in English (e.g. so-called *diary drop*), “...though recoverability no doubt plays a role, register-specific omission phenomena cannot purely be analyzed in functional terms [...] they are subject to syntactic constraints”.

The missing-subject retorts of (58) are clearly different from other “register-specific omission phenomena” that Haegeman and others (e.g. Kay 2002) have previously described, though, since the latter are not contingent upon VP-omission. Indeed, for this reason, Haegeman’s influential truncation analysis of these other subject-omission phenomena would seem to be the wrong approach for the subjectless retorts I describe here. I believe a very different analysis is called for: one that provides a natural explanation for the data in (62).

I argue that, despite appearances, subjectless retorts are not derived by any sort of subject omission operation *per se*, or by truncation of left peripheral structure. Instead, I claim that the subject and the VP are eliminated together, in

one fell swoop, by an independently-attested phenomenon: TP ellipsis (TPE).<sup>47</sup> Given that subjects normally surface in Spec-TP, a position higher than those occupied by material that is pronounced in subjectless retorts (i.e. the auxiliaries and polarity particle), we are forced to conclude that this pronounced material has escaped TPE by undergoing movement to a position above the ellipsis site.

This conclusion presents a handful of problems. Foremost, under standard assumptions, the pronounced elements in subjectless retorts such as (58) do not form a constituent. That is, the tensed auxiliary and the PE marker (and any lower auxiliaries) are not thought to form a constituent that excludes the surface subject and the VP. Thus, if the pronounced elements in such retorts have undergone movement allowing them to escape a large ellipsis site, then they must not have moved as a single constituent: they must have moved independently of one another, and yet wound up in the same surface order that they would have been pronounced in had they not undergone this movement. This is a non-trivial challenge for the present analysis. Still, a few analytical options present themselves.

The first option is to simply reject standard assumptions: one could argue that the surviving elements in subjectless retorts do form a constituent, thereby allowing them to escape ellipsis with a single step of movement. This could perhaps be conceived of as VP movement followed by remnant movement of the minimal constituent containing both the tensed auxiliary and the polarity particle (and any lower auxiliaries), but not containing the subject (or the landing site of the moved VP). This sort of analysis is compatible with Koopman and Szabolcsi's (2000) general derivation for verbal complexes (see in particular *op. cit.*: ch. 4 on

---

<sup>47</sup>Assuming, as is typical, that external arguments are first-merged below the auxiliary complex, one might wish to entertain the possibility that a smaller constituent than TP – say, just VP – undergoes ellipsis in such cases, but that it does so prior to subject raising, thus deleting the subject along with the predicate. I reject this approach, though, for the simple reason that it is not independently attested in English: none of the other manifold VPE configurations involve ellipsis of the subject in this way. (See McCloskey 1991 on Irish Gaelic, a language whose subjects independently remain in-situ within the extended verbal projection, and thus can be elided by VPE.)

“VP-splitting” and PredP), but it would require the remnant-moved constituent to be large enough to contain the tensed auxiliary, but not so large that it contains the subject; this of course would mean that “TP” is not the elided constituent, but some other functional projection in the TP-layer. I will not explore this possibility further.

If we maintain standard assumptions, though, then the surviving elements must have moved independently, as described above, and we are left to explain how their order is preserved. Here, one could invoke Cyclic Linearization (Fox and Pesetsky 2005): the moved elements are pronounced in their underlying order because that order had been fixed at an earlier stage of the derivation, and failing to preserve that order would yield a structure that could not be linearized. This approach has appeal, but its implementation would face at least one significant problem: which Spell-Out domain is responsible for fixing the order of these elements? It would have to be a domain that contained all of the relevant elements (in the surface positions they would otherwise occupy, e.g. T < polarity particle < auxiliaries), but which was properly below the lowest landing site of their eventual movement. This would seem to implicate TP as the point at which their order is fixed, but TP is almost universally argued not to be a Spell-Out domain (both in Cyclic Linearization and other frameworks). Perhaps their order is fixed upon merger of a low projection in the C-layer, followed by their (order-preserving) movement to higher positions within that layer. This approach enjoys the particular advantage of accounting for the necessity of ellipsis in subjectless retorts (cf. *\*Has TOO left!*): movement of the pronounced elements creates precedence violations with the subject, whose order had been previously fixed as preceding those elements; an independently-attested strategy for repairing such precedence violations involves deleting one or more of the elements in violation (Fox and Pesetsky 2005:13). Thus, a linearization-based analysis shows promise, but I do not attempt to work out its details here.



Relevant to the present discussion is Laka's (1990) *Tense C-Command Condition*, given below (*op. cit.*:9):

- (65) Tense must c-command at S-structure all propositional operators of the clause.

Laka (1990:41) suggests that this condition is related to the independent semantic requirement that event variables be bound by Tense. She argues that the TCC must be satisfied by a  $T^0$  with phonological content, and its empirical effects can be seen in English *do*-support (*op. cit.*:§1.3) and Basque auxiliary-fronting (*op. cit.*:§1.2.5, but cf. Haddican 2004). Assume that the null operator normally found in CIP retorts (see (42)) is absent in the subjectless variety, forcing the clause-internal PE marker to move to PolP in the C-layer (akin to movement of *yes/no* in LEP retorts: see (42)). If the clause-internal polarity particle counts as a propositional operator, which seems reasonable given its close ties to regular sentential polarity (e.g. negation, traditionally regarded as a propositional operator), then such a configuration would violate the TCC. If satisfaction of the TCC is capable of triggering movement, as Laka suggests it is in her discussion of Basque auxiliary-fronting, then this might explain the subsequent movement of  $T^0$  to a position where it c-commands the fronted polarity particle.<sup>48</sup> If Laka's suggestion is correct that the TCC has independent semantic motivations, then this explanation has promise.

It would raise a question, however: why do we not see the effects of the TCC (that is:  $T^0$ -fronting) in the derivation of LEP retorts? Recall that such retorts involve movement of the polarity particles *yes* and *no* to PolP. If subsequent  $T^0$ -fronting took place to satisfy the TCC, the result would be ungrammatical:

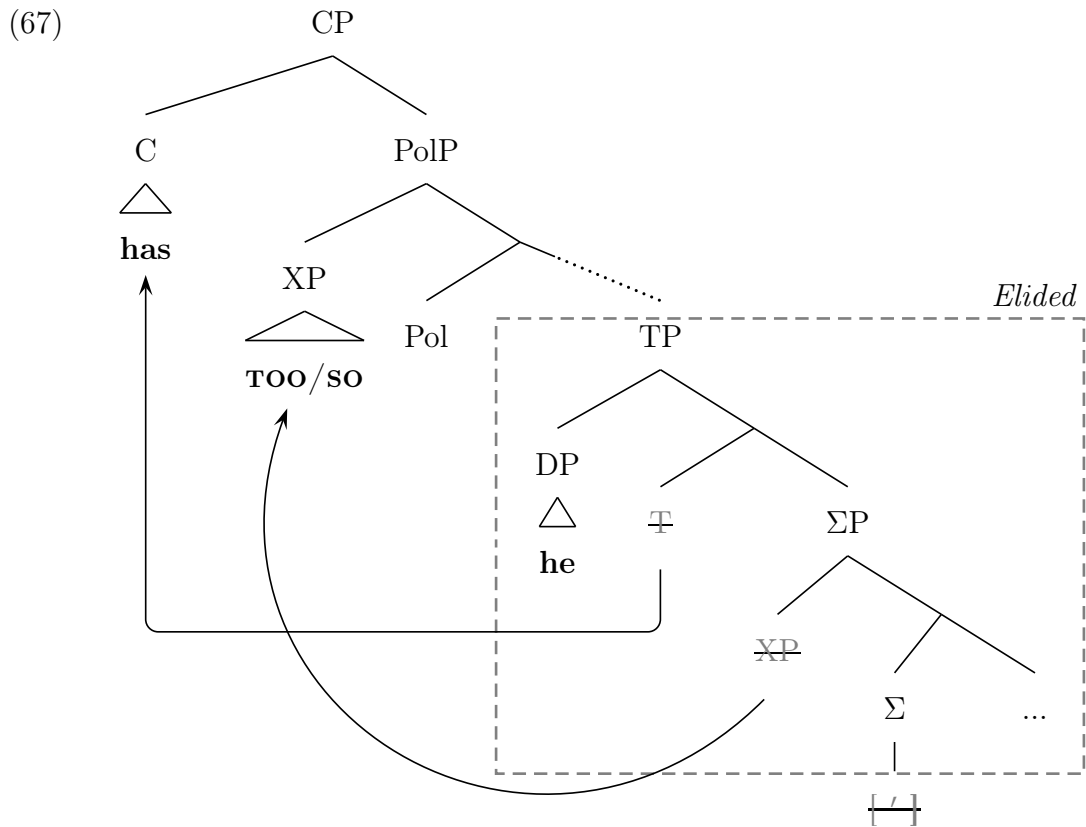
---

<sup>48</sup>This cannot explain why any auxiliaries below  $T^0$  apparently also move, but to positions below the fronted polarity particle. This is a non-trivial problem for a TCC-driven approach.

(66) \*Has YES he!

Thus, if the TCC is a factor in deriving subjectless CIP retorts, then we must explain why it is apparently not a factor in LEP retorts. The simplest explanation is that the operators *yes* and *no* (and *op*) are illocutionary operators (on par with e.g. question operators, etc.), and therefore do not trigger TCC effects. In other words, they contribute strictly to the illocutionary force of the retort, rather than to its propositional content, the latter perhaps mediated by the concord process described at the end of §3.2.4.

A derivation that assumes independent movement of the surviving elements in subjectless retorts is below. I depict  $T^0$  moving directly into  $C^0$  for simplicity, though nothing relies on this; in particular, it may be that *has* moves to some other left-peripheral projection:



### 3.2.6 Summary of retorts

The preceding discussion comprises a definition and description of English retorts, a particular type of reversing assertion that makes use of polarity particles to reject admission of a discourse-salient assertion to the common ground. We saw two different surface realizations of retorts (the CIP and LEP strategies), and I argued that the two share a common derivational (morphosyntactic, prosodic) core.

In the next section, I draw behavioral parallels between retorts and tag questions with respect to ellipsis and related anaphoric phenomena. This establishes an empirical foundation for the discussion of high-VPE in §3.5.

## 3.3 Retorts and tag questions as high-VPE configurations

Retorts share with dependent tag questions a cluster of properties relating to anaphoricity, including neophobia, preferential pronominalization and ellipsis, etc. I describe these similarities below. Later, we will see their potential relevance to the theory of ellipsis.

### 3.3.1 Neophobia

Both retorts and tags are neophobic: they must comprise entirely discourse-given material (modulo polarity and force). Inclusion of new material – e.g. adverbs, modals, auxiliaries, subjects, etc. – is impossible:

- (68) Franklin should have gone to clown college<sub>{,/.}</sub>
- |    |                                 |                     |
|----|---------------------------------|---------------------|
| a. | shouldn't he (*really) have?    | <i>Tag question</i> |
| b. | NO he (*really) shouldn't have! | <i>LEP retort</i>   |
| c. | He (*really) should NOT have!   | <i>CIP retort</i>   |

### 3.3.2 Pronominal subjects

About the subjects of tag question clauses, Rezac (2013:10) says, “[their] antecedent is unambiguous and [they] are obligatorily pronominalized”, which, he argues, makes tag subjects special with respect to a few phenomena. Specifically, Rezac cites Ross (1973), who shows that tag subjects can take certain idiomatic DPs as antecedents without incident, despite that such DPs normally resist pronominalization (see Ross 1973 for qualifications regarding speaker variation):

- (69) Some *headway<sub>i</sub>* has been made on problem X,  
a. hasn’t *it<sub>i</sub>*?  
b. ?\*but *it<sub>i</sub>* hasn’t (been made) on problem Y.

