

ABSTRACT

Among the arguments for successive cyclicity, the one emerging from the Vacuous Movement Hypothesis (VMH) is perhaps the least straightforward, for both empirical and theoretical reasons. Theoretically, Chomsky's (1986) formulation of the VMH ('vacuous movement is not obligatory at S-Structure') is weak and its status in current theorising dubious. Empirically, the VMH predicts that highest-subject *wh*-questions and relative clauses should, with the subject-*wh* *in situ* and SpecCP available as an escape hatch, be more transparent to extraction than other *wh*-questions and relatives; at first glance, this seems to be at best only partially confirmed by the data. This paper argues that a closer examination of the empirical facts reveals that the VMH's consequences in the realm of extraction do in fact materialise, in a way that dictates a strengthening of the VMH into a *prohibition* on vacuous movement, where by 'vacuous movement' I understand movement crossing no phonologically or semantically visible material. As such, the VMH supports an analysis of successive-cyclic A'-extraction as proceeding through SpecCP and it also procures an account of the hitherto mostly elusive distribution of highest-subject *that*-relatives and 'contact relatives' in different varieties of English.

1 Introduction

Does the subject A'-move in English highest-subject *wh*-questions and relative clauses such as (1a) and (1a')?¹

- | | | | |
|-----|------------------------------------|-----|---|
| (1) | <i>wh</i> -questions | | <i>relative clauses</i> |
| a. | who {ate/*did eat} the croquettes? | a'. | the people *(who/that) ate the croquettes |
| b. | what did they eat? | b'. | the croquettes (which/that) they ate |
| c. | why did they eat them? | c'. | the reason (why/that) they ate them |
| d. | when did they eat them? | d'. | the time (when/that) they ate them |
| e. | where did they eat them? | e'. | the place (where/that) they ate them |

While (1b-e) and (1b'-e') must involve A'-fronting in the overt syntax, for (1a,a') it is not obvious that the subject *wh*-phrase has undergone any overt A'-movement. As is well-known, both highest-subject *wh*-questions and highest-subject relative clauses behave differently from other *wh*-questions or relative clauses in English — thus, in highest-subject *wh*-questions there is no *do*-support, and in highest-subject relatives there must be an overt *wh*-operator or alternatively an overt complementiser at the left edge; so-called 'contact relatives' (Jespersen 1961:Vol. III, pp. 81, 132ff.), which are not introduced by a *wh*-element or *that*, are ungrammatical in standard English if the highest subject is relativised. Especially the lack of *do*-support in (1a) seems to receive a straightforward account if the *wh*-constituent of a highest-subject *wh*-question actually remains *in situ*, in SpecTP: the empirical generalisation about the distribution of *do*-support in root *wh*-questions can then be stated quite simply in terms of whether or not SpecCP is occupied — a straightforward Verb Second effect.²

1 By 'highest-subject *wh*-questions/relatives' I am referring to *wh*-questions and relative clauses in which the subject that is the clausemate of the interrogative/relative C-head is questioned/relativised. Note that the properties described in the main text are unique to *highest-subject wh*-questions and relative clauses — they do not manifest themselves when the subject of a lower clause is questioned or relativised, as seen in (i).

(i) a. who does Sue think ate the croquettes? a'. the people (who/that) Sue thinks ate the croquettes

2 Whether this is a desirable way of looking at the distribution of *do*-support (in particular, whether one should, or even could, capitalise in simultaneous filling of the specifier and head of a functional projection, a 'doubly-filled Comp effect') is not at issue in the context at hand, which is intended to be largely pre-theoretical. See e.g. Koopman & Szabolcsi (2000) for discussion of the theoretical status of simultaneous filling of spec and head.

Logically speaking, there are two broad perspectives one can entertain with respect to the application of A'-movement in questions and relative clauses. Either one could argue that *wh*-operators *uniformly* undergo A'-fronting — e.g. to check a strong/EPP feature;³ or one could argue that *wh*-operators must undergo A'-fronting only when they are demonstrably displaced. The latter option still leaves room for two alternative points of view when it comes to the analysis highest-subject *wh*-questions and relatives: if there is an *obligation* to A'-front only in cases in which displacement is not string-vacuous, then in string-vacuous cases there could either be an *option* to A'-front or A'-fronting could be *forbidden* in such contexts. In *Barriers*, Chomsky (1986:49-50) takes the former tack: he formulates his Vacuous Movement Hypothesis (VMH) as in (2a), which makes vacuous movement optional. A strong version of the VMH, as in (2b), would forbid all vacuous movement, and would hence effectively force the subject of highest-subject *wh*-questions and relative clauses in English to stay *in situ* in SpecTP.

- | | | |
|-----|----|---|
| (2) | a. | <i>Vacuous Movement Hypothesis — weak version</i> |
| | | 'vacuous movement is not obligatory at S-Structure' (Chomsky 1986:49-50) ⁴ |
| | b. | <i>Vacuous Movement Hypothesis — strong version</i> |
| | | vacuous movement is prohibited |

Chomsky (1986) adopts (2a), not (2b), because the latter would lead us to expect that '*wh*-island effects will be removed for embedded *wh*-subjects' (p. 48), *quod non* — (3) is very poor.

- (3) *what do you wonder [_{CP} who saw *t*]

He goes on to rationalise his weak formulation of the VMH with an appeal to acquisition and markedness:

- (4) 'The VMH makes intuitive sense. It would mean that UG requires that *wh*-phrases appear in the position of specifier of CP at LF but that the language learner assumes that there is syntactic movement only when there is overt evidence for it. We might suppose that the unmarked case for a language with overt *wh*-movement is that it always takes place at S-Structure, so that nonmovement of subject in English would have a somewhat marked character, accounting for the persistence of weak island effects even with *wh*-subjects, as in [(3)].' (Chomsky 1986:50)

This particular rationale is rather tenuous. On the one hand, it ascribes a central role to the overt evidence that is available to the language learner; on the other, it tries to do away with the guiding role of overt evidence by assuming a general markedness consideration that will insist on overt *wh*-movement even when there is no apparent evidence for it in the linear string. Chomsky's attempt here to have his cake and eat it, too, is reflective of a general uncertainty in the principles-and-parameters literature regarding vacuous movement.

In current theorising (couched in the 'minimalist program'), one might argue that there simply could not be a role to play for anything like the VMH: after all, if all overt displacement (at least as far as phrasal movement is concerned) is driven by the need to check an EPP-property of some functional head, then whenever the F-head in question is present it will force a constituent with matching features into its specifier position to satisfy its EPP-property, completely without regard to the question of whether the movement operation that this constituent is undergoing is crossing any phonologically overt material or not. This conclusion seems particularly acute if phonological features are only supplied at Spell-Out. So (a) if 'EPP' is key, (b) if EPP-satisfaction is possible via movement to specifier positions only, and (c) if we assume 'late insertion', it seems the VMH, whether formulated as in (2a) or as in (2b), must necessarily be false.

3 For relative clauses the text statement would be sufficient; for *wh*-questions, it would (for English) have to be taken to be restricted to *single* questions — after all, English is not a multiple *wh*-fronting language.

4 For Chomsky (1986:50), 'more general considerations, presumably having to do with questions of scope at LF, require that *wh*-phrases move to the specifier position of CP at LF whenever they remain *in situ* at S-Structure'.

In this paper, I will not reject the VMH, however. On the contrary, I will argue that the empirical facts support a *strong* interpretation of the VMH in terms of a *prohibition* on vacuous movement, where by ‘vacuous movement’ I understand movement crossing no phonologically *or semantically* visible material. The version of the VMH that I will argue for in this paper, then, reads as in (5):

- (5) *Vacuous Movement Hypothesis — my version*
movement that does not cross phonologically or semantically visible material is prohibited

I will show that this version of the VMH supports an analysis of successive-cyclic A'–extraction as proceeding through intermediate SpecCP positions (*contra* Rackowski & Richards 2005), and that it also brings forth an account of the distribution of highest-subject *that*-relatives and ‘contact relatives’ in different varieties of English, an issue that has so far remained largely elusive.

2 The Vacuous Movement Hypothesis and extraction: The role of focus

Chomsky (1986:51) derives one of his arguments⁵ for the VMH from Chung & McCloskey’s (1983) observation that at least *some* highest-subject relatives and *wh*-questions do not set up islands for extraction. They base this on sentences of the type in (6) and (7) (taken directly from Chung & McCloskey 1983), and the contrast between such sentences and cases in which a *non-subject* is involved on the lower cycle as well as cases in which a *non-locally* moved subject is involved — cf. (8a) *vs* (8b,c).⁶

- (6) a. that’s the one trick that I’ve known a lot of people who’ve been taken in by
b. isn’t that the song that Paul and Stevie were the only ones who wanted to record?
c. this is a paper that we really need to find someone who understands
(7) a. this is the kind of crisis that you never know who(se spouse) is manipulating for whose benefit
b. what kinds of gifts are there rules about who can give to whom?
c. these are the dialects that we want to find out who speaks to whom
(8) a. that’s one trick that I know a lot of people who’ve been taken in by
b. *that’s one trick that I know a lot of people that the police have taken in with
c. *that’s one trick that I know a lot of people who no one will admit have been taken in by

5 Chomsky’s (1986:50) initial motive for considering and effectively adopting the (weak) VMH in (2a) comes from the examples in (i): ‘If the VMH is correct, we should expect to find that in such examples as [(i)] (due to Luigi Rizzi), the (a)-case should be more acceptable than the (b)-case, with trace indicated only for the relative ... The conclusion seems correct, though the facts are hardly crystal clear’ (p. 50).

- (i) a. he is the man to whom I wonder [_F [who knew [which book to give *t*]]]
b. he is the man to whom I wonder [who John told [which book to give *t*]]

Chomsky then goes on to derive the relative transparency of *whether*-questions from the idea that *whether* is a C–head raising to SpecCP at LF. Raising *whether* from C to SpecCP is problematic, of course; besides, this approach could only capture Ross’s (1967[1986]:19–20) set of judgements, reproduced in (ii), if one treated *why* as a non-moved *wh*-phrase not occupying SpecCP. While there is reason to believe *why*-type *wh*-phrases do indeed differ from other *wh*-phrases in terms of (non)movement, the claim that *why* does not, at S-Structure, occupy the position of fronted *wh*-phrases is less straightforward, esp. for cases where ‘*why*’ is phrasal.

- (ii) a. he told me about a book which I can’t figure out {*why*/*whether*/*when*} I should read
b. which books did he tell you {*why*/*whether*/*when*} he wanted to read?

6 Chung & McCloskey (1983:711) duly note that ‘[t]he judgments of grammaticality ... are of course delicate and marginal’, and add in a footnote that ‘[w]e are not, of course, claiming that all examples of the general type in [(6)/(7)] are fully acceptable’.

On the assumption that the *wh*-operator introducing the relative or embedded *wh*-question is (or at least can be) *in situ* in SpecTP in (6), (7) and (8a) (but not, of course, in (8b,c)), the SpecCP escape hatch is available in principle for extraction.⁷ In the *Barriers* theory, that then makes (7) straightforwardly grammatical. For (6), the inherent barrierhood of the relative clause and the inherited barrierhood of the relativised noun phrase should continue to deliver an ill-formed output even with a stop-over in SpecCP made *en route* (something that Chomsky 1986 does not address), but this problem is easily sidestepped in current phase-based locality theory.

But empirically, ‘the persistence of weak island effects even with *wh*-subjects’ (Chomsky 1986:50) in *single wh*-questions such as (3), repeated below, suggests that the fact that Chung & McCloskey’s examples in (7) all involve *multiple* embedded *wh*-questions is not innocuous. This is confirmed by the fact that the single-question examples in (7') are significantly worse than their multiple-question counterparts in (7).⁸

- (3) *what do you wonder [_{CP} who saw *t*]
(7') a. *this is the kind of crisis that you never know who(se spouse) is manipulating for their own benefit
b. *what kinds of gifts are there rules about who can give to politicians?
c. *these are the dialects that we want to find out who speaks to their neighbours

2.1 The single/multiple question distinction and the role of focus

Why should the single *vs* multiple *wh*-question distinction make such a difference when it comes to extraction from highest-subject *wh*-questions? An answer presents itself based on É. Kiss’s (1993) analysis of the ambiguity of English multiple *wh*-questions such as (9a). É. Kiss argues that such questions support two LF representations (cf. (9b,c)), with the higher *wh* represented as a universal QP and the lower one as the focused constituent of the question. This analysis is supported for Hungarian (cf. (10)), where there are two distinct counterparts to English (9a), both having both *wh*’s fronted, but in different surface orders: the higher of the fronted *wh*’s syntactically and semantically functions as a universal QP, sitting in SpecDistP (cf. Beghelli & Stowell 1997); the lower one is interpreted as the focus, occupying SpecFocP. This is depicted in (10a,b).

- (9) a. who bought what?
b. $\forall x, x$ a person, WH y, y a thing, x bought y
c. $\forall y, y$ a thing, WH x, x a person, x bought y

7 Chung & McCloskey (1983) themselves actually present their data as support for a GPSG account of ‘extraction’ phenomena in terms of ‘slash categories’. For a filler-gap dependency, GPSG exploits what one might call an ‘accounting mechanism’ to keep track of the fact that the filler for a particular position down the tree has not been found yet and we are still ‘looking for it’ — the ‘slash category’ (cf. (ic)). In combination with a(n English-specific) constraint (originally proposed by Maling & Zaenen 1982) to the effect that a slash category *S* with more than one element to the right of the slash is ill-formed (cf. (ii), which can of course be readily updated in a *Barriers*-style phrase structure; Koster’s 1986:168–70 account of (some of) the Chung & McCloskey 1983 facts bases itself on a ‘uniqueness condition on domain definition’ that is highly reminiscent of (ii)), this straightforwardly rules out *Wh*-Island Condition violations as well as CNPC cases involving relative clauses. The operative idea in Chung & McCloskey’s (1983) analysis of the facts in (6) and (7) is Gazdar’s (1981:166) claim that local subject extraction does *not* involve a slash category: instead of *S/α*, what we have is a simple VP (cf. (iii)). (A more radical version of this proposal is Grimshaw’s (1997) idea (embracing the strong VMH) that local subject *wh*-questions/relatives are in fact VPs, with the subject in SpecVP.)

- (i) a. who can you trust?
b. [_{CP} who, [_C can [_{IP} you [_{VP} trust [_{NP} *t*]]]]]
c. [_{CP} who [_{C/NP} can [_{IP/NP} you [_{VP/NP} trust NP/NP]]]]]
(ii) **S/α*, β
(iii) a. [_{NP} [_{NP} the spy] [_R [_{NP} who] [_{VP} came in from the cold]]]
b. [_Q [_{NP} who] [_{VP} wants another piece of pizza]]

8 I thank Jim McCloskey and Jason Merchant for their help with these examples. See also Trotta (2000:74–77) for an earlier claim to the effect that the fact that the examples in (7) involve multiple *wh*-questions is significant. Trotta also appeals to the VMH.

- (10) a. ki MIT vett? (Hungarian)
 who what-ACC bought
 ‘for every person in the discourse, tell me what it was that (s)he bought’
 [DistP ki [Dist ∅ [FocP mit [Foc vett ...]]]
 b. mit KI vett?
 what-ACC who bought
 ‘for every item in the discourse, tell me who it was that bought it’
 [DistP mit [Dist ∅ [FocP ki [Foc vett ...]]]

As Den Dikken & Giannakidou (2002:53) point out (see also Lipták 2001:115⁶), É. Kiss’s (1993) analysis of the Hungarian examples in (10) (updated to include Beghelli & Stowell’s DistP) is strongly supported by the fact that the initial *wh*-phrase in Hungarian double *wh*-questions of this type must be specific/D-linked, hence resists being ‘radically non-D-linked’ (Pesetsky 1987). The primed examples in (11) fail for the same reason that their English paraphrases given below them fail as well.

- (11) a. ki MI A FENÉT vett? (Hungarian)
 who what-the-hell-ACC bought
 ‘for every person in the discourse, tell me what the hell it was that (s)he bought’
 [DistP ki [Dist ∅ [FocP mi a fenét [Foc vett ...]]]
 a’. *ki a fene MIT vett?
 who-the-hell what-ACC bought
 ‘for every-the-hell person in the discourse, tell me what it was that (s)he bought’
 *[DistP ki a fene [Dist ∅ [FocP mit [Foc vett ...]]]
 b. mit KI A FENE vett?
 what-ACC who-the-hell bought
 ‘for every item in the discourse, tell me who the hell it was that bought it’
 [DistP mit [Dist ∅ [FocP ki a fene [Foc vett ...]]]
 b’. *mi a fenét KI vett?
 what-the-hell-ACC who bought
 ‘for every-the-hell item in the discourse, tell me who it was that bought it’
 *[DistP mi a fenét [Dist ∅ [FocP ki [Foc vett ...]]]