I will not attempt to explain this asymmetry; see Ross (1973). Rezac (2013:11) claims that pronominal reference to idiomatic DPs such as *headway* occurs “in tags alone”, but this is not quite right. Crucially, it occurs in retorts as well:

- (70) A: Some *headway<sub>i</sub>* has been made on problem X.  
B1: No *it<sub>i</sub>* hasn’t! *LEP*  
B2: *It<sub>i</sub>* has NOT! *CIP*

Thus, tags and retorts pattern the same with respect to this particular anaphoric process. From the perspective of Rezac’s quote above, what retort subjects have in common with tag subjects is that they are both obligatorily pronominal, and both are unambiguous.

This tag-retort connection persists with other types of antecedent DPs. Citing others, Rezac (2013) notes that pronouns anteceded by weak definites such as *the hospital* lack the non-unique interpretation that characterizes such weak

definites:<sup>49</sup>

(71) #Bill is in the hospital<sub>i</sub>, and John is in it<sub>i</sub> too. (unique rdg. only)

However, Rezac points out that this effect disappears when the pronoun occurs in a tag question. That is, pronominal tag subjects retain the weak definite non-unique reading:

(72) The hospital<sub>i</sub> is the best place to be if you're injured, isn't it<sub>i</sub>?

Once again, retorts pattern like tags in this respect:

(73) A: The hospital<sub>i</sub> is the best place to be if you're injured.  
B1: No it<sub>i</sub> isn't! *LEP*  
B2: It<sub>i</sub> is NOT! *CIP*

The similarities continue with other types of antecedents as well. For instance, certain non-referential subjects can serve as antecedents to pronouns in run-of-the-mill examples (cf. the idiomatic examples above), yet they stubbornly disallow coreference with a tag or retort pronoun:

(74) At least four men<sub>i</sub> are in the room, and they<sub>i</sub> look angry.  
(75) At least four men<sub>i</sub> are in the room<sub>{,/.}</sub>  
a. \*aren't they<sub>i</sub>?  
b. \*No they<sub>i</sub> aren't! *LEP*  
c. \*They<sub>i</sub> are NOT! *CIP*

---

<sup>49</sup>See Rezac (2013: ex. (15)) for a possible class of exceptions in a non-tag, non-retort context.

Moreover, speakers who reject epithet subjects in tags (but allow them elsewhere) also reject them in retorts, and vice-versa (and the contrapositive holds as well). In other words, the data in (76) are all judged alike within speakers, even though that judgment varies across speakers.

- (76) John<sub>i</sub> isn't here<sub>{,/.}</sub>
- a. %is the son of a bitch<sub>i</sub>?
  - b. %Yes the son of a bitch<sub>i</sub> is!
  - c. %The son of a bitch<sub>i</sub> is TOO/SO!

Finally, although the following is strictly speaking not solely a pronominal phenomenon, tags and retorts exhibit the same interpretational behavior when they take split (coordinated) antecedents. In a typical (i.e. non-tag, non-retort) coordinated-antecedent VPE configuration such as (77), two readings are available: a “distributive” reading, in which the subjects of the VPE clause are assigned as participants to the predicates from the coordinated antecedent in a one-to-one fashion (according to the order of the predicates in the antecedent, mimicking the effect of the adverb *respectively*); and, a “collective” reading, in which the subjects of the VPE clause are interpreted as participants in a conjunction of the antecedent predicates (Sailor 2009:34):<sup>50</sup>

- (77) John sang and Mary danced a jig. In fact, Bill and Sue did too.
- a. = sing and dance a jig (respectively) *Distributive rdg.*
  - b. = (both) sing and dance a jig *Collective rdg.*

A different pattern emerges for tags and retorts, however. As noted in Sailor (2009:§3.2.3.1), tags with split antecedents (“coordinated-antecedent tags”) lack

---

<sup>50</sup>I assume the collective reading is availed by the presence of coordinated subjects in the VPE clause (something that relevant tag and retort examples lack: see below), but further investigation is necessary.

the collective reading; only the distributive reading is available. This is also true of retorts:

- (78) John sang and Mary danced a jig<sub>{,/.}</sub>
- a. didn't they? (✓ distributive; #collective)
  - b. No they didn't! (✓ distributive; #collective)
  - c. They did NOT! (✓ distributive; #collective)

Thus, tags and retorts continue to pattern alike. Next, we examine their shared properties with respect to ellipsis.

### 3.3.3 Ellipsis

As mentioned earlier, a VPE configuration is a high-VPE configuration if at least VoiceSynP is included in the ellipsis site. This is diagnosable a number of different ways, although independent constraints on the shape of tag and retort clauses, e.g. neophobia, restrict the diagnostic possibilities.<sup>51</sup> Nevertheless, the available diagnostics all indicate that the VPE involved in tags and retorts is high-VPE.

Foremost, both straightforwardly resist voice mismatches:

*No voice mismatch in tag questions and retorts*

- (79) Your car should have been fixed by the mechanic last week<sub>{,/.}</sub>
- a. \*shouldn't he have?
  - b. \*No he shouldn't have!
  - c. \*He should NOT have!

This cannot be an effect of neophobia, since e.g. active/passive alternations do not

---

<sup>51</sup>For example, the status of strict interpretations for elided reflexives cannot be tested, since neophobia prevents the introduction of a new subject in retort and tag clauses.

*prima facie* require the introduction of (or, for that matter, removal of) any material, and active/passive counterparts are widely thought to be mutually-entailing (Merchant 2013b and references therein).

Secondly, VP-level adverbials are obligatorily recovered in tags and retorts. In other words, such modifiers cannot be ignored when determining the interpretation of the tag or retort clause (even though, for negative retorts, rejecting the content of the unmodified predicate would entail rejection of the modified one):<sup>52</sup>

*Obligatory recovery of verbal modifiers in tag questions and retorts*

- (80) Jordy carefully reviewed the book<sub>{,/.}</sub>
- a. didn't he?
    - (i) = didn't he carefully review it?
    - (ii) #didn't he review it? (Not necessarily carefully)
  - b. No he didn't!
    - (i) = he didn't carefully review it
    - (ii) #he didn't review it (carefully or otherwise)
  - c. He did NOT!
    - (i) = he didn't carefully review it
    - (ii) #he didn't review it (carefully or otherwise)

Finally, both retorts and tags allow omission of auxiliaries that cannot be omitted in low-VPE configurations, e.g. Prog<sup>0</sup><sub>be</sub>:<sup>53</sup>

*Auxiliary omission in tag questions and retorts*

- (81) John will be dying to get out of there<sub>{,/.}</sub>
- a. won't he?

---

<sup>52</sup>One might argue that this is not a property of ellipsis in such examples, but rather is due to other factors (e.g. the scope of the question in the tag example). I take no stance on this, and include such examples for completeness.

<sup>53</sup>See §3.4 for discussion of dialectal microvariation in this domain.



- b. No he won't!
- c. He will NOT!

Thus, retorts and tags seem to be the consummate high-VPE environments. If anything, these judgments are even clearer in retorts and tags than in run-of-the-mill coordinated root clause examples.

I would like to suggest that these behavioral similarities with respect to ellipsis are directly correlated with the other properties uniting retort and tag clauses described above (i.e. pronominalization and neophobia). Exactly how this should be theoretically implemented, however, is a challenge. Recall from earlier the supposition that while extra-syntactic (e.g. pragmatic/discourse) factors may be at play in the differences we see between low- versus high-VPE configurations, those factors do not operate directly on surface representations. As I showed, their effects are clearly diagnosable in the syntax, meaning their domain of direct influence is structural, not strictly phonological or linear (or, for that matter, semantic; cf. Kehler 2002).

For the sake of argument, assume that Kehler's (2002) discourse coherence model captures VPE's sensitivity to the coordination/subordination distinction seen earlier, such that when the E-clause stands in a *Resemblance* relation with its A-clause, the ACD problem is avoided and maximal parallelism is presupposed, yielding the high-VPE pattern; on the other hand, when the E-clause bears a *Cause-Effect* relation to the A-clause, the presupposition of maximal parallelism is absent, and low-VPE arises so as to avoid ACD. The relation between retorts/tags and their antecedents could be seen as consistent with the definition of a Resemblance relation.<sup>54</sup> What about the other properties uniting retorts and

---

<sup>54</sup>Kehler (2002:15) begins to define the Resemblance family of relations thus: "...the hearer identifies a relation  $p_1$  that applies over a set of entities  $a_1, \dots, a_n$  from the first sentence  $S_1$ , and a corresponding relation  $p_2$  that applies over a corresponding set of entities  $b_1, \dots, b_n$  from the second sentence  $S_2$ . Coherence results from inferring a common (or contrasting) relation  $p$  that subsumes  $p_1$  and  $p_2$ , along with a suitable set of common (or contrasting) properties  $q_i$

tags discussed here, e.g. neophobia, preferential pronominalization and ellipsis, etc.? These are presumably governed by extra-syntactic factors as well, e.g. discourse coherence. As Kehler (2002:45) himself notes, “...the repetition of a full lexical noun phrase or proper name in a context that licenses a pronoun can be misleading and lead to unwanted inferences. [...] We would expect the same to be true of VP-ellipsis in light of the anaphoric properties it shares with pronouns.”

The purpose of pointing out the above similarities is to suggest that they may be related to the obligatoriness of ellipsis in these phenomena. Like (LEP) retorts, tag question clauses that do not undergo VPE are heavily degraded:

*Obligatory VPE in tag questions*

(82) ??John gave a book to Mary, didn't he give one to her?

Elsewhere, VPE is optional in effectively every environment compatible with it (see below), so one might wonder whether the appearance of VPE in these phenomena is illusory, with some non-elliptical process responsible for the conspicuous absence of the predicate. Such an approach seems highly improbable for tag questions, at least: as Sailor (2009:§3.2) shows in detail, tag questions exhibit the full array of properties characteristic of uncontroversial (high-)VPE. The same can definitively be shown to hold for retort clauses as well.

So, why would VPE, an operation that is optional nearly everywhere else, be obligatory in these particular phenomena?<sup>55</sup> Obligatory ellipsis is hardly unknown in the literature: it is assumed or argued for explicitly in various instantiations of the “move-and-delete” derivation in which some element moves out of an XP prior to eventual ellipsis of that XP, in structures where non-ellipsis of XP would

---

of the arguments  $a_i$  and  $b_i$ .”

In the case of retorts and tags,  $S_1$  is the antecedent assertion/host clause,  $S_2$  is the retort/tag clause,  $p_1 = p_2$ , and  $a_i = b_i$ . The latter two equivalences characterize an existing subtype of Resemblance, namely the *Elaboration* relation (Kehler 2002:18-19), which may very well be the right way to think of retorts and tags with respect to discourse coherence.

<sup>55</sup>I thank Jim McCloskey for emphasizing the importance of this question.

render an ill-formed sentence (e.g. in sluicing, pseudogapping, etc.; for survey and discussion, see Merchant 2010, Sailor and Thoms to appear:§§4-5, and Sailor and Schütze 2013).

Obligatory ellipsis could be accomplished by analytical fiat: one could, for example, appeal to Merchant’s (2001) influential E-feature, and claim that some functional head shared by the relevant phenomena – say, a variety of the high Pol(arity) head, or another left-peripheral head associated with some type of force or speaker mood shared by the two utterance types – has the E-feature hard-coded into its lexical entry. This would effectively shift the burden onto selection, an operation known to be independently responsible for the introduction of obligatory derivational components. However, this approach would not constitute an explanation of the observed facts, but merely a formalization; it would not address the initial puzzle in any satisfying way.

Another approach might be to embrace the aforementioned similarity to move-and-delete phenomena, and seek an account of obligatory ellipsis from that literature. Fox and Pesetsky (2005), Thoms (2010), Sailor and Thoms (To appear), a.o. have suggested that the matter might be dictated by concerns of linearization associated with the “move” component of the move-and-delete derivation: if, for example, an XP contains an element that cannot be linearized (e.g. a problematic lower copy of a higher moved element), then the derivational recourse is to delete the offending element, which, in some cases, requires deletion of the containing XP. If we wish to appeal to such an approach to explain obligatory ellipsis in e.g. retorts and tag questions, we would need to identify the relevant offending element. I have already argued that movement is involved in the derivation of these phenomena earlier in this chapter, so those movement operations provide a reasonable place to start; however, since seemingly identical instances of such movement (e.g. T-to-C movement,<sup>56</sup> movement of *yes/no/op*, etc.) are attested

---

<sup>56</sup>But see Merchant (2003) for an apparent case of T-to-C requiring ellipsis.

in non-elliptical environments, this approach would seem to be a non-starter, at least without some degree of stipulation to once again uniquely distinguish the phenomena to be accounted for.

The seed of a third alternative can be found in recent work by Rezac (2013), who suggests that tag question clauses involve a particular left-peripheral polarity head (familiar to us as  $\text{Pol}^0$ , but which he calls  $\Sigma^0$ ) which encodes the “special” question interpretation characteristic of tags, and, crucially, which presupposes the givenness of everything else in the clause other than itself, including its own propositional argument, Rezac suggests. This approach is not without some degree of intuitiveness, though it would have to be loosened to accommodate retorts. Specifically, the question interpretation characterizing tags must originate from a head that is distinct from the one responsible for taking a propositional argument, ensuring givenness. This is desirable on independent grounds, since there is ample evidence independent of tags for a high polarity projection (see references cited above). Moreover, there is also a great deal of independent evidence for a peripheral projection related to interrogative mood and/or clause type more generally (Rizzi 2001), either of which is demonstrably distinct from any high instantiation of polarity. Divorcing the question interpretation of tags from  $\text{Pol}^0$ , which hosts contrastive polarity but ensures that all other content is given, allows us to accommodate retorts in a straightforward way: they involve  $\text{Pol}^0$ , as tag questions do, but the two utterance types differ in their force due to variation on the other functional head(s) associated with such meaning. Moreover, it seems right that the semantics of retorts should involve a component taking a proposition as an argument, giving you the correct interpretation that the scope of the denial is limited just to polarity and not any other part of the content expressed by the antecedent assertion. Such an approach is promising, but many details remain to be worked out.

### 3.3.4 Summary

Summarizing to this point, I have shown that retorts and tag questions are prototypical high-VPE configurations, bringing the relevant diagnostic patterns out even more clearly than the examples with coordinated root clauses I appealed to in the previous chapter. Following intuitions from the literature on pronominalization and discourse coherence, I suggested that the various noteworthy properties uniting retorts and tag questions are related to their status as definitive high-VPE environments.

What remains to be determined is the syntactic analysis of high-VPE; I address this in §3.5. First, though, I take a moment to briefly describe some surprising facts revealing dialectal microvariation in ellipsis size within high-VPE contexts in English.

## 3.4 Microvariation in high-VPE

I have argued that the choice between low-VPE and high-VPE is not arbitrary, but determined by the ellipsis-antecedent environment (see in particular §2.5). However, to this point I have not said anything about the choice among different ellipsis sizes within high-VPE. As we saw in chapter 1, low-VPE has a fixed size, targeting only VoiceInflP (though its size can appear to fluctuate due to movement: see §2.3.1); on the other hand, high-VPE can come in different sizes, with a structure as small as VoiceSynP undergoing ellipsis, or one as large as PerfP. This was clearest in data involving auxiliary omission (§1.4.4), which I exemplify again below with data from one of the consummate high-VPE configurations just discussed, namely tag questions:<sup>57</sup>

---

<sup>57</sup>The forthcoming discussion will reveal that (83b)-(83d) are all subject to dialectal variation. I omit the % judgment marks here for expository purposes; see below.

*Variability in high-VPE size*

- (83) The cakes should have been being eaten,
- a. \*shouldn't they have been being ~~[<sub>VP</sub> eaten]~~?
  - b. shouldn't they have been ~~[<sub>voiceSYNP</sub> being eaten]~~?
  - c. shouldn't they have ~~[<sub>PROG</sub> been being eaten]~~?
  - d. shouldn't they ~~[<sub>PERF</sub> have been being eaten]~~?

We now have an understanding of why (83a) is bad: VPE must delete a structure at least as large as VoiceInflP (see the previous chapter), which contains *being*. However, the ellipsis possibilities in (83b)-(83d) require additional treatment. I discuss some empirical matters regarding this variability below, and defer theoretical discussion to the next section.

Like the low- vs. high-VPE distinction, it turns out that the choice among the different possible ellipsis sizes within high-VPE is not arbitrary. Unlike the low/high distinction, though, it appears that fluctuation within high-VPE is dialectally conditioned (although within-speaker variability exists too). The brief discussion below is based on an informal cross-dialectal survey whose results should be taken tentatively; nevertheless, the results are suggestive: speakers of American English (AE)<sup>58</sup> generally prefer eliding less material in high-VPE, while speakers of other varieties of English – including those from England, Scotland, Australia, and even Canada – generally prefer to elide more. For ease of reference, I will refer to this collection of varieties as the “Commonwealth Englishes” (CE).<sup>59</sup>

This microvariation in high-VPE size was first observed in Sailor (2009: fn. 20) for examples involving tag questions of the sort above in (83). However, the pattern also extends to retorts, as expected, as well as other high-VPE environ-

---

<sup>58</sup>I use the conventional term “American English” to refer strictly to the standard variety of English spoken in the United States, not North America generally. I mean no offense to my Canadian comrades for this exclusive use of “American”.

<sup>59</sup>I make no claims about other varieties of English spoken in the Commonwealth of Nations that I have not investigated.

ments (e.g. simple coordination structures). I will focus on the two extreme ends of the high-VPE size spectrum: the smallest constituent, VoiceSynP (as in (83b)), and the largest constituent, PerfP (as in (83d)). This is partly because this dichotomy exhibits the microvariation most clearly, and partly because more data from the intermediate possibility, namely ellipsis of ProgP, is necessary before a clear conclusion can be drawn.

The trend regarding dialectal microvariation in high-VPE size can be stated as follows. Speakers of AE generally accept or prefer the smallest ellipsis site within high-VPE – i.e. VoiceSynP – even though this can have the effect of stranding several redundant (given) auxiliaries outside the ellipsis site. They reject or disprefer the largest ellipsis site within high-VPE, i.e. PerfP. On the other hand, speakers of CE generally reject or disprefer the smallest high-VPE size, while accepting or preferring the largest. This is illustrated below in all of the high-VPE environments we have seen so far (though the effect is smallest in simple coordination environments). First, consider “small” high-VPE:

*Microvariation in “small” high-VPE (ellipsis of VoiceSynP)*

- (84) The paper should have been accepted, shouldn't it have been? *Tag*  
 AE: ✓  
 CE: ??/\*
- (85) A: It should have been accepted. B: NO it shouldn't have been! *LEP*  
 AE: ✓  
 CE: ??/\*
- (86) A: It should have been accepted. B: It should NOT have been! *CIP*  
 AE: ✓  
 CE: ??/\*
- (87) This paper should have been accepted, and that one should have been too. *Coord.*

AE: ✓  
CE: ?/??

Stacked auxiliaries stranded adjacent to the small high-VPE site (VoiceSynP) are fully acceptable in AE. In CE, however, this pattern is strongly dispreferred in the tag and retort contexts we have discussed in this chapter. There seems to be some internal variation among CE speakers in simple coordination contexts, as in (87). Now consider “large” high-VPE:

*Microvariation in “large” high-VPE (ellipsis of PerfP)*

- (88) The paper should have been accepted, shouldn't it? *Tag*  
AE: ??/\*  
CE: ✓
- (89) A: It should have been accepted. B: NO it shouldn't! *LEP*  
AE: ??/\*  
CE: ✓
- (90) A: It should have been accepted. B: It should NOT! *CIP*  
AE: ??/\*  
CE: ✓
- (91) (I'm betting that) John will have been to Paris by then, and his wife will too. *Coord.*  
AE: ?/??  
CE: ✓

Here, the pattern is completely reversed: when the high-VPE site is large (PerfP) and the redundant auxiliaries are elided, AE speakers' responses reflect general unacceptability (with the same minimized effect in coordination); on the other hand, CE speakers fully accept such sentences.



It may be that the patterns reported above change for CE speakers if inflectional material is made to differ between the A-clause and the E-clause. Tim Stowell (p.c.) reports that the presence of a contrast between the two clauses in the material occupying T<sup>0</sup> – modals in particular – leads to an exceptional preference for a small high-VPE site in his Canadian English. In my own American English, I find a large high-VPE site even more degraded than usual in such environments, meaning the two dialects seem to converge on a preference for (a) below:

- (92) Mary **might** have been dying to go to the concert...
- a. ...but Bill **must** have been ~~{dying to go}~~.
  - b. \*...but Bill **must** ~~{have been dying to go}~~.

Perhaps contrastive material in T<sup>0</sup> induces a general preference for small high-VPE in both AE and CE (although this cannot be tested in tags and retorts due to neophobia). Further investigation is needed.

Again, these results should be taken tentatively until a systematic investigation has been done. If the trend holds, though, it would represent an entirely novel area of microvariation among the English dialects, which would be a significant finding. In particular, the differential behavior of American English and Canadian English in this domain is especially striking. These varieties are known to vary in some ways (lexically, phonologically), but claims of syntactic variation between them are vanishingly few. If systematic investigation in this domain upholds the tentative results I have put forward here, then a novel domain of microcomparative work examining the syntactic behavior of these two closely related varieties of English opens up along with it.

Many open questions and issues remain to be investigated. One concerns the minimized effect in coordination environments, which is particularly noteworthy because, among the high-VPE environments examined above, it is the only one

which is not neophobic by nature. Another question that must be looked into involves the behavior of both AE and CE with respect to ellipsis of PerfP in high-VPE (“medium” high-VPE, so to speak). The above discussion is also restricted to high-VPE environments; one might wonder whether the choice to omit or pronounce VoiceInfl<sup>0</sup><sub>BE</sub> in low-VPE environments (see §2.3.1) is similarly varied across English dialects. Very tentative results suggest that the same general pattern arises there as it does here, with CE speakers generally opting for omission of as much inflectional material as possible when the option is available, while AE speakers tend to leave such material intact.

However, the most significant lingering question that arises from the preceding discussion concerns the source of these apparent high-VPE size preferences. I conclude with some remarks about this question in the next section.

### 3.5 Toward an analysis of high-VPE

In the previous chapter, I offered an analysis of the fine structure of low-VPE. A similar approach ought to be possible for high-VPE. As we saw above in (83) and elsewhere, patterns of auxiliary omission would seem to provide the most straightforward evidence of the variable amount of material that can be omitted in high-VPE. How can such variability be reconciled with the structural approach to ellipsis size taken here? I close this chapter with a few brief remarks on this question.

One possibility is simply that what you see is what you get: there are as many distinct sizes of high-VPE as there are attested possibilities diagnosed by these missing auxiliaries, as assumed in the previous section. Thus, in addition to VoiceSynP ellipsis, we would also have ProgP ellipsis and PerfP ellipsis, perhaps reflecting the output of a single predicate ellipsis operation that varies in size (within speakers) under the influence of presently unknown factors (in the spirit,

but not the implementation, of Ross 1969). This would seem to conflict with the suggestion from Holmberg (2001), Gengel (2009), Sailor (2012a), Aelbrecht and Harwood (2013), a.o. that VPE is non-pronunciation of a cyclic domain (see chapter 4). While some have proposed that cyclic domains (e.g. phases) can vary contextually, i.e. across sentences (den Dikken 2007), I am not aware of any claim that such variation is also possible within a single sentence, as would be necessary to capture the ellipsis possibilities in (83). It may be that the premise regarding the set of elidable constituents needs to be revisited. Perhaps the set is slightly larger than just the phases: for example, Bošković (2013) has recently argued that ellipsis is the non-pronunciation of either a phase or a phase complement, allowing for minor (binary) variation in ellipsis size similar to what we see above (which Bošković appeals to as support for his analysis, although he explicitly predicts (83d) to be ungrammatical, contrary to fact). Alternately, it may be that the premise is simply incorrect, and that cyclic domains play no role in the size of ellipsis. Of course, this would leave unaddressed the myriad parallels that exist between ellipsis and other syntactic operations (e.g. movement) with respect to the XPs they can(not) operate on, discussed previously.