This analysis of multiple *wh*-questions presents the beginnings of an answer to the question of why the single vs multiple *wh*-question distinction makes such a big difference when it comes to extraction from highest-subject *wh*-questions. Before I can lay out the full answer, I need to put in place one further piece to the puzzle — the assumption in (12):

- (12) when a subject-*wh* gets a *focus* interpretation, it must *not* be in SpecTP in overt syntax

Given that SpecTP is the quintessential A-position for topics, (12) is part of a broader generalisation that bars foci from occupying topic positions. In keeping with (12), a focused subject-*wh* must be *ex situ*. In English embedded questions, *ex situ wh*-phrases raise to SpecCP (as shown in Den Dikken & Giannakidou 2002; see also Den Dikken 2003a) — this is evident from at least two mutually reinforcing pieces of empirical evidence: (i) the placement of *ex situ wh*-phrases *vis-à-vis* fronted topics, shown in (13) (cf. Pesetsky 1989), and (ii) the distribution of *wh-the-hell*, partially illustrated in (14).

⁶ Though the dates of publication of these two works do not reveal this, Den Dikken & Giannakidou (2002) is historically prior to Lipták (2001), and is, to my knowledge, the first context in the literature in which these Hungarian data are presented. In Den Dikken & Giannakidou (2002) (where no reference is made to É. Kiss’s analysis of multiple *wh*-questions), the position of the initial *wh*-phrase in (10) and (11) is identified as SpecTopP. In Den Dikken (2003a), the DistP structure is used in lieu of TopP.

- (13) a. ?a book like this, why should I buy
 b. ?Bill doesn’t know why a book like this, he should buy
 (14) a. who the hell would buy a book like this?
 a’. [C_Q [FocP who the hell would buy a book like this]]
 b. I *(don’t) know who the hell would buy a book like this
 b’. ... *(Neg) ... [C_P who the hell [C_Q [would buy a book like this]]

The relative order of topics and *ex situ wh*-phrases is different in English root and embedded *wh*-questions, as (13) shows. Such word-order variation could in principle be accounted for in either of two ways: by keeping the position of *wh*-phrases constant and varying the placement of topics depending on whether we are dealing with a root or embedded clause, or by keeping the position of topics constant and varying the placement of *wh*-phrases. It turns out that the latter approach gives us an interesting purchase on the facts in (14). If *wh*-phrases are in a relatively low structural position in root *wh*-questions, one that is crucially below the Q-operator in C, whereas in embedded clauses *ex situ wh*-phrases must raise to SpecCP, we expect to find an asymmetry between root and embedded questions when it comes to the (in)ability on the part of the Q-operator in C to license *wh-the-hell*, which Den Dikken & Giannakidou (2002) argue is a polarity item. In particular, we expect that since the nonveridical Q-operator in (14a’) c-commands *wh-the-hell*, it can license the polarity item, while in (14b’) the lack of c-command between C_Q and *wh-the-hell* should force the latter to depend for its licensing on a nonveridical element in the matrix clause. This turns out to be exactly right: (14b) succeeds if the matrix clause harbours a nonveridical operator that can license *wh-the-hell* (here, the matrix negation), but not otherwise. If *ex situ wh*-phrases could raise to a position below C_Q in English embedded questions, the obligatoriness of the matrix negation in (14b) would be difficult to account for; conversely, if *ex situ wh*-phrases raised to SpecCP in matrix questions, the grammaticality of (14a) would be a major mystery. The facts in (14) fall out straightforwardly if *ex situ wh*-phrases target SpecCP in embedded questions but raise to a position below C, which (14a’) identifies as SpecFocP, in root questions in English. With that conclusion in place, the word-order facts in (13) immediately fall into place as well, with topics systematically targeting a position outside FocP but below CP.

I take this to establish that the landing-site of *ex situ wh*-phrases in English embedded clauses is SpecCP. Going back now to (12), the ban on focalised *wh*-subjects staying in SpecTP, we are led to the conclusion that the subject-*wh* in English embedded clauses has to raise to SpecCP whenever it is a focus, whereas it will be allowed (and, on a strong interpretation of the VMH *à la* (2b), possibly forced) to be in SpecTP when it does not get a focus interpretation. The key insight to take away from this is that the SpecCP position of English embedded *wh*-questions is of necessity occupied by a fronted *wh*-constituent whenever the *wh*-constituent is a non-subject *or* is a *focused* subject; it is only in embedded highest-subject *wh*-questions in which the subject-*wh* is NOT read as a focus that the embedded SpecCP position is not necessarily occupied by the *wh*-phrase in the left periphery of the subclause. This is what holds the key to an understanding of the facts canvassed in the foregoing discussion, repeated below.

- (6) a. that’s the one trick that I’ve known a lot of people who’ve been taken in by
 b. isn’t that the song that Paul and Stevie were the only ones who wanted to record?
 c. this is a paper that we really need to find someone who understands
 (7) a. this is the kind of crisis that you never know who(se spouse) is manipulating for whose benefit
 b. what kinds of gifts are there rules about who can give to whom?
 c. these are the dialects that we want to find out who speaks to whom
 (7’) a. ?*this is the kind of crisis that you never know who(se spouse) is manipulating for their own benefit
 b. ?*what kinds of gifts are there rules about who can give to politicians?
 c. ?*these are the dialects that we want to find out who speaks to their neighbours

Let me spell out in detail how É. Kiss's (1993) analysis of multiple *wh*-questions, in conjunction with the hypothesis in (12) and Den Dikken & Giannakidou's (2002) conclusion that *ex situ wh*-phrases are in SpecCP in English embedded *wh*-questions, explains the facts in (6), (7) and (7'). Consider first the relative clause cases in (6), which are the easiest ones to account for. These examples are grammatical because the *wh*-operator of relative clauses is never a focus. This is immediately apparent from the information-structural function of the *wh*-operator: it does not provide or request new information. It is also clear from the word-order facts of languages such as Hungarian, which allow us to diagnose focus movement on the basis of an inspection of the linear string: whereas a focused constituent that is raised into SpecFocP forces the finite verb to surface in Foc⁰, to its immediate right, constituents that do not raise to SpecFocP do not bring about such a Verb Second effect.¹⁰ The contrast in (15) shows unequivocally that the fronted *wh*-operator in relative clauses is not in SpecFocP.¹¹

- (15) a. ki <(*be-)jött (be) a szobába? (Hungarian)
 who PV came PV the room-into
 'who came into the room?'
 b. az ember aki (be-)jött (*be) a szobába
 the man A-who PV came PV the room-into
 'the man who came into the room'

Now that we have established that the *wh*-operator in relative clauses is not a focus, the grammaticality of the examples in (6) follows straightforwardly: since nothing tells us that the *wh*-operator is *ex situ* (in particular, word order does not, nor does (12), which is moot in relatives), by the VMH we will leave the *wh*-operator *in situ* in SpecTP; as a result, SpecCP is available as an intermediate landing-site for extraction from the highest-subject relative clause.¹²

In *single* embedded subject *wh*-questions, the subject-*wh* is necessarily a focus (recall Hungarian (15a): the preverb *be* has to appear to the right of the finite verb), hence *ex situ*, in SpecCP, in the case of English. It thereby blocks extraction — hence the ungrammaticality of the examples in (7'). But *multiple wh*-questions such as (7) can be assigned an interpretation *à la* (9b) (repeated below), with the subject-*wh* read as a universal QP and the non-subject-*wh* functioning as the focus of the question. On such a reading of a highest-subject multiple question, the subject-*wh* is *in situ* in SpecTP in the overt syntax (raising covertly to SpecDistP; I am assuming that 'QR' is covert movement in English), and since the focused non-subject-*wh* does not occupy SpecCP either (as word order shows), SpecCP is thus available as an intermediate landing-site for *wh*-extraction.

- (9) a. who bought what?
 b. $\forall x, x$ a person, WH y, y a thing, x bought y
 c. $\forall y, y$ a thing, WH x, x a person, x bought y

On the alternative reading in which the subject-*wh* is read as the focus of the *wh*-question (cf. (9c)), by contrast, the focused subject-*wh* is forced (by hypothesis (12)) to vacate SpecTP and raise to SpecCP in overt syntax. By so raising, it prevents A'-extraction across it just as in *single wh*-questions.

10 As Horvath (2000, 2005) shows, movement to the position identified in this paper as SpecFocP is undertaken only by a subset of focused constituents in Hungarian. In particular, only focused constituents that receive an 'exhaustive identification' reading raise to this position. Wedgwood (2005:sect. 4.6) shows, however, that putting a constituent in the immediately preverbal focus position does not appear to be sufficient to procure 'exhaustive identification'.

11 In the glosses to (15), 'PV' stands for 'preverb', a particle-like element that, in neutral (i.e., widest-scope, full-clausal focus) sentences, immediately precedes the verb but which in focus-fronting contexts must follow the finite verb.

12 As is apparent from the main-text discussion, I am not assuming a Vergnaud/Kayne-style raising analysis of relative clauses. See (the text to) Den Dikken (2006:309, n. 100) and references there for relevant discussion.

This analysis of extraction from *single vs* multiple highest-subject *wh*-questions makes an interesting prediction — one that is not made by any previous accounts, to my knowledge. Since, on my analysis, the SpecCP position of highest-subject multiple *wh*-questions is available as an escape hatch ONLY if the subject-*wh* is read as a distributive QP, it is predicted that extraction from multiple *wh*-questions disambiguates them towards an interpretation in which the subject-*wh* takes *wide scope* and distributes over the *in-situ wh*, being forced into a *specific* reading (cf. É. Kiss's 1993 Specificity Filter). This prediction is confirmed by the fact that the subject-*wh* must necessarily be specific in multiple *wh*-questions involving extraction, as the contrast in (16) shows.¹³

- (16) the dialects that we want to find out {*who*/**how many people*} speak(s) where

The combination of É. Kiss's (1993) analysis of multiple questions and the hypothesis in (12) thus derives the facts in (6), (7) and (7'), and correctly predicts (16). In highest-subject relative clauses and also in *multiple wh*-questions whose subject-*wh* is not a focus, the subject-*wh* is *in situ* in SpecTP, in agreement with the Vacuous Movement Hypothesis; in *single wh*-questions as well as in *multiple wh*-questions whose subject-*wh* is focused, even the highest subject is forced to be *ex situ* due to (12), which bans focused subjects from SpecTP, which is quintessentially a topic position hence by nature incompatible with foci occupying it. For *wh*-phrases in English embedded questions, 'being *ex situ*' means (as Den Dikken & Giannakidou 2002, Den Dikken 2003a show) 'being in SpecCP' — and with SpecCP occupied in embedded questions introduced by an *ex situ wh*-phrase, this means that this position cannot be used as an escape hatch for A'-extraction. The extraction facts then follow, on the traditional assumption that A'-extraction from CP must proceed successive-cyclically, via an intermediate touch-down in SpecCP. The representations in (17) sum up the essence of the account outlined so far.¹⁴

- (17) a. $*[_{CP} wh_k [_{C'} C \dots [_{\nu P} \overline{wh}_k [_{\nu P} \nu \dots [_{CP} who_i [_{C'} C \dots [_{TP} \overline{wh}_i [_{TP} T [_{\nu P} \overline{wh}_k [_{\nu P} \nu \dots \overline{wh}_k \dots]]]]]]]]]]$
 b. $[_{CP} wh_k [_{C'} C \dots [_{\nu P} \overline{wh}_k [_{\nu P} \nu \dots [_{CP} \overline{wh}_k [_{C'} C \dots [_{TP} who_i [_{TP} T [_{\nu P} \overline{wh}_k [_{\nu P} \nu \dots \overline{wh}_k \dots]]]]]]]]]]$

2.2 On successive cyclicity

Though this is not immediately apparent from the schematic representation in (17b), there is a difference between the derivation (17b) depicts and a 'garden variety' case of successive-cyclic extraction out of an embedded CP, instantiated by (17c), the representation of a sentence like *what do you think that Mary read?*

- (17) c. $[_{CP} wh_k [_{C'} C \dots [_{\nu P} \overline{wh}_k [_{\nu P} \nu \dots [_{CP} \overline{wh}_k [_{C'} C \dots [_{TP} SU [_{TP} T [_{\nu P} \overline{wh}_k [_{\nu P} \nu \dots \overline{wh}_k \dots]]]]]]]]]]$

Whereas in (17c), the head of the embedded CP is not interrogative, hence not endowed with the feature [+WH], in (17b) we are dealing with a question-CP with a [+WH] C-head. Successive-cyclic *wh*-extraction from a [+WH] CP, with the single *wh*-phrase satisfying the [+WH] features of both the matrix and the embedded C-heads, is normally entirely impossible: **what do you wonder Mary read?* What, then, makes (17b) converge?

13 Thanks to Jason Merchant for his help with this and similar examples. He finds the version of (16) with *how many people* 'much worse' than its variant with *who*. Testing the prediction on the basis of the distribution of *wh-the-hell* (aggressively non-D-linked/non-specific; apparently never *in situ*) is not straightforward in English given the fact that the use of *wh-the-hell* in English multiple *wh*-questions presents additional complications: see Den Dikken & Giannakidou (2002:35ff.) for discussion.

14 (17) assumes, in line with standard assumptions in the current literature, that νP is a phase, and that hence escaping from νP under A'-extraction always proceeds via intermediate adjunction to νP .

I would like to argue that (17b) is grammatical thanks to the fact that, *prior* to the embedded C ever getting engaged in a relationship with wh_k , it first establishes a *wh*-Agree relationship with the subject-*wh*, which checks C's uninterpretable [+WH] feature — though not its EPP-property. C's EPP-property is subsequently checked, separately, by wh_k . It would be impossible for wh_k to check *both* C's [+WH] feature *and* its EPP-property, and then to move on into the matrix SpecCP position: I am assuming (in line with what I argued in Den Dikken 2003a) that the morphological [+WH] feature of *wh*-expressions is uninterpretable, hence erased upon the completion of the first phase within which this feature is checked, which would be the embedded CP if wh_k checked C's [+WH] feature there. The ungrammaticality of **what do you wonder Mary read?* then follows; and the grammaticality of the derivation in (17b) does, too: here the checking of C's [+WH] feature is taken care of by *who*, independently of wh_k . Naturally, (17a) is grammatical as well: here *who* checks both C's [+WH] feature and its EPP-property. But since (17a) delivers a focus reading for *who*, this derivation is unavailable in multiple *wh*-questions in which *who* is interpreted as a wide-scope universal quantifier (cf. (9b)). For the (9b) reading of the embedded multiple question, we need (17b), which the analysis correctly declares grammatical *without* at the same time allowing undesired successive-cyclic extraction from embedded questions: the analysis makes the desired cut between successive-cyclic extraction from embedded questions that do and ones that do not have a *wh*-constituent in them other than the extractee.

The idea that C in (17b) *wh*-Agrees with the *in situ* subject-*wh* also answers another question raised by (17b): the question of how wh_k can be raised to the embedded SpecCP across *who*, in keeping with (18).¹⁵

- (18) *Closest* (Rackowski & Richards 2005:579)
 a goal α is the closest one to a given probe if there is no distinct goal β such that for some X (X a head or maximal projection), X c-commands α but does not c-command β

By (18), we expect C in (17b) to only be able to Agree with who_i — after all, who_i is closer to C than wh_k at the point of Merge of C, with the occupant of SpecTP asymmetrically c-commanding the occupant of the vP-adjoined position.¹⁶ It would certainly be illegitimate, from the point of view of *Closest*, for C not to Agree with *who* at all and then to attract wh_k across it. Of course, if C checked *all* of its features against *who* (as in (17a)), that would not be able to result in attraction of wh_k up to C either. But there is a middle way: C can check a subset of its features against *who*, and then check the remaining feature(s) against wh_k . By Richards' (1998) Principle of Minimal Compliance, the establishment of an Agree relationship between C and *who* will allow C to ignore *who* for the remainder of the derivation — concretely, once C has *wh*-Agreed with *who*, it is free to establish an Agree relationship with wh_k that will result in the erasure of its EPP-property.