Another possibility involves more than one ellipsis operation: low- and high-VPE configurations would be united in undergoing ellipsis of strictly VoiceInflP, then a separate “auxiliary ellipsis” operation would apply on top of that to eliminate the relevant auxiliaries above VoiceInflP to generate the various surface patterns of auxiliary omission we have seen characterizing high-VPE, as in (83). Indeed, this is just what Akmajian and Wasow (1975:§7) propose; however, Sag (1976:25-29) argues strenuously against this, proposing to capture the same set of data with only one VPE rule (alongside certain assumptions about constituency that are incompatible with contemporary theory). The Minimalist credo is to reduce the size of the grammar by eliminating ad hoc operations, so maintaining a special auxiliary ellipsis operation in addition to VPE would require strong sup-

porting evidence which, to my knowledge, does not exist. If the ellipsis possibilities in (83) can be accounted for using only those operations which can be independently motivated (perhaps nevertheless interacting in some complex way to be determined), then auxiliary ellipsis can be eliminated, as Sag proposes.

At present, I do not know whether such an account is possible, though I assume it is. We have already seen that variable ellipsis possibilities within a single low-VPE configuration can be reduced to the interactions of movement and a fixed-size VPE operation. I suspect that the same is true for the analogous state of affairs in high-VPE; but, lacking a specific proposal, I leave the matter open.

## CHAPTER 4

### On Grammatical Privilege

#### 4.1 Introduction

In this chapter,<sup>1</sup> I discuss an array of empirical phenomena beyond VPE that preferentially target what I am calling VoiceInflP (diagnosed by *being*), including “VP” topicalization, pseudoclefting, predicate questions and answers, etc. As we will see, in general, if an operation applies that divides up the verbal-inflectional domain, then with overwhelming frequency, its output lumps *being* and nothing higher together with the verb. Thus, VoiceInflP enjoys some sort of syntactic *privilege* that its neighbors do not.

Why should this be the case? In chapter 2, we saw that low-VPE always targets VoiceInflP, whose head is realized as *being* in progressive passive clauses. If the low-VPE site is always exactly VoiceInflP, this explains why *being* can never be stranded by VPE; however, we are left wondering what is so special about VoiceInflP, and why low-VPE is so picky about what it operates on, preferring this particular constituent to the many inflection-related projections immediately local to it (*vP*, *ProgP*, etc.). I suggested in §2.5 that the choice between low-VPE vs. high-VPE might be dictated by economy principles, with the grammar avoiding antecedent-containment whenever possible. While this can explain the distribution of these different sizes of VPE, it does not clearly extend to the observation I lay out below, namely that (low-)VPE is not alone in singling out VoiceInflP. I lay

---

<sup>1</sup>This chapter has its foundations in earlier work, Sailor (2012a).

out the data first, and then discuss their consequences.

## 4.2 The grammatical privilege of VoiceInflP

In this section, I discuss several empirical phenomena that pattern like low-VPE, in that they always and only target the constituent I have identified as VoiceInflP, despite an array of seemingly reasonable local alternatives. The preponderance of the data that follow clearly show that VoiceInflP has a privileged status in the grammar: it can participate in a host of operations that other projections local to it cannot. The data and discussion are below; later, in §4.3, I explore what the theoretical status of this privilege might be, and what consequences it might have for the theory of grammar.

Before we begin looking at the data, a few procedural points are in order. Except where mentioned otherwise, all of the examples below involve progressive passive clauses – i.e., those involving VoiceInfl<sup>0</sup><sub>being</sub> – in order to make clear that VoiceInflP is in fact the maximal constituent being manipulated in each of the phenomena discussed here.<sup>2</sup> The reader will naturally wonder how these phenomena behave when applied to clauses containing different inflectional configurations (i.e. different aspects, grammatical voices, etc.); though this is in some cases straightforward, I delay treatment of this question until later. Finally, to fully illustrate that VoiceInflP is privileged among the INFL projections in this respect, I give examples involving manipulation of each of the other INFL projections wherever possible, all of which turn out to be ungrammatical.

---

<sup>2</sup>See chapter 2 on the behavior of VoiceInfl<sup>0</sup> in non-progressive passives. As in chapter 2, the phenomena we discuss here can be shown to treat VoiceInflP the same even when it is not headed by *being*, though it requires a more delicate treatment, as we saw before. I return to this briefly below.

### 4.2.1 Argument 1: “VP” topicalization

“VP” topicalization (henceforth “VPT”) exhibits a number of striking similarities to VPE. In fact, the conditions which favor VPE are so similar to those licensing VPT that Johnson (2001) suggests VPT is involved in the derivation of VPE, a position recently argued against by Aelbrecht and Haegeman (2012). Our only concern here, though, is the category of the constituent manipulated by VPT.

In chapter 2, we saw that *being* could not be stranded outside of a VPE site, which follows if VPE always targets a fixed size of structure, and *being* is generated within that structure. By hypothesis, this structure is VoiceInflP. In VPT, a similar pattern arises: *being* cannot be stranded by the fronting operation. Extending the reasoning from chapter 2, this indicates that VPT does not target *v*P (or VP), but rather VoiceInflP:

*VPT only targets VoiceInflP*

- (1) If Mary says that the cakes will have been being eaten, then...
  - a. \*...[<sub>v</sub>P eaten], they will have been being  $t_{vP}$
  - b. ...[VoiceInflP being eaten], they WILL have been  $t_{\text{VoiceInflP}}$
  - c. \*...[<sub>Prog</sub>P been being eaten], they WILL have  $t_{\text{ProgP}}$
  - d. \*...[<sub>Perf</sub>P have been being eaten], they WILL  $t_{\text{PerfP}}$
  - e. \*...[<sub>TP</sub> WILL have been being eaten], they  $t_{\text{TP}}$

Thus, when it occurs, VPT always and only targets VoiceInflP, clearly indicating that phrase’s privileged interface status. That is, when all the inflectional heads are filled (see below for other environments), the only constituent capable of undergoing the movement that characterizes VPT is the constituent I have identified as VoiceInflP. For concreteness, I assume that VPT involves a single step of movement to a projection (assume TopP) in the clausal left periphery.<sup>3</sup> The tree below

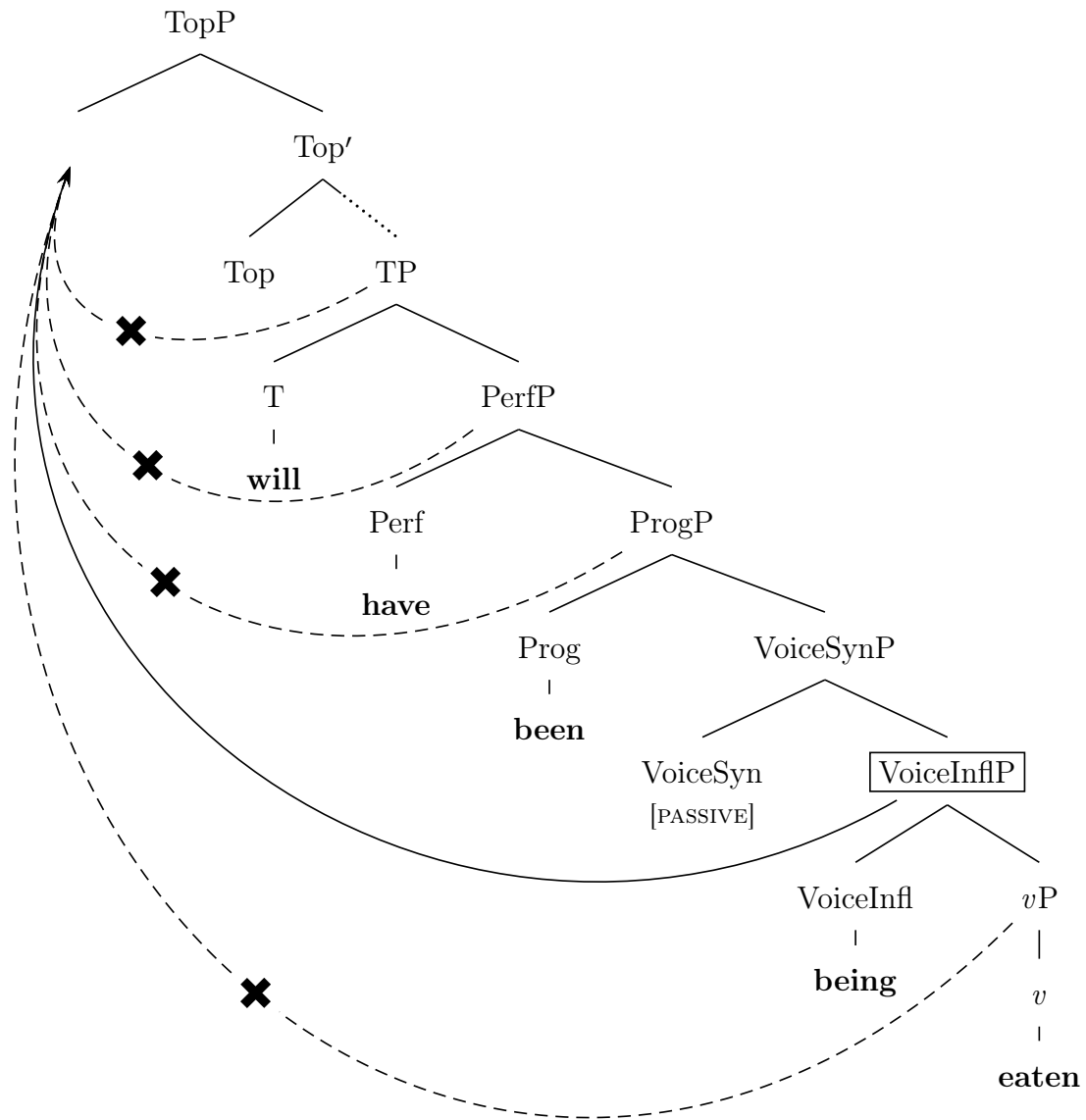
---

<sup>3</sup>It is possible that VPT involves movement to a low-peripheral position (i.e. a low TopP: Belletti

reflects the possible and impossible movements revealed by the data in (1), where impossible movements are represented by dashed lines marked with **✗**:

*VPT moves VoiceInflP*

(2) ...being eaten, they will have been



2004) before movement to the clausal periphery; however, I leave this aside. See Aelbrecht and Haegeman (2012) for discussion. For clarity, I obscure the movements presumed to occur in the derivation of the passive in the tree below.



#### 4.2.2 Argument 2: Reduced relative clauses (“whiz-deletion”)

Occurrences of so-called *whiz* (*wh-* + *is*) *deletion* are reduced relative clauses of the type in (3), characterized by the absence of both an overt relativizer and a tensed form of BE:

*Reduced relative clauses (“whiz”-deletion)*

- (3) a. The nun who is beating that man → The nun beating that man  
b. The images which were shown on TV → The images shown on TV

Reduced relatives can be either progressive or passive, as above, but they can also be both, shown below.

The crucial observation for our purposes is that reduced relatives are sensitive to the same cutoff point as VPE and, as we have just seen, VPT: that is, reduced relatives can contain inflectional material up to *being*, but not anything higher. I assume that reduced relatives are the smallest stretches of structure capable of bearing a relative-modificational relationship to some nominal head; however, since the purpose of this section is simply to illustrate the special syntactic status of the constituent headed by *being* (by hypothesis: VoiceInflP), I will not commit myself to a specific analysis of reduced relative clauses here (though I assume the correct analysis does not actually involve true deletion of the *wh-* + *is* string, which is a non-constituent with no antecedent). See Iatridou et al. (2001) for discussion and a list of references involving reasonable non-deletion accounts of reduced relatives. What is important for us is the following observation: when the content of a reduced relative involves a progressive, passive verbal predicate,<sup>4</sup> then the reduced relative apparently cannot contain any inflectional projection above

---

<sup>4</sup>Obviously, reduced relative clauses can involve other predicate (and inflectional) types as well, e.g. APs and PPs. This is also true for every other phenomenon we consider in this chapter: i.e., VPE and VPT can involve non-verbal predicates (and verbal predicates in non-progressive-passive contexts). Since this issue is not specific to reduced relatives, I leave it aside.

VoiceInflP (whose head is *being*), nor can it be smaller than VoiceInflP in such a context:

*Reduced relatives (“whiz”-deletion) only target VoiceInflP*

- (4) Someone has been being beaten all day long, and frankly...
- a. \*...the man [<sub>vP</sub> beaten] deserves it.
  - b. ...the man [<sub>VoiceInflP</sub> being beaten] deserves it.
  - c. \*...the man [<sub>ProgP</sub> been being beaten] deserves it.
  - d. \*...the man [<sub>PerfP</sub> has been being beaten] deserves it.

Once again, we see that when the grammar has to split up the INFL-domain, it consistently draws the line at the same constituent: VoiceInflP. (See §4.2.7 for evidence suggesting the reduced relatives in (3) are also VoiceInflPs, despite lacking *being*.)

### 4.2.3 Argument 3: Pseudoclefting

Specificational pseudoclefts also exhibit a sensitivity to the constituent headed by *being* in progressive passives. Such pseudoclefts superficially appear to involve a copula separating a *wh*- clause from some sort of predicate (in either order):

- (5) a. [What John bought] was a book about syntax.  
b. A book about syntax was [what John bought].

Many competing analyses for (specificational) pseudoclefts have been proposed in the literature; for a detailed survey, see den Dikken (2006). Our concern here is not with their proper analysis, but rather with a particular empirical pattern exhibited by the predicate portion of such pseudoclefts. That is, the size of that portion of the pseudocleft seems to be constrained in a now-familiar way. As shown below,

in a progressive-passive context involving a lexical verb, the predicate portion of such pseudoclefts once again corresponds exactly to the constituent headed by *being*, namely VoiceInflP. I show this below using inverse pseudoclefts like that in (5b) in order to rule out the irrelevant predicational reading (but the canonical order patterns the same way):<sup>5</sup>

---

<sup>5</sup>There are some speakers who find it ungrammatical to pseudocleft any verbal constituent (i.e., (6) is entirely bad); I leave this aside, as it seems to bear on pseudoclefting in general, and not the INFL-domain in particular.

*Specificational pseudoclefts only target VoiceInflP*

- (6) A: John should have been being praised. B: No, ...
- a. \*...[<sub>vP</sub> CRITICIZED] is what John should have been being.
  - b. ...[VoiceInflP being CRITICIZED] is what John should have been.
  - c. \*...[<sub>ProgP</sub> been being CRITICIZED] is what John should have.
  - d. \*...[<sub>PerfP</sub> have been being CRITICIZED] is what John should.
  - e. \*...[<sub>TP</sub> should have been being CRITICIZED] is what John.

Some of these strings might be ruled out for a variety of reasons (e.g. (6e)), but the important point here is that the grammar once again exhibits preferential treatment for the constituent headed by *being* in such contexts – VoiceInflP – rather than other local alternatives.

Various accounts have argued that the *wh*- clause in specificational pseudoclefts is a free relative clause, while other accounts have argued that the *wh*- clause is a genuine *wh*- question, with the remaining material serving to answer that question in a sort of topic-comment configuration: see den Dikken (2006:§5). A common theme connecting these analyses is that there is some sort of dependency (syntactic or semantic) between the *wh*- clause and the remaining predicate portion. Thus, the pattern above may reflect a privileged aspect of VoiceInflP that we have already seen evidence of, i.e. VPT (if the predicate portion of such pseudoclefts involves VPT followed by clausal ellipsis, as den Dikken 2006 discusses). It may also reflect a privileged aspect of VoiceInflP that we will see more examples of in a moment: namely, that it is the only INFL projection capable of being rendered as a *wh*- type pronominal (and perhaps as non-*wh*- pronominal types as well, e.g. *do it/do so*, but I leave this aside). Regarding the latter point, an initial position one could take would be to assume that a *wh*- phrase has no independent syntactic category, but instead bears whatever category is licensed in the position from which it originates or with which it otherwise bears a dependency (i.e., *who*

is a DP because it originates in positions where DPs are licensed). If this were on track, then the pattern in (6) could be seen as not importantly different from the VPT examples in §4.2.1: among the INFL projections, only VoiceInflP is capable of moving, even if it is realized phonologically as a *wh*- phrase. Additional evidence of this is laid out in the next subsection, where we examine predicate *wh*- questions directly.

#### 4.2.4 Argument 4: Predicate *wh*- questions

The next argument comes from predicate *wh*- questions, an empirical domain that has received very little attention in the literature, despite providing fertile ground for linguistic inquiry. As the following examples show, the gap left by a predicate *wh*- question must correspond to the constituent we now know of as VoiceInflP, or the result is ungrammatical:<sup>6</sup>

*Wh*- phrases in predicate questions can only correspond to VoiceInflP

- (7) If Galileo shouldn't have been being PERSECUTED, then...
- a. \*...what SHOULD he have been being  $t_{vP}$  ?
  - b. ...what SHOULD he have been  $t_{\text{VoiceInflP}}$  ?
  - c. \*...what SHOULD he have  $t_{\text{ProgP}}$  ?
  - d. \*...what SHOULD he  $t_{\text{PerfP}}$  ?

Once again, VoiceInflP exhibits special grammatical behavior, whereas the other INFL projections do not. As noted above, this pattern might be a direct reflection of the same general effect seen with pseudoclefts; in other words, these data and those from pseudoclefts might not compose two distinct arguments for the privilege of VoiceInflP, but rather be superficially different demonstrations of the same basic type of privilege. As I suggested in the previous subsection, the nature of that

---

<sup>6</sup>I leave aside cases that require the use of the light verb *do*, since such cases form a class that clearly involves a different syntax than what is considered here.

privilege for these two empirical domains might be based in pronominalization: it may be that VoiceInflP is the only inflectional projection capable of being rendered as a *wh*- (or perhaps other types of) pronominal. Regardless, the basic empirical point stands: the constituent headed by *being*, VoiceInflP, is once again seen to be special in ways that its neighbors are not.

This empirical pattern continues in the next subsection, which considers answers to the types of predicate questions discussed here.

#### 4.2.5 Argument 5: Fragment answers to predicate *wh*- questions

In a demonstration of a canonical constituency test, a single VoiceInflP can be made to stand alone in composing a complete utterance, just in case it serves as a fragment answer to a predicate *wh*- question of the type we saw in the previous subsection. Moreover, fragment answers corresponding to other projections in the INFL-domain are not possible with the relevant reading:

*Predicate fragment answers can only correspond to VoiceInflP*

- (8) If the room shouldn't have been being DECORATED, what SHOULD it have been  $t_{\text{VoiceInflP}}$  ?
- a. #...[ $v\text{P}$  made child-proof].<sup>7</sup>
  - b. ...[VoiceInflP being made child-proof].
  - c. \*...[ProgP been being made child-proof].
  - d. \*...[PerfP have been being made child-proof].
  - e. \*...[TP should have been being made child-proof].

---

<sup>7</sup>Although (8a) is syntactically well-formed, it is interpreted as the non-progressive “it should have been made child-proof”, similar to the aspectual mismatch effect we saw in chapter 2 (which is not obvious since the marker of progressive, *being*, is “inside” the *wh*- phrase: see §4.2.4). Based on arguments discussed there, we cannot confirm that this is truly a  $v\text{P}$  fragment answer and not a VoiceInflP fragment answer with an empty VoiceInfl<sup>0</sup>, so I set this datum aside. Note that we cannot appeal to e.g. an obligatorily-progressive idiom here, since those discussed in chapter 2 cannot be passivized (and I am unaware of any progressive-passive idioms).

This pattern is perhaps unsurprising, given what we saw for the predicate *wh*-questions, above: if all predicate *wh*- words stand in for a VoiceInflP, then we would expect exactly the pattern we see here, namely that their corresponding answers spell out the questioned VoiceInflP. In other words, the pattern in (8) might be reducible to a more general phenomenon relating to pronominalization or anaphora, in which case it would not stand alone as an independent argument for the privilege of VoiceInflP. This would have little bearing on the position I take here, which is merely to point out the special behavior of this particular constituent. We can see it either as special in X different environments, or in all environments involving a general phenomenon Y (e.g. pronominalization); the crucial observation is simply that this constituent is special in a significant way.

#### 4.2.6 Argument 6: Participle preposing

The last empirical point I discuss here regarding the privilege of VoiceInflP involves participle preposing of the sort first discussed in Emonds (1970), where progressive and passive participles appear to the left of the heads that presumably select them. As before, we see that *being* cannot be left behind (Samko 2013), showing that VoiceInflP (and only VoiceInflP) is undergoing the preposing operation. Independent constraints on participle preposing prevent us from using a fully-realized inflectional array in the way we have been doing so far; however, the same points can be made piecemeal:

*Participle preposing only targets VoiceInflP*

- (9) Your new desk will be being delivered today; however, there's bad news:
- a. \*<sub>vP</sub> [delivered later] will be being all the tools needed to build it.
  - b. <sub>VoiceInflP</sub> [being delivered later] will be all the tools needed to build it.
  - c. \*<sub>ProgP</sub> [be being delivered later] will all the tools needed to build it.
  - d. \*<sub>TP</sub> [will be being delivered later] all the tools needed to build it.

For completeness, the following example involving perfect aspect shows that PerfP, unsurprisingly, cannot prepose either:

- (10) \*<sub>[PerfP</sub> have been bringing up the rear] will everyone who will have failed to arrive on time.

Samko (2013) argues that this sort of participle preposing has special information-structural licensing properties with syntactic effects that might be seen to distinguish the phenomenon from the superficially-similar phenomenon of VPT we discussed earlier. See *op. cit.* for an in-depth analysis of this participle preposing phenomenon, the details of which are not essential for us here. The crucial observation is simply that VoiceInflP can undergo an operation – in this case, movement (which may be distinct from VPT) – that its neighboring projections cannot, once again revealing its privileged status in the grammar.

#### 4.2.7 Afterword: Confirming the results using GET-passives

Recall that in §2.4, we saw evidence that the GET in (certain kinds of) GET-passives occupies the VoiceInfl head. Crucially, we saw that this variety of GET never undergoes raising, even when such movement would be possible (or required) by an analogous VoiceInfl head realized by a form of BE. This allowed us to use data from GET-passives to eliminate the possibility that some lexical quirk of *being*, rather than a general property of VoiceInfl<sup>0</sup>, was responsible for certain patterns involving low-VPE. Here, we can extend the argument to these other phenomena we have just looked at. Below, we will see that VoiceInfl<sup>0</sup><sub>getting</sub> patterns like VoiceInfl<sup>0</sup><sub>being</sub> in every way with respect to these phenomena, showing that the patterns discussed throughout this section cannot be due to a lexical property of *being* in particular, but instead must arise from basic syntactic properties of the inflectional domain in progressive-passive sentences.



GET-passives in VPT

- (11) If Mary says that John will have been getting slandered, then...
- a. ...getting slandered, he will have been.
  - b. \*...slandered, he will have been getting.

GET-passives in reduced relatives (“whiz-deletion”)

- (12) The man [getting beaten] deserves it.

GET-passives in pseudoclefting

- (13) A: John should have been getting praised. B: No, ...
- a. \*...<sub>[vP CRITICIZED]</sub> is what John should have been getting.
  - b. ...<sub>[VoiceInflP getting CRITICIZED]</sub> is what John should have been.

Pseudoclefting provides us with a way of distinguishing two homophonous instances of *getting*: one which occupies VoiceInfl<sup>0</sup>, as shown above, versus one which is a regular main verb (with a nominal object), as below:

- (14) A: John should have been getting praise. B: No, ...
- a. \*...getting CRITICISM is what John should have been.
  - b. \*...receiving CRITICISM is what John should have been.