So C's *wh*-Agree relationship with *who* in SpecTP is what allows wh_k to raise past it into SpecCP, as in (17b), which is thereby legitimate. In (17a), by contrast, C checks both its [+WH] feature and its EPP-property against *who*, making it impossible for C to attract wh_k , which must hence try to make its way out of the embedded CP in one fell swoop, under the force of the matrix *v*'s attractions. This fails, for familiar reasons: *who* is unattractable, un-Agree-able, 'defective' intervener. For essentially the same reasons, a derivation very much like (17b) but featuring fell-swoop movement of wh_k from CP without a touch-down in SpecCP, depicted in (17d), crashes as well — once again because *who* is a harmful intervener: while the

15 The analysis at the same time raises non-trivial questions about Superiority which I am not in a position to address in this paper. The question of whether Superiority can be accounted for in terms of *Closest* is not a straightforward one: as is well-known, Superiority is not absolute; 'inferior' *wh*'s can, even in languages otherwise observing Superiority, be raised past 'superior' ones under certain circumstances (D-linking being paramount among these; cf. Pesetsky 1987). I suspect that this D-linking effect will ultimately be translatable into the analysis I am developing here, but I do not at this time have the appropriate translation in place.

16 I assume strict c-command here, not m-command. An m-command approach would completely forfeit a structural account of subject/non-subject Superiority (since, by m-command, the subject-*wh* in SpecTP and the vP-adjoined non-subject-*wh* will be in a mutual command relationship: for both, the first maximal projection dominating them is TP), which may be too significant a loss (though assessing this is not straightforward given that the status of Superiority is not crystal clear in the approach that I am outlining in this paper; recall fn. 15).

embedded C can *wh*-Agree with *who* and then 'neutralise' it for the purposes of attraction of wh_k , the matrix *v* cannot accomplish something similar because the embedded CP sets up a phase boundary in between *v* and *who* across which no Agree relationships are establishable. It does not matter, therefore, whether *who* is in SpecTP or SpecCP as far as attraction of wh_k *directly* by the matrix *v* is concerned: the derivation crashes either way.

- (17) d. *_{[CP wh_k [_C C ... [_{vP} \overline{wh}_k [_{vP} *v* ... [_{CP} [_C C ... [_{TP} *who*_i [_{TP} T [_{vP} \overline{wh}_k [_{vP} *v* ... \overline{wh}_k ...]]]]]]]]]]]}

The only derivation that converges is the one in (17b) — a successive-cyclic derivation involving a touch-down in the embedded SpecCP made possible by the fact that the highest-subject *wh*-phrase in the embedded question has remained *in situ*, in SpecTP.

The theory that I am advocating for English, therefore, is one in which, in CP-complementation structures, in order for *v* to successfully attract a subconstituent of the embedded CP up to its edge, that constituent must raise to SpecCP prior to the establishment of an Agree relationship between it and the matrix *v*. While for 'simple' cases of successive-cyclic extraction from CP there may be alternative ways of getting the facts to fall out, for the complex examples addressed in this section, there really seems to be no escape from touching down in SpecCP on one's way out of CP.

Rackowski & Richards (2005) argue (on the basis of a detailed investigation of Tagalog) that long A'-extraction from a complement-CP, while still successive cyclic in key respects, crucially eschews an intermediate touch-down in SpecCP, and that this is because matrix *v* Agrees morphologically with the embedded CP: the Agree relationship between *v* and CP 'lifts' the phasehood of CP (by the Principle of Minimal Compliance) and thus allows matrix *v* to attract a *wh*-constituent on the edge of the embedded vP up to its edge in one fell swoop, much as in (17d). This proposal could only be compatible with the facts discussed in this section if the subject-*wh* in SpecTP in (17d) somehow did not interfere with the Agree relationship that matrix *v* wants to establish with the copy of wh_k on the edge of the embedded vP. While Rackowski & Richards (2005:584) do indeed independently stipulate (in order to account for the ban on subextraction from subjects) that 'subjects, like adjuncts, never enter into an Agree relation with *v*', it would be straightforwardly false to say that *v* can never Agree with a [+WH] nominative subject — after all, *who do you think left?* is grammatical. In the case at hand (i.e., (17d)), therefore, we cannot simply ignore the subject-*wh* in the embedded SpecTP. And if we cannot ignore *who*, it will obstruct an Agree relationship between matrix *v* and the copy of wh_k on the edge of the embedded vP in exactly the same way that *who* in the embedded SpecCP in (3), above, sets up an intervention effect for attraction of *what*.¹⁷

It seems, therefore, that Rackowski & Richards' (2005) approach does not ultimately accommodate our data — it cannot make the desired cut between *in situ* and *ex situ* subject-*wh*'s because, in their theory, even *in situ* subject-*wh*'s should obstruct *wh*-extraction across them. More generally, it is clear from *wh*-extraction in languages such as Hungarian (discussed in Den Dikken 2005) that one cannot take fell-swoop extraction from CP to be the one and only option made available by UG: successive-cyclic movement via a stopover in SpecCP must exist; it seems that English uses it for its long-distance *wh*-extraction derivations.

2.3 The persistence of 'ECP effects' under adjunct extraction

One remaining puzzle is the persistence of 'ECP effects' under adjunct extraction, even in contexts of the type in (7) — cf. the contrast between (7c), repeated below, and (19).

- (7c) these are the dialects that we want to find out who speaks to whom
 (19) *these are the occasions on which we want to find out [who speaks these dialects to whom]

17 See Rackowski & Richards (2005:594) for discussion of *wh*-island effects from the point of view of their theory. They predict categorical opacity for *wh*-questions, which does not seem descriptively adequate (recall Ross 1967[1986], and fn. 5, above).

Though I am not in a position at this time to develop this suggestion in detail, it seems to me that we can gain a perspective on this question if we assume that adjuncts (or at least, the adjuncts in question) originate in vP-adjoined positions *below* the Merge site of the subject (see (20a)), and are prevented from exiting from vP via intermediate adjunction to vP (as in (20b)) on pain of a violation of the ban on adjoining to the same projection twice.

- (20) a. *... [_{CP} *wh*_k [_C C ... [_{TP} *who*_i [_T T [_{vP} *wh*_o_i [_{vP} *wh*_k [_{vP} v ...]]]]]]]]]
 b. *... [_{CP} *wh*_k [_C C ... [_{TP} *who*_i [_T T [_{vP} *wh*_k [_{vP} *wh*_o_i [_{vP} *wh*_k [_{vP} v ...]]]]]]]]]]]

Thus, the *wh*-adjunct in (20a) is c-commanded both by the physical subject in SpecTP and by its lower copy in the vP-internal subject position. This means that after C has established a *wh*-Agree relationship with *who* in SpecTP, the next *wh*-element that it will ‘see’ lower down in the tree is the subject’s own lower copy, *not* the *wh*-adjunct below the subject’s base position.¹⁸ Due to the particular phrase-structural configuration that adjuncts and subjects are in, therefore, in (19) C is in effect forced to do all of its Agree-ing with the subject-*wh* alone, and cannot establish an Agree relationship with the adjunct.

Note that this suggestion will not make vP-adjoined *wh*-adjuncts radically immobile: whenever the subject is not a *wh*-constituent, C can *wh*-Agree with the adjunct without difficulty, the subject not interfering with this relationship this time: the adjunct originates on the edge of the vP phase, and is not separated from C by the full chain of the *wh*-subject. It is the precisely the fact that the full chain of the *wh*-subject separates the *wh*-adjunct from C that, I would suggest, is preventing the adjunct from being attracted up to C — even when the *wh*-subject stays *in situ*, in SpecTP.

3 Forced evacuation of SpecTP, and the nature of ‘vacuous movement’

The analysis of the ‘lifting’ of *wh*-island effects in highest-subject relatives and multiple questions that I presented in section 2 assumes that movement of the subject-*wh* from SpecTP to the local SpecCP is forced when it is a focus — (12), repeated below, is a key ingredient of the account.

- (12) when a subject-*wh* gets a *focus* interpretation, it must *not* be in SpecTP in overt syntax

Forced evacuation of SpecTP due to focality raises two pressing questions that need to be addressed:

- Q1 how can (12) be implemented in a formal theory of movement triggers?
 Q2 what does (12) entail for the strong interpretation of the Vacuous Movement Hypothesis?

3.1 What drives foci out of SpecTP

The statement in (12) is an ‘anti-focus’ condition that is in some sense the opposite of what we find at work in scrambling-type phenomena in the VP-domain: constituents that themselves resist focus are not supposed to stay in the VP, and must scramble out (or, in clitic-doubling languages, undergo clitic doubling) instead (see Reinhart 1995, Zubizarreta 1998). In our case, we find that constituents that are themselves focused are not supposed to stay in SpecTP, which is quintessentially a topic or ‘anti-focus’ position. In both cases, forced evacuation seems to be a function of information-structural properties of the constituents involved.

18 A technical question arises here regarding locality and ‘equidistance’ — one would expect that the subject’s lower copy and the base position of the *wh*-adjunct should be equidistant, hence that there should be no blockage to C attracting the adjunct across the lower copy of the *wh*-subject. In the particular case under discussion, however, extraction proceeds across the entire chain of the subject, whose upper copy, in SpecTP, is not equidistant from *wh*_k. I suggest that because the upper copy of the *wh*-subject is not in the same minimal domain as the *wh*-adjunct, this ultimately prevents C from Agree-ing with the *wh*-adjunct across the full chain of the *wh*-subject. Technical details remain to be worked out — the text suggestion is tentative.

It is difficult to implement (12) in a feature-driven movement system such as the one espoused in the original minimalist program, where a positive morphological property (a ‘strong feature’) of the probe is held responsible for overt displacement. English does not necessarily displace its foci in overt syntax; and even in languages such as Hungarian, which are famous for their overt-syntactic movement of focused constituents, it has become clear thanks to work by Kenesei (1986), Horvath (2000, 2005), É. Kiss (1998), Szendrői (2003), and Wedgwood (2005), among others, that being focused *per se* is not what drives constituents up to the designated focus position, SpecFocP, in syntax: *in situ* foci are grammatical even in Hungarian; only foci that have the semantic property of ‘exhaustive quantification’ are compelled to be *ex situ*.¹⁹ One might conceivably hold an ‘exhaustivity feature’ responsible for displacement of foci in Hungarian (as in Horvath’s 2000, 2005 work); but plainly, such a feature would not be within the set of allowable triggers for movement in the minimalist program as originally defined, where only morphological features are possible triggers. Of course, in present-day minimalism the link between overt displacement and morphological features has been substantially weakened, perhaps even rendered null and void, with the advent of ‘EPP-properties/features’ (also ‘edge features’ in Chomsky 2004) as the sole triggers of movement. The relationship between ‘EPP-properties/features’ and morphological properties of functional heads is at best obscure, and at worst non-existent (cf. esp. the liberal deployment of ‘edge features’ on phase-heads regardless of morphological content).

Van Riemsdijk (1995), Platzack (1996), Koenenman (2000), Zwart (2004), Franks (in press), and Van Craenenbroeck (2006), among others, have experimented in recent years with an alternative outlook on triggering overt displacement, phrased not in terms of featural properties of a probe but instead in terms of properties of the moved constituent itself (closer in spirit to the ‘Greed’ approach of early minimalism). Platzack (1996), in particular, suggests that there exists an operation called ‘Repel F’, which ensures that bearers of a particular feature F be ‘expelled’ from a certain structural domain. For the case of scrambling, then, what one would say, along these lines, is that constituents with the feature [+topic] (or perhaps rather [–focus]; cf. Kallulli 2000 for discussion related to clitic doubling) should be expelled from the VP, which is the typical focus domain. And for (12), we would say that constituents bearing the feature [+focus] should be expelled from the SpecTP position, which is quintessentially a topic position.

Statements of this sort raise the important question of the role that information-structural properties can play in syntax — i.e., how the structure of the discourse can influence the structure of the individual sentences which the discourse is made up of. Assigning individual sentence constituents features such as [+topic], [–focus] or [+focus] involves ‘prejudging’ their role in the discourse at a point at which discourse analysis is in fact impossible. Discourse analysis must arguably follow the syntactic analysis of the individual sentences that constitute the discourse.²⁰ But even if we allowed a cyclic intertwining of syntactic and discourse analysis,²¹ we face the question of how to encode discourse function.

19 Recall from fn. 10, however, that the converse of the text statement following the semi-colon is not true: Wedgwood (2005:sect. 4.6) shows that constituents placed in the immediately preverbal focus position do not automatically get an ‘exhaustive identification’ interpretation.

20 Consider, for instance, the fact that anaphoric dependencies are often sensitive to grammatical function (subject, object, etc.), not just within sentences but across sentences as well — a case in point being the determination of the antecedents of the anaphoric demonstratives and pronouns in the Dutch examples in (i), where (ia) involves a complex sentence and (ib) a sequence of two sentences in discourse: the anaphoric restrictions on *deze* ‘this (one)’ and *hem* ‘him’ in the two examples are the same, and apparently sensitive to the grammatical function of the antecedent (*deze*, in particular, resisting coreference with the subject, *Jan*).

- | | | | | | | | | | | | | |
|-----|----|------------------|------|-------------------|-----|--------|------------|----------------------------------|----------------------------------|---------------------------------|-------|-------|
| (i) | a. | Jan _i | gaf | Piet _j | een | schop. | waarna | <i>deze</i> _{j<i>1</i>} | <i>hem</i> _{j<i>2</i>} | een | dreun | gaf |
| | | Jan | gave | Piet | a | kick | whereafter | this | him | a | slap | gave |
| | b. | Jan _i | gaf | Piet _j | een | schop. | Daarna | gaf | <i>deze</i> _{j<i>1</i>} | <i>hem</i> _{j<i>2</i>} | een | dreun |
| | | Jan _i | gave | Piet | a | kick | thereafter | gave | this | him | a | slap |

21 The cyclic Spell-Out model of current minimalism would not in and of itself enable this, however: Spell-Out ‘freezes’ or ‘consolidates’ a portion of the bottom-up syntactic derivation, but does not assign it an LF/PF interpretation at that stage.

Encoding discourse function in the form of a feature of a syntactic constituent would, in keeping with the Inclusiveness Condition (only material that is part of the numeration can be introduced into the derivation), have to be tantamount to encoding discourse function in the form of a feature of a *head* inside this constituent. But features such as [+focus] or [+topic] are markedly different from such features as [+plural], [1st person], or [past tense] in that they are not the province of lexical/functional *heads*: they are ‘assigned from without’, not ‘projected from within’ a syntactic category. And as Horvath (2000, 2005) shows convincingly with reference to Hungarian, there appears to be no way to syntactically constrain the ‘percolation’ or ‘projection’ of a discourse-related feature from the head that carries the prosodic exponent of the feature to the phrasal constituent that undergoes the syntactic movement operation ‘triggered’ by the feature: pied-piping under focus movement is much more liberal than even the most liberal case of pied-piping under feature-driven movement (i.e., *wh*-movement).²² Horvath concludes from this that ‘focus movement’ is unlikely to be based on a feature-matching Agree relationship involving a feature [+focus] of (the head of) the constituent serving a particular discourse function. A parallel conclusion can be drawn for the feature [+topic], and whatever other ‘information-structural’ features one might appeal to.

It would seem much more natural to approach the syntactic distribution of topics and foci from the opposite end, i.e., from the perspective of the discourse. In particular, I would suggest (in line with the ‘syntax of interpretation’ model pursued in Barbiers 1995, for instance) that what syntax delivers (i.e., the configuration that results from applications of Merge and Move) receives a particular interpretation at the interpretive level, and that this interpretation should be compatible with the semantics and information structure of the utterance, or else it will be filtered out. For topics, this entails that they want to evacuate the VP, which is the typical domain of focus (cf. Reinhart 1995, Zubizarreta 1998). For foci, by contrast, this means (among other things) that they do not want to surface in SpecTP, the A-position *par excellence* for topics.

The constraint in (12), therefore, is an interface condition. A translation of (12) into a feature-driven model of movement *à la* minimalism seems pointless, or even fundamentally flawed. One may continue to use the functional projections TopP and FocP in one’s theory of syntax — so long as one does not ‘translate’ the labels of the heads of these functional projections into features that serve as triggers for syntactic displacement. Topics and foci may well have designated positions in the syntactic left periphery (cf. Rizzi 1997), and those designated positions may well be labelled ‘SpecTopP’ and ‘SpecFocP’ positions, but it is much more likely to be the function of these constituents in the larger discourse of which the sentence that contains them is a part than some property of some functional head in the structure of the individual sentence itself, that dictates that these constituents surface in the positions in which they do.

Syntactic movement that is interpretively innocuous and takes place purely for the sake of the satisfaction of some formal property of some functional head does arguably exist. Head movement may be in fact be the paradigm case of such movement — though the status of head movement as a syntactic operation has been called into question in recent work, at least in part because of the fact that it generally seems not to have any interpretive repercussions (but see, for instance, Lechner 2005).