This stark contrast between (13b) and (14a) here illustrates two different usages of *get*, the second of which behaves like a main verb (not associated with passive), as seen in the synonymous example in (14b) involving *receive*, an uncontroversial main verb. Thus, the GET associated with passives is not a main verb, consistent with it being in VoiceInfl<sup>0</sup>.

We continue now with the remainder of the examples:

GET-passives in predicate wh-questions and answers

- (15) Q: If Galileo shouldn't have been getting PERSECUTED, then...
- a. ...what SHOULD he have been (\*getting)?
  - b. A: Getting PRAISED.

GET-passives in participle preposing

- (16) Your new desk will be getting delivered today; however, there's bad news:
- a. \*[<sub>vP</sub> delivered later] will be getting all the tools needed to build it.
  - b. [<sub>VoiceInflP</sub> getting delivered later] will be all the tools needed to build it.

Summing up, we see that *getting* patterns like *being* with respect to all the relevant data. Given our earlier conclusion that *getting* is also in VoiceInfl<sup>0</sup>, this provides independent support for the claim that VoiceInflP is the relevant source of these patterns, not a lexical property of *being* in particular. The data also continue to show that VoiceInflP, and not e.g. *vP*, has special status in syntax.

#### 4.2.8 Summary

Throughout this section, I have attempted to show that the string [ *being* + *vP* ] has special grammatical privilege, allowing it to participate in an array of syntactic operations that even other nearby INFL constituents (*vP*, ProgP, etc.) cannot.<sup>8</sup> This string corresponds to the projection I identified as VoiceInflP in chapter 2, since the highest overt head can (but need not) be *being*.

In the next section, I explore what the source and status of this grammatical

---

<sup>8</sup>We saw in chapter 2 that the associates of existentials surface in Spec-VoiceInflP, which might be taken as an additional example of privilege for this projection. I do not discuss it here, but see Harwood (2013a) for arguments that this distribution of existential associates diagnoses the phasal status of this constituent. See below for more on phases.

privilege is. Further study may in fact yield evidence that the constituent I have labored to identify and describe here actually corresponds to some silent functional projection above VoiceInflP. If, as I explore in the next section, the constituent being diagnosed above is some sort of cyclic domain with parallels in the nominal and clausal regions (e.g. a phase), then perhaps the likeliest candidates for such other functional projections would be those associated with Belletti’s (2004) “low periphery” (TopP, FocP, etc.; see also Butler 2003 for low ForceP and FinP), since it seems likely that each cyclic domain is endowed with its own periphery of this sort. In the event that such a discovery is made, it would be entirely consistent with both the thesis and supporting data I put forward here, so long as the constituent in question is still found to be located beneath what I refer to as ProgP (Bjorkman 2011’s “AspP”) and VoiceSynP (in low-VPE contexts).

### 4.3 On the theoretical status of privileged XPs

In the previous section, we identified a variety of ways in which VoiceInflP is special: even though it is surrounded by featurally-similar projections, these projections resist undergoing an array of operations that VoiceInflP can undergo. Thus, VoiceInflP stands out as having some sort of interface privilege: it gets to participate in operations that its next-door neighbors cannot. Certainly, VoiceInflP is not the only such privileged XP (henceforth, PXP), but it will be the primary subject of the discussion to follow. Of central concern here is understanding what the theoretical status of this privilege is, and what it might mean for the theory at large.

Since at least *Barriers* (Chomsky 1986), we have known that certain specific maximal projections – e.g., that of the main verb – hold special significance in the syntactic derivation (and of course work on the cycle within syntax goes back before this: see Bresnan 1971). Traditionally, this significance has been identified

and characterized in terms of constraints on extraction, e.g. the enforcement of successive-cyclic movement, and these special nodes have previously been labeled bounding nodes, barriers, and, most recently, phases (Chomsky 2000, 2001). Thus, the theory already contains some notion of privilege for particular constituents. Perhaps, then, PXP's such as VoiceInflP are simply phases (or Spell-Out domains, or whatever analytical near-equivalent one invokes to subsume the empirical coverage of bounding nodes/barriers, etc.). Below, I will entertain a few arguments in favor of this position (based on arguments made previously in Sailor 2012a; see also Harwood 2013a).

It should be noted from the outset that the core notions discussed in this section – Spell-Out domains, phases, etc. – are not well-understood in the literature, and are indeed matters of ongoing disagreement and debate at even the most fundamental levels, not the least of which concerns the seemingly-simple matter of defining their fundamental empirical properties (thus allowing for their identification with surface diagnostics). To my knowledge, no consensus exists on such matters; see Gallego (2010) for an overview of some of the debate, which we will return to later. This raises some questions: if there is no accepted notion of what phases/Spell-Out domains are, or how they can be empirically established, then what theoretical value does the notion have, and why would we want to equate the privilege we observed above to such a construct?<sup>9</sup> The answer, however unsatisfying, is this: although there is no agreement on exactly what phases/Spell-Out domains do (and don't do) in the grammar, or how to precisely identify them, there is general agreement that such constructs must exist in some form (meaning of course that more work is necessary to reveal their core properties). Since a guiding question of the present discussion is “why this constituent, and not the one right next to it?”, it makes sense to appeal to the only property in contemporary syntactic theory that affords a particular category with the ability to behave

---

<sup>9</sup>I thank Gary Thoms for emphasizing the relevance of these questions.

differently than its neighbors in all contexts in which it occurs.<sup>10</sup> That property is phasehood (or phase complementhood; I will not bother to distinguish them carefully here) – ill-understood as it may be – and thus I will start with the controversial assumption (to be discussed further shortly) that phasehood can be directly diagnosed on empirical grounds. The suggested connection between privilege and phasehood is tentative, and indeed may find itself at odds with some future consensus in the field regarding the definition of the cyclic domains in syntax. This would nevertheless leave intact the core finding of the previous section, namely that particular XPs (in this case, VoiceInflP) enjoy syntactic privilege of a specific sort in the grammar, and that we require an understanding of what this privilege is.

The practical justification for this approach is as follows. The privilege of VoiceInflP that we saw in chapter 2 and above is entirely consistent with claims in the earlier literature on phase theory, which argued that phases could be identified on the basis of particular interface properties, such as being phonetically and semantically independent or isolable (i.e. being movable, or having all arguments saturated: see Chomsky 2004:124). This straightforwardly describes what we have seen for VoiceInflP: in the case of low-VPE, VoiceInflP is assigned a null phonetic value at PF, making it phonetically independent by definition. The same reasoning applies to its mobility in VPT, participle preposing, etc. Thus, if such behavior is diagnostic of phasehood, then the privilege exhibited by VoiceInflP may well diagnose its status as the verbal phase. I discuss some of these arguments in more detail below.

---

<sup>10</sup>Where this behavior is understood to be importantly different from run-of-the-mill selectional/featural differences that distinguish categories generally.

### 4.3.1 Ellipsis

There is an intuitive connection to be made between the special behavior of Voice-InflP we have witnessed and the notion of phasehood (or the slightly more theory-neutral notion of a cyclic Spell-Out domain). Let us focus on the first empirical phenomenon that drew our attention to this projection, namely VPE. If we arm ourselves only with Minimalist technology, then the most parsimonious treatment of ellipsis is one appealing to cyclic Spell-Out: that is, ellipsis arises when a Spell-Out domain fails to be realized at the sensorimotor interface ( $\approx$ PF). This is precisely the approach originally hinted at in Chomsky (1995), and later employed by Holmberg (2001) and developed in more depth by Gengel (2009).<sup>11</sup>

As Gengel points out, the set of constituents capable of undergoing ellipsis aligns closely with, and is perhaps identical to, the set of constituents corresponding to the Spell-Out domains (i.e., the constituents which, during the course of the derivation, are sent to the interfaces for computation and pronunciation). In one particular theory of cyclic Spell-Out – that of Chomsky’s (2001) *phases* – the Spell-Out domains are the complements selected by  $C^0$ ,  $v^0$ , and (it is implied)  $D^0$ , also known as the *phase heads*. Under standard assumptions, the syntactic categories of a given language’s Spell-Out domains are fixed across clauses,<sup>12</sup> and those categories selected by the phase heads are typically taken to be TP, VP, and NP, which corresponds exactly to the elliptical domains of sluicing, VPE, and NP ellipsis. Thus, an approach that treats ellipsis as the null realization of a Spell-Out domain has conceptual and theoretical appeal in aligning the small inventory of ellipsis domains with the small inventory of domains that are argued to be privileged in the grammar on entirely independent grounds (e.g. due to constraints they impose on the derivation).

---

<sup>11</sup>See Aelbrecht (2010: §3.2.4) for potential counterarguments involving extraction asymmetries.

<sup>12</sup>See Gallego (2010: §2.4.2) and Bošković (2013) and references therein, who argue that phases can fluctuate in size within a language.

Before going any further, though, it is worth saying something about the inventory of syntactic categories generally, and about those that are traditionally taken to be relevant to the cycle in particular. Chomsky's own work developing the Minimalist Program and Bare Phrase Structure (Chomsky 1995, 2000, etc.) endorses only an extremely impoverished inventory of core (non-adjunct) categories, one comprising perhaps nothing more than C, T, *v*, V, P, D, and N. At this point it should be clear that I do not assume such a sparse inventory; indeed, there is overwhelming evidence in favor of a far richer array of categories: see, for example, much influential work in Syntactic Cartography (Rizzi 1997, Cinque 1999, and many others). If we endow the inventory with the categories I appeal to here (along with many others stemming from work in the Cartographic approach), it seems unlikely that the proper categorial identity of the phases would turn out to be those typically assumed, i.e. CP, *v*P, and NP. Indeed, these categories may not even exist within a more articulated approach to phrase structure: for example, what was previously thought of as the atomic category CP came to be thought of as a "layer" or a "region" containing several distinct functional projections, none of which actually bears the label CP (Rizzi 1997). To that end, from the perspective of the present analysis, it seems that the PXP in the verbal-inflectional domain is not *v*P or VP, but rather what I have identified as VoiceInflP. Thus, while I will draw upon prior work on phases and Spell-Out domains for the remaining discussion, I assume that the constituents that those works have identified as being relevant to the interfaces do not actually correspond to the traditionally-held labels CP, *v*P, and NP, and that those labels are simply approximations of the relevant categories within what is otherwise a highly articulated array of projections. Indeed, work of the present sort is what is necessary to reveal a more precise view of just what the relevant categories might be.

Returning now to ellipsis, relating this phenomenon to an effect of Spell-Out has its advantages. For example, any theory of ellipsis must necessarily burden

the grammar with a process for suppressing pronunciation; however, if ellipsis sites are simply phonetically-null Spell-Out domains, then this suppression is less a “process”, and more a “diacritic” added to a Spell-Out domain (under the appropriate licensing configuration) instructing PF to render the nodes within the domain as silent. This is reminiscent of the phonological properties that Merchant (2001, 2004) assigns to his [E] feature, which also make use of PF instructions for non-pronunciation. However, the chief difference between the Spell-Out-based approach discussed above and Merchant’s deployment of [E] is that the former, as I said, involves syntactic objects that are privileged at the interface for reasons independent of ellipsis, whereas the latter involves a morphosyntactic feature whose lexical distribution does not obviously correspond to independent interface principles.<sup>13</sup>

### 4.3.2 Propositionality

Chomsky (2001:12, 2004:124) also claims that phases can be characterized as being “propositional” in nature. This is also potentially compatible with VoiceInflP, perhaps providing an additional argument for its phasal status. The main reason to adopt such an approach is the following. A great deal of recent literature argues that the array of projections typically thought of as being at the left edge of the clause – namely, those dedicated positions associated with topic, focus, polarity, and so on – can also be found at the left edge of the clause-internal verbal domain (Jayaseelan 2001, Butler 2003, Belletti 2004). This implies a significant structural similarity between the clausal and verbal domains, in addition to their shared

---

<sup>13</sup>As I mentioned, even if ellipsis is the simple non-pronunciation of a Spell-Out domain, the PF component presumably still needs some indication that the domain in question can be felicitously rendered silent; i.e., that it stands in an ellipsis-licensing configuration in the syntax, and a mutual-entailment relation to a discourse-salient antecedent in the semantics (*e-givenness*: Merchant 2001). The [E] feature provides a convenient way of ensuring these requirements are met during the derivation. An advantage of the Spell-Out-based approach advocated here, then, is that it radically constrains the set of heads that can bear [E], limiting it to just the set of phase heads (and, perhaps, the ellipsis licensers as well).



status as Spell-Out domains. Belletti (2004) goes as far as to suggest that their shared structural “peripheries” are in fact the source of their privileged status at the interfaces.