22 For illustration, consider the Hungarian pair in (i), taken from Horvath’s work. Note that one could of course forgo ‘percolation’ or ‘projection’ (Selkirk 1984, 1995) of the [+focus] feature from the carrier of the prosodic exponent of focus, and say instead that the feature [+focus] is assigned to the syntactic head of the pied-piped constituent in (ia), i.e., *néhány* ‘some’. But such a move, by divorcing the feature [+focus] from its prosodic reflex, undermines the rationale for calling the feature in question [+focus]: then one might as well call the feature [+Q], some totally abstract morphosyntactic feature that happens (essentially by accident or divine providence) to be present on the head of focus-moved constituents. Such a move would quite transparently amount to a last-resort way of salvaging a feature-driven account of focus movement — fine in principle, though hardly insightful because of the entirely abstract nature of the feature in question.

(i)	a.	[<i>néhány</i> <u>MARILYN MONROE-RÓL</u> írt könyvet]	írt	könyvet]	olvastam	el	<i>t</i>
		some Marilyn Monroe-about	written	book-ACC	read-PST-1SG		PV
		‘it’s a few books written about MARILYN MONROE that I have read’					
	b.	*a filmszínésznő [néhány <u>akiről</u> írt könyvet]	írt	könyvet]	elolvastam		<i>t</i>]
		the movie-actress some who-about	written	book-ACC	PV-read-PST-1SG		

In the phrasal domain, classical ‘EPP-driven movement’ (i.e., NP-movement to SpecTP and, by minimal extension, to SpecvP as well) often does not seem information-structurally innocuous. But movement to SpecCP for the satisfaction of C’s [+WH] may in fact, surprisingly, be a good example of phrasal movement that is not inherently ‘interpretive’. The case of English (and also Serbo-Croatian; cf. Bošković 2002) is instructive here: *wh*-phrases raise to SpecFocP in root clauses but to SpecCP in embedded clauses; yet regardless of whether they raise to SpecFocP or SpecCP, they have a constant information-structural signature — they are the focus of the question (at least, in single *wh*-questions).

It seems, then, that purely formal, feature-driven movement really does exist. But a *purely* formal-feature-driven approach to *all* syntactic movement seems infeasible: it seems that the theory has to countenance, alongside purely formal-feature-driven movement, the existence of overt movement that does not get ‘driven from within’, but whose output is ‘evaluated from without’, by the interpretive component (LF and discourse pragmatics). It is entirely possible that only the kind of syntactic movement that is ‘driven from within’ is strictly speaking under the jurisdiction of a minimalist, derivational theory of syntax, and that movement that is ‘evaluated from without’ is regulated primarily by a representational theory of the type espoused in OT-syntactic work.²³ But that is a broader question that is, it seems to me, in principle independent of the more limited conclusion that not all narrow-syntactic movement is formal-feature-driven movement.

3.2 *Vacuous movement: PF and LF*

Now that we have gained a perspective on the nature of (12), what remains to be addressed is its consequences for the strong interpretation of the VMH that I am advocating in this paper. Plainly, the movement that focused *wh*-subjects must undergo in English root questions (SpecTP-to-SpecFocP; cf. (21a)²⁴) and in embedded questions (SpecTP-to-SpecCP; cf. (21b)) is string vacuous in standard English, where the C-head of embedded questions is necessarily phonologically empty. On the traditional interpretation of the VMH according to which ‘vacuous movement’ means ‘string-vacuous movement’ (hence constrains movement that has no effect on the *PF*-output), (21a,b) are an anomaly if, as per the strong VMH, vacuous movement is prohibited.

- (21) a. [_{FocP} *wh*_i [_{Foc_o} [_{TP} *wh*_i [T ...]]]]
 b. [_{CP} *wh*_i [_{C_o} [_{TP} *wh*_i [T ...]]]]

A rethinking of ‘vacuity’ such that it takes *both* interpretive components (not just *PF* but also *LF*) into consideration, as in (5), repeated below, allows us to reconcile obligatory string-vacuous movement in focus contexts with a strong version of the VMH. Movement of the focused subject to SpecFocP, as in (21a), crosses the *Foc*-head, which is semantically active: *Foc* is intrinsically meaningful — the existence of overt focus particles occupying *Foc* (see e.g. Kayne 1998) lends support to this assumption; Horvath’s (2000, 2005) covert ‘exhaustive identification’ operator may reside under *Foc*, too (though Horvath herself does not place it there: she adjoins it to the constituent undergoing focus movement). And movement to SpecCP crosses the *C*-head, which, in questions, is semantically visible as well: it harbours the question operator (recall (14)), and thereby types the clause as a question.

23 For interesting work that aims to combine formal-feature-driven derivational syntax with evaluative/representational OT-syntax, see Broekhuis & Dekkers (2000).

24 Recall that in English root *wh*-questions, the CP-layer is not activated in overt syntax at all, overt *wh*-movement targeting the specifier position of *FocP* rather than *SpecCP* (see Den Dikken & Giannakidou 2002, Den Dikken 2003a). The *FocP* structure in (21a) does not conflict with what was said in section 3.1, for reasons made explicit there. This substructure may or may not be an integral component of the structure of English embedded *wh*-questions. If it is, then the *wh*-phrase raises straight into *SpecCP* without a stopover in *SpecFocP* (cf. Den Dikken 2003a for discussion), crossing both *Foc* and *C*. If it is not, *FocP* and *CP* are presumably conflated in English embedded questions. (21b) abstracts away from this question for the sake of simplicity.

- (5) *Vacuous Movement Hypothesis — my version*
 movement that does not cross phonologically or semantically visible material is prohibited

So (22a) and (22b) are both ruled in thanks to the fact that movement crosses semantically visible material; and so, of course, are (22c) and (22d), where phonologically visible material is crossed. For (22d), this is straightforward — though standard English does not employ this strategy: root *wh*-questions do not activate CP in overt syntax; embedded *wh*-questions do not have C overtly filled in the standard dialect, though there are varieties of English (Belfast English, for instance) in which (22d) does in fact materialise in embedded contexts. That (22c) is ruled in by the VMH may seem to be an embarrassment — after all, **who do you think that left?* is ungrammatical in standard English. But arguably *that*-trace effects are not within the jurisdiction of the VMH — rather, some other principle of the grammar must be held responsible for the fact that (22c) does not come out grammatical across the board, depending on language-particular properties.²⁵

- (22) a. $[_{FocP} wh_i [Foc_{[\sigma_{PHON}; +SEM]} [_{TP} \overline{wh}_i [T \dots]]]]$
 b. $[_{CP} wh_i [C_{[\sigma_{PHON}; +SEM]} [_{TP} \overline{wh}_i [T \dots]]]]$
 c. $[_{CP} wh_i [C_{[-PHON; +SEM]} [_{TP} \overline{wh}_i [T \dots]]]]$
 d. $[_{CP} wh_i [C_{[+PHON; +SEM]} [_{TP} \overline{wh}_i [T \dots]]]]$
 e. $[_{CP} wh_i [C_{[\sigma_{PHON}; +SEM]} [_{TP} \overline{wh}_i [T \dots]]]]$

There is one member of the typology in (22) left to be discussed, (22e). This is an interesting case whose resolution depends directly on one’s assumptions regarding the nature of the Vacuous Movement Hypothesis — as a condition on derivations or representations. I will turn to this question next.

3.3 *On the nature of the VMH*

The derivation in (22e) involves local raising of the highest subject into the specifier of a CP whose head, at the point in the derivation at which *wh*-movement takes place, is both semantically and phonologically empty. If this is all there is to (22e), then obviously it will be rejected as a violator of the VMH, which in turn means that grammatical sentences such as *who do you think left?*, which are commonly analysed precisely along the lines of (22e), cannot in fact involve a derivation of this sort — either the subject is being extracted directly from its base position inside vP (cf. Chomsky 2004:16, and Rizzi 1982 for Italian), or the phonologically and semantically empty CP layer is not projected at all. For English, Bošković (1997) and Grimshaw (1997) have argued that the latter is the right way to go.