Butler (2003, following the original Belletti manuscript) further develops this proposal by bringing in data from modals, specifically to build on Chomsky’s suggestion that the verbal Spell-Out domain is propositional. In brief, Butler reviews existing arguments that expressions of root modality take scope underneath the surface position of subjects (Spec-TP) and concludes that root modality is associated with specific projections (ForceP and FinP) in the “low”, verbal periphery. This is directly analogous to its epistemic counterpart, which is associated with the same projections in the “high”, clausal periphery. As Butler points out, modal expressions are propositional operators, meaning they only combine with complete propositions. Given that root modals scope underneath surface subjects but must combine with a proposition, it follows that there is a complete proposition located below the TP position in the clausal hierarchy. By parity of reasoning, the scopal properties of such modals can be thought of as diagnosing the verbal Spell-Out domain, which, as Chomsky claims, is propositional in nature. In the details of Butler’s (2003) analysis, which I have simplified here, he argues that root necessity modals are generated in the low ForceP, whereas root possibility modals are generated in the low FinP, and it is the head of FinP that, for Butler, selects the verbal Spell-Out domain. He uses the label *vP* for this projection, as is typical; however, as I have argued throughout, we know there must be an inflectional position, VoiceInfl, immediately above the position of the main verb. Nothing in Butler’s argumentation would suggest that this inflectional position is above the low periphery (and thus outside the propositional/Spell-Out domain we are concerned with). Given the preceding arguments suggesting that VoiceInflP, and not *vP*, is the privileged domain in that region of the clause, it stands to reason that VoiceInflP is located beneath the low periphery, and is thus part

of the proposition that the aforementioned modal types combine with. Following Chomsky's claims, this is equivalent to saying that VoiceInflP reflects the edge of the verbal Spell-Out domain, as I have been suggesting.

### 4.3.3 Summary and outlook

We have just seen that VoiceInflP is capable of many syntactic feats that its local neighbors cannot perform. I have suggested here that the interface privilege characterizing VoiceInflP might be fruitfully connected to the preexisting notion of interface privilege known as phasehood or Spell-Out domain status. Of course, if this connection proves to be incorrect, no part of the preceding discussion is eroded: we would be left with a collection of empirical observations that require explanation.

Indeed, there may be reasons to doubt that phasehood status is the correct attribution for PXP. Recent work on phase theory has retreated from and even turned against the assumption that phases should be diagnosable on their interface and semantic properties. For example, Chomsky's remarks in Gallego (2010:55) (and the surrounding discussion there) argue against the notions of propositionality and PF-isolability being definitional properties of phases, despite that Chomsky himself had, only a few years earlier, offered those as the only clear external properties of phases. Gallego argues instead that phases are best characterized by particular Case and agreement properties (i.e., properties relating to the valuation/deletion of uninterpretable features) rather than simply by their own external interface properties, or even by the movement-related effects they induce, which are now apparently regarded as epiphenomenal (Chomsky 2008, Gallego 2010:§2.2). If the claims of these more recent works in phase theory are to be sustained, then perhaps the privilege we identified above – which is based on decidedly syntactic properties related to movement, ellipsis, etc. – should not be taken to diagnose phasehood. As Gallego (2010:335) points out, there are other types

of cyclic domains that have been argued for in the literature (relating to stress, tone sandhi, focus, etc.) that may or may not align with the inventory of phases. Likewise, Müller (2010:41) notes that although phases are typically thought of as being domains of cyclic Spell-Out, even the earliest work on phase theory argues that it is in fact the complement of the phase head (the “phase complement”) that is the Spell-Out domain, meaning the two domain-based constructs – phases and regions of cyclic Spell-Out – are correlated, but not aligned.

Thus, open questions remain about the privilege that certain projections can have in the grammar; there may be multiple different types or layers of privileged domains. The methodology employed here clearly isolates and highlights one such domain type, and it does so in a restrictive way, such that other potential candidates – that is, other local projections – can be shown to behave differently with respect to the relevant diagnostics. Determining whether this methodology diagnoses phases or Spell-Out domains versus some other type of cyclic domain will require a more precise understanding of what phases are and how they can be identified, something that is independently needed quite badly in the field.

## BIBLIOGRAPHY

- Aelbrecht, L. 2010. *The Syntactic Licensing of Ellipsis*. John Benjamins, Amsterdam/New York.
- Aelbrecht, L. and L. Haegeman. 2012. VP Ellipsis is not licensed by VP Topicalization. *Linguistic Inquiry* 43:591–614.
- Aelbrecht, L. and W. Harwood. 2013. To Be or Not To Be Elided: VP Ellipsis Revisited. lingBuzz/001609.
- Ahn, B. 2011. Explorations of Voice, Identity and the Self. Presented at Parallel Domains: A workshop in honor of Jean-Roger Vergnaud, Los Angeles, CA.
- Ahn, B. 2013. Universality and Subject-Oriented Reflexivity. Presented at ICL 19, University of Geneva.
- Ahn, B. In Progress. Giving Reflexivity a Voice: Twin Reflexives in English. Ph.D. thesis, UCLA.
- Aijmer, K. 2002. *English Discourse Particles*. John Benjamins.
- Akmajian, A. and T. Wasow. 1975. The Constituent Structure of VP and AUX and the Position of the Verb BE. *Linguistic Analysis* 1:205–245.
- Arregui, A., C. Clifton, L. Frazier, and K. Moulton. 2006. Processing Elided Verb Phrases with Flawed Antecedents: The Recycling Hypothesis. *Journal of Memory and Language* 55:232–246.
- Baltin, M. 2002. Movement to the Higher V is Remnant Movement. *Linguistic Inquiry* 33:653–659.
- Belletti, A. 2004. Aspects of the low IP area. In *The Structure of CP and IP*, edited by L. Rizzi, volume 2 of *The Cartography of Syntactic Structures*. Oxford University Press, pages 16–51.
- Bjorkman, B. 2011. BE-ing default: the morphosyntax of auxiliaries. Ph.D. thesis, MIT.
- Bolinger, D. 1946. Thoughts on ‘Yep’ and ‘Nope’. *American Speech* 21:90–95.

- Bouton, L. 1970. Antecedent-Contained Proforms. In *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*. University of Chicago Press, pages 154–167.
- Bošković, Ž. 2013. Now I’m a Phase, Now I’m Not a Phase: On the Variability of Phases with Extraction and Ellipsis. *Linguistic Inquiry* 45.
- Brasoveanu, A., D. Farkas, and F. Roelofsen. 2011. Polarity Particles and the Anatomy of N-Words. In *Proceedings of Sinn und Bedeutung 16*, edited by A. A. Guevara, A. Chernilovskaya, and R. Nouwen. MIT Working Papers in Linguistics, pages 99–112.
- Breitbarth, A., K. De Clercq, and L. Haegeman. 2013. The Syntax of Polarity Emphasis. *Lingua* 128:1–8.
- Bresnan, J. 1971. Sentence Stress and Syntactic Transformations. *Language* 47:257–281.
- Butler, J. 2003. A Minimalist treatment of modality. *Lingua* 113:967–996.
- Chomsky, N. 1957. *Syntactic Structures*. Mouton.
- Chomsky, N. 1986. *Barriers*. Linguistic Inquiry Monograph. MIT Press.
- Chomsky, N. 1995. *The Minimalist Program*. MIT Press.
- Chomsky, N. 2000. Minimalist Inquiries: The Framework. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, edited by R. Martin, D. Michaels, and J. Uriagereka. MIT Press, pages 89–155.
- Chomsky, N. 2001. Derivation by Phase. In *Ken Hale: A Life in Language*, edited by M. Kenstowicz, chapter 1. MIT Press, pages 1–52.
- Chomsky, N. 2004. Beyond Explanatory Adequacy. In *Structures and Beyond*, edited by A. Belletti, volume 3 of *The Cartography of Syntactic Structures*, chapter 3. Oxford University Press, pages 104–131.
- Chomsky, N. 2008. On Phases. In *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, edited by R. Freidin, C. Otero, and M. L. Zubizarreta. MIT Press, pages 133–166.

- Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. Oxford University Press.
- Clifton, C. and L. Frazier. 2010. Imperfect Ellipsis: Antecedents beyond Syntax? *Syntax* 13:279–297.
- Constant, N. 2012. English Rise-Fall-Rise: A Study in the Semantics and Pragmatics of Intonation. *Linguistics and Philosophy* 35:407–442.
- den Dikken, M. 2006. Specificational Copular Sentences and Pseudoclefts. In *The Blackwell Companion to Syntax*, edited by M. Everaert and H. van Riemsdijk, volume 4, chapter 61. Blackwell Publishing, pages 292–409.
- den Dikken, M. 2007. Phase extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33:1–41.
- Emonds, J. 1970. Root and Structure-Preserving Transformations. Ph.D. thesis, MIT.
- Ernst, T. 2009. Speaker-Oriented Adverbs. *Natural Language and Linguistic Theory* 27:497–544.
- Farkas, D. and K. Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27:81–118.
- Farkas, D. and F. Roelofsen. To appear. Polar Initiatives and Polarity Particle Responses in an Inquisitive Discourse Model. *Language* to appear.
- Fiengo, R. and R. May. 1994. *Indices and Identity*. MIT Press, Cambridge, MA.
- Fox, D. and D. Pesetsky. 2005. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31:1–45.
- Gallego, Á. 2010. *Phase Theory*, volume 152 of *Linguistik Aktuell/Linguistics Today*. John Benjamins, Amsterdam/New York.
- Gengel, K. 2009. Phases and Ellipsis. *Linguistic Analysis* 35:21–42.
- Goldberg, L. 2005. Verb-Stranding VP Ellipsis: A Cross-Linguistic Study. Ph.D. thesis, McGill University.
- Grant, M., C. Clifton Jr., and L. Frazier. 2012. The role of non-actuality im-

- plicatures in processing elided constituents. *Journal of Memory and Language* 66:326–343.
- Haddican, W. 2004. Sentence Polarity and Word Order in Basque. *The Linguistic Review* 21:87–124.
- Haegeman, L. 2012. *Adverbial Clauses, Main Clause Phenomena and the Composition of the Left Periphery*. Oxford University Press.
- Haegeman, L. 2013. The Syntax of Registers: Diary Subject Omission and the Privilege of the Root. *Lingua* 130:88–110.
- Harley, H. 2002. WCO, ACD, and QR of DPs. *Linguistic Inquiry* 33:659–664.
- Harwood, W. 2012. There are several positions available: English intermediate subject positions. In *Proceedings of ConSOLE XIX*, edited by E. Boone, K. Linke, and M. Schulpen. Leiden University Centre for Linguistics, pages 1–25.
- Harwood, W. 2013a. Being Progressive is Just a Phase. [lingbuzz/001663](http://lingbuzz/001663).
- Harwood, W. 2013b. Rise of the Auxiliaries: A Case for Auxiliary Raising vs. Affix Lowering. [lingbuzz/001755](http://lingbuzz/001755).
- Hestvik, A. 1995. Reflexives and Ellipsis. *Natural Language Semantics* 3:211–237.
- Holmberg, A. 2001. The Syntax of ‘Yes’ and ‘No’ in Finnish. *Studia Linguistica* 55:140–174.
- Holmberg, A. 2012. On the Syntax of ‘Yes’ and ‘No’ in English. [lingBuzz/001334](http://lingbuzz/001334).
- Holmberg, A. 2013. The Syntax of Answers to Polar Questions in English and Swedish. *Lingua* 128:31–50.
- Horn, L. 1989. *A Natural History of Negation*. University of Chicago Press.
- Hornstein, N. 1998. Movement and Chains. *Syntax* 1:99–127.
- Hualde, J. I. and J. Ortiz de Urbina, eds. 2003. *A Grammar of Basque*. Mouton de Gruyter, Berlin.
- Iatridou, S., E. Anagnostopoulou, and R. Izvorski. 2001. Observations about the Form and Meaning of the Perfect. In *Ken Hale: A Life in Language*, edited by

- M. Kenstowicz, chapter 6. MIT Press, pages 189–238.
- Jayaseelan, K. 2001. IP-internal topic and focus phrases. *Studia Linguistica* 55:39–75.
- Johnson, K. 2001. What VP Ellipsis Can Do, and What it Can't, but not Why. In *The Handbook of Contemporary Semantic Theory*, edited by M. Baltin and C. Collins. Blackwell, pages 439–479.
- Johnson, K. 2004. How to be Quiet. In *Proceedings from the 40th Annual Meeting of the Chicago Linguistic Society*, edited by N. Adams, A. Cooper, F. Parrill, and T. Wier. University of Chicago Press, pages 1–20.
- Kay, P. 2002. English Subjectless Tagged Sentences. *Language* 78:453–481.
- Kehler, A. 2002. *Coherence, reference, and the theory of grammar*. Center for the Study of Language and Information, Stanford, CA.
- Kennedy, C. 2003. Ellipsis and Syntactic Representation. In *The Interfaces: Deriving and Interpreting Omitted Structures*, edited by S. Winkler and K. Schwabe. John Benjamins, pages 29–53.
- Kertz, L. 2010. Ellipsis Reconsidered. Ph.D. thesis, UC San Diego.
- Kim, C., G. Kobele, J. Runner, and J. Hale. 2011. The Acceptability Cline in VP Ellipsis. *Syntax* 14:318–354.
- Kim, C. and J. Runner. 2011. Discourse Parallelism and VP Ellipsis. In *UMass Occasional Papers in Linguistics: Ellipsis*. UMass.
- Klima, E. 1964. Negation in English. In *The Structure of Language*, edited by J. Fodor and J. Katz. Prentice Hall, pages 246–323.
- Koopman, H. 1996. The spec-head configuration. *Syntax at Sunset: UCLA Working Papers in Syntax and Semantics* 1(37–65).
- Koopman, H. and A. Szabolcsi. 2000. *Verbal Complexes*, volume 34 of *Current Studies in Linguistics*. MIT Press.
- Kramer, R. and K. Rawlins. 2010. Polarity Particles and Ellipsis: A (Somewhat) Cross-Linguistic Perspective. Presented at UCSC Workshop on Polarity Parti-



- cles.
- Kramer, R. and K. Rawlins. 2011. Polarity particles: an ellipsis account. In *Proceedings of NELS 39*, edited by S. Lima, K. Mullin, and B. Smith. UMass, pages 479–492.
- Kratzer, A. 1996. Severing the external argument from its verb. In *Phrase Structure and the Lexicon*, volume 33 of *Studies in Natural Language and Linguistic Theory*. Dordrecht: Kluwer, pages 109–137.
- Laka, I. 1990. Negation in Syntax: On the Nature of Functional Categories and Projections. Ph.D. thesis, MIT.
- Lasnik, H. and T. Stowell. 1991. Weakest Crossover. *Linguistic Inquiry* 22:687–720.
- Legate, J. A. 2003. Some Interface Properties of the Phase. *Linguistic Inquiry* 34:506–515.
- Lieberman, M. and I. Sag. 1974. Prosodic Form and Discourse Function. In *Proceedings from the 10th Annual Meeting of the Chicago Linguistic Society*. University of Chicago Press, pages 402–415.
- Lipták, A. 2003. The Expression of Sentential Emphasis in Hungarian, *ms.* Leiden University.
- Lobeck, A. 1995. *Ellipsis: Functional heads, licensing, and identification*. Oxford University Press.
- López, L. 1999. VP-Ellipsis in Spanish and English and the features of Aux. *Probus* 11:263–297.
- Mahajan, A. 2005. Slooping in Hindi, *ms.* UCLA.
- Matsuo, A. 2001. AspP-shell Structure in VP-Ellipsis and ACD. In *Proceedings of WCCFL 20*, edited by K. Megerdooomian and L. Bar-el. Cascadilla Proceedings Project, pages 386–399.
- May, R. 1985. *Logical Form*. MIT Press, Cambridge, MA.
- McCloskey, J. 1991. Clause Structure, Ellipsis and Proper Government in Irish.

- Lingua* 85:259–302.
- McCloskey, J. 2005. On Goldberg. Presented at the Workshop on Identity in Ellipsis, University of California, Berkeley.
- Merchant, J. 2001. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford University Press.
- Merchant, J. 2003. Subject-Auxiliary Inversion in Comparatives and PF Output Constraints. In *The Interfaces: Deriving and Interpreting Omitted Structures*, edited by K. Schwabe and S. Winkler. John Benjamins, pages 55–77.
- Merchant, J. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27:661–738.
- Merchant, J. 2010. A taxonomy of elliptical repair. Handout presented at SyntaxFest 2010, Indiana University.
- Merchant, J. 2013a. Ellipsis: A Survey of Analytical Approaches, *ms.* University of Chicago.
- Merchant, J. 2013b. Voice and Ellipsis. *Linguistic Inquiry* 44:77–108.
- Moulton, K. 2008. Small antecedents: Presupposition accommodation in ellipsis resolution, *ms.* UMass.
- Müller, G. 2010. On Deriving CED Effects from the PIC. *Linguistic Inquiry* 41:35–82.
- Orfitelli, R. 2012. Parsimony in passivization: Lexically defining the core characteristics of the *get*-passive, *ms.* UCLA.
- Pope, E. 1972. Questions and Answers in English. Ph.D. thesis, MIT.
- Postma, G. and W. van der Wurff. 2007. How to say ‘no’ and ‘don’t’: Negative imperatives in Romance and Germanic. In *Imperative Clauses in Generative Grammar*, edited by W. van der Wurff. John Benjamins, pages 205–249.
- Potsdam, E. 1997. NegP and subjunctive complements in English. *Linguistic Inquiry* 28:533–541.
- Rezac, M. 2013. Three Consequences of the Pronoun + Ellipsis Analysis of English Tags, *ms.* CNRS-IKER.

- Rizzi, L. 1997. The Fine Structure of the Left Periphery. In *Elements of Grammar*, edited by L. Haegeman. Springer, pages 281–337.
- Rizzi, L. 2001. On the Position Int(errogative) in the Left Periphery of the Clause. In *Current Studies in Italian Syntax: Essays offered to Lorenzo Renzi*, edited by G. Cinque and G. Salvi. North-Holland, pages 287–296.
- Roberts, I. 2010. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. MIT Press, Cambridge, MA.
- Ross, J. R. 1969. Auxiliaries as Main Verbs. In *Studies in Philosophical Linguistics*, edited by W. Todd. Great Expectations Press, pages 77–102.
- Ross, J. R. 1973. A Fake NP Squish. In *New Ways of Analyzing Variation in English*, edited by C.-J. N. Bailey and R. W. Shuy. Georgetown University Press, pages 96–140.
- Sag, I. 1976. Deletion and Logical Form. Ph.D. thesis, MIT.
- Sailor, C. 2009. Tagged for Deletion: A Typological Approach to VP Ellipsis in Tag Questions. Master's thesis, UCLA.
- Sailor, C. 2011. Remarks on Retorts. Presented at NELS 42, University of Toronto.
- Sailor, C. 2012a. Inflection at the Interface, *ms.* UCLA.
- Sailor, C. 2012b. On Embedded Polar Replies. Presented at the Workshop on the Syntax of Answers to Polar Questions, Newcastle University.
- Sailor, C. 2012c. Remarks on Replies: On Emphatic Polarity in English. Presented at the 2012 LSA Annual Meeting, Portland, OR.
- Sailor, C. and B. Ahn. 2010. The Voices in our Heads: The VoiceP in English. Presented at the Workshop on Morphological Voice and its Grammatical Interfaces, University of Vienna.
- Sailor, C. and C. T. Schütze. 2013. Is there repair by ellipsis?, *ms.* UCLA.
- Sailor, C. and G. Thoms. To appear. On the Non-Existence of Non-Constituent Coordination and Non-Constituent Ellipsis. In *Proceedings of WCCFL 31*. Cascadia Proceedings Project.

- Samko, B. 2013. Participle Preposing, *ms.* UCSC.
- SanPietro, S. A., M. Xiang, and J. Merchant. 2012. Accounting for Voice Mismatch in Ellipsis. In *Proceedings of WCCFL 30*, edited by N. Arnett and R. Bennett. Cascadilla Proceedings Project, pages 303–312.
- Sauerland, U. 2005. Don't Interpret Focus: Why a Presuppositional Account of Focus Fails, and How a Presuppositional Account of Givenness Works. In *Proceedings of Sinn und Bedeutung 9*, edited by E. Maier, C. Bary, and J. Huitink. Nijmegen Centre of Semantics, pages 370–384.
- Schütze, C. T. 2004. Synchronic and diachronic microvariation in English 'do'. *Lingua* 114:495–516.
- Silverman, K., M. Beckman, J. Pitrelli, M. Ostendorf, C. Wightman, P. Prince, J. Pierrehumbert, and J. Hirschberg. 1992. TOBI: A Standard for Labeling English Prosody. In *Second International Conference on Spoken Language Processing*.
- Stowell, T. In progress. Parenthetical Adverbials, *ms.* UCLA.
- Tenny, C. 1987. Grammaticalizing aspect and affectedness. Ph.D. thesis, MIT.
- Thoms, G. 2010. 'Verb Floating' and VP-Ellipsis: Towards a Movement Theory of Ellipsis Licensing. *Linguistic Variation Yearbook* 10:252–297.
- Thoms, G. 2012a. Verb Movement and Ellipsis in Scandinavian, *ms.* Ms., University of Edinburgh.
- Thoms, G. 2012b. Yes and No, Merge and Move, Ellipsis and Parallelism, *ms.* Presented at the Workshop on the Syntax of Answers to Polar Questions, Newcastle University.
- Valmala, V. 2009. On the Position of Central Adverbial Clauses. *Anuario del Seminario de Filología Vasca "Julio de Urquijo" (ASJU) / International Journal of Basque Linguistics and Philology* 43:951–970.
- van Craenenbroeck, J. 2010. *The Syntax of Ellipsis: Evidence from Dutch dialects.* Oxford University Press.

- van Craenenbroeck, J. Forthcoming. VP-Ellipsis. In *The Blackwell Companion to Syntax*. Blackwell, second edition.
- van Craenenbroeck, J. and J. Merchant. 2013. Ellipsis Phenomena. In *The Cambridge Handbook of Generative Syntax*, edited by M. den Dikken. Cambridge University Press, pages 701–745.
- Wagner, M. 2006. Givenness and Locality. In *Proceedings of SALT XVI*, edited by M. Gibson and J. Howell. Cornell University, pages 295–312.
- Ward, N. 2006. Non-Lexical Conversational Sounds in American English. *Pragmatics & Cognition* 14:129–182.
- Wood, J. 2008. ‘So’-inversion as Polarity Focus. In *Proceedings of WECOL 38*, edited by M. Grosvald and D. Soares. UC Davis, pages 304–317.
- Yoshida, M. 2010. “Antecedent-Contained” Sluicing. *Linguistic Inquiry* 41:348–356.
- Zanuttini, R. 1997. *Negation and Clausal Architecture*. Oxford University Press.
- Zwicky, A. and G. Pullum. 1983. Cliticization vs. Inflection: English N’T. *Language* 59:502–513.