25 Pesetsky & Torrego (2001) have a concrete proposal regarding the ill-formedness of **who do you think that left?* that relies crucially on T-to-C movement. For them, C has an uninterpretable tense feature (*uT*) that must be checked. In *wh*-questions, T-to-C movement is driven by the need to get C’s *uT* feature checked. But T-to-C movement is not systematically obligatory in all root *wh*-questions — in English highest-subject *wh*-questions, it never takes place. That is because raising the subject to SpecCP ‘kills two birds with one stone’, as it were: the raised subject checks both C’s *uWh*-feature and C’s *uT*-feature (thanks to the fact that ‘nominative Case’ is an instantiation of *uT* itself); so raising T to C is uneconomical in highest-subject *wh*-questions (cf. (ia)). Assuming that *that* is the surface spell-out of overt T-to-C raising in embedded clauses, Pesetsky & Torrego then straightforwardly explain the ban on *that* in cases of highest-subject *wh*-extraction. The Pesetsky & Torrego account of *that*-trace effects is compatible with the discussion in this paper, and does not seem to interfere directly with anything said here. (I will not discuss here the question of how dialects of English allowing *that* in C in local subject questions (such as **I wonder who that kissed Mary*) should be analysed; Pesetsky & Torrego have an account for this. It remains to be investigated whether such dialects allow *that* to show up in multiple questions whose highest subject is a *wh*-phrase interpreted as a universal quantifier (*I wonder who that ate what*; cf. (9b)) — it is certainly an interesting question to ask whether the dialects in question make a distinction between the two readings of ... *who (that) ate what* when it comes to the overt realisation of the complementiser: my analysis of multiple *wh*-questions would predict that, with *that* present (assuming that it occupies C), a wide-scope universal quantifier reading for *who* should be unavailable.)

- (i) a. $[_{CP} SUBJECT_i [C (*+T_i) [_{TP} t_i [T_i [_{\sigma} \dots]]]]]$
 b. $[_{CP} NON-SUBJECT_n [C (*+T_n) [_{TP} SUBJECT_i [t_i [_{\sigma} \dots t_n \dots]]]]]$

It is not a necessary conclusion, however, that the VMH leads to rejection of (22e). Much depends here on the question of whether the VMH is conceived of as a condition on derivations or representations, and whether the LF representation of (22e) is identical with its overt-syntactic signature — i.e., whether C remains semantically empty throughout the derivation. If T raises to C in the LF derivation of the embedded clause of sentences such as *who do you think left?*,²⁶ then by LF C is no longer semantically empty, given that tense is semantically meaningful.²⁷ If the VMH is evaluated at LF, then, the outcome for (22e) will be positive with T raising covertly to C. Without wishing to push a particular interpretation of the VMH, let me contemplate the pros and cons of derivational and representational approaches to the VMH within a somewhat broader conceptual context.

The VMH as stated in (5) (in line with the traditional formulation; cf. (2)) is plainly a condition on *derivations*, not a condition on (LF/PF) *representations*: it is stated in terms of *movement* operations themselves, not in terms of their outputs. Thus, if at the point of the derivation at which a certain movement operation takes place, the positions that this movement operation crosses are all devoid of phonological and semantic content, the derivation is rejected *at that point*, with no hope of retroactive salvation. In other words, operations such as T-to-C movement *at LF* (the existence of which in itself is a contentious issue in an Agree-based system) could have no effect on output as far as the VMH is concerned, if the VMH is indeed a condition on derivations. That is, even with T raising to C at LF in, say, (22e), the derivation depicted by this structure will still be filtered out on this interpretation of the VMH.

Interpreting the VMH as a condition on derivations only makes sense if phonological and semantic features are *fully specified* in the syntactic component — if, that is, phonological and semantic features ‘are there’ already in syntax. If phonological and semantic features are supplied only after Spell-Out (as in Distributed Morphology), the VMH *qua* condition on derivations will (trivially) rule out *all* movement prior to Spell-Out. In a late insertion model such as DM, therefore, the VMH should either be rejected wholesale, or be reinterpreted as a condition on *representations*, checking/filtering outputs of syntactic movement rather than adjudicating syntactic movements as they happen. With the VMH in (5) interpreted as a condition on representations, including those of the LF component, the desired results in the domain of the *wh*-movement constructions discussed in section 2 will continue to carry over — but the derivational and representational perspectives on the VMH yield potentially different results for (22e), which, as we have seen, is ruled out on a derivational interpretation of the VMH, but will be grammatical on a representational approach if T raises to C at LF so that, by the time the VMH is evaluated, semantically meaningful material (*tense*) will have ‘arrived’ in C and the dependency between *wh*_i and its trace is no longer a ‘vacuous movement’ dependency.

A representational approach to the VMH thus has two potential advantages: (i) it makes the VMH compatible with a late insertion approach to phonological and semantic features *à la* Distributed Morphology, and (ii) it allows us to preserve (22e) (with T-to-C movement at LF rendering C meaningful by LF). To what extent (ii) is a genuine pro of the representational approach to the VMH depends on one’s outlook on the syntax of sentences such as *who do you think left?* — the embedded clause here could be ‘dressed down’ to TP or even less (Bošković 1997, Grimshaw 1997), or the subject could be launched from a position lower than SpecTP, crossing the T-head (Chomsky 2004:16). I will not take a stand here on the possibility of smaller-than-CP accounts of the embedded clause. But the discussion of highest-subject relative clauses in section 4 bears directly on the feasibility of launching the *wh*-subject from a position lower than SpecTP, and may, depending on one’s assumptions regarding the distribution of covert T-to-C movement, also rule on the fate of a representational outlook on the VMH as a filter on LF and PF outputs.

26 Recall from fn. 25 that Pesetsky & Torrego (2001) have the complementiser *that* be the surface spell-out of overt T-to-C raising in embedded contexts. So for Pesetsky & Torrego, T-to-C raising certainly can happen in embedded contexts. What can happen in overt syntax in some contexts could happen covertly in others.

27 By saying that tense is meaningful, I am not making any claims about tense *morphology*, which often is not meaning-bearing at all or is not associated with a temporal interpretation (see sequence-of-tense phenomena, where embedded past tense is semantically meaningless, and the past-tense forms of English modals, which do not have a temporal meaning). It is the *temporal operator* in T that is meaningful, not the tense feature or the associated tense morphology.

4 The Vacuous Movement Hypothesis and highest-subject contact relatives

So far I have only been concerned in this paper with cases of extraction from highest-subject relatives and *wh*-questions. I have so far sidestepped the question of why highest-subject relatives and *wh*-questions differ markedly from one another when it comes to the distribution of overt material in C. The Vacuous Movement Hypothesis in (5) contributes to this venerable puzzle in English syntax in an interesting way, as I will show in this section.

4.1 The VMH and the ban on highest-subject contact relatives in standard English

Let me start by reiterating that, as I pointed out in section 3.2, the VMH *per se* does not rule on the (im)possibility of extraction of the highest subject across a lexical complementiser — that is, the fact that a structure of the type in (22c), repeated below in a slightly modified form as (23), fails in standard English cases of long subject extraction (**who do you think that left?*) has nothing to do with the VMH, which in fact rules this in.

(23) $*[_{CP} wh_i [C_{[-\text{PHON}; \text{SEM}]=\text{that}} [_{TP} \overline{wh}_i [T \dots]]]$

But what I would like to argue here is that the VMH in (5) does allow us to understand the ‘anti-*that*-trace’ effect seen in standard English relatives such as (24): the fact that the complementiser *must* be overt in highest-subject relative clauses whenever the *wh*-operator is null.

(24) the guy $*(\text{that})$ fixed the sink is a lousy plumber

The fact that in the standard language (24) is ungrammatical with *that* omitted follows straightforwardly from the VMH in (5), in conjunction with the hypothesis that null operators cannot stay in an A–position — a hypothesis that receives independent support from the fact that null operators cannot pied-pipe a preposition, as seen in (25).

(25) a. a problem *Op* I’d like to talk [about *t*] a’. a problem *Op* to talk [about *t*]
 b. $*\text{a problem [about } Op] \text{ I’d like to talk } t$ b’. $*\text{a problem [about } Op] \text{ to talk } t$
 c. a problem [about which] I’d like to talk *t* c’. $^? \text{a problem [about which] to talk } t$

The root of the problem with (25b,b’) need not concern us here;²⁸ I will take these examples at face value to support the claim that null operators cannot stay in A–positions.²⁹ The combination of this claim and the VMH in (5) now gives us the following verdicts on the configurations in (26):

28 One possibility explored in the literature is that (25b,b’) fail because of a violation of the PRO Theorem — assuming that *Op* = PRO (the only reasonable option if *Op* is to be a member of the typology of empty noun-phrase types), (25b,b’) violate the ban on government of PRO. The status of both PRO and the PRO Theorem in current theory are dubious, however. I will not explore alternative accounts of the impossibility of pied-piping involving null operators.

29 The ban on null operators in A–positions makes a prediction in the context of extraction from highest-subject relatives, in light of the discussion in section 2 of this paper: such extraction should succeed only from *wh*-relatives (which, by the VMH, have the *wh*-operator in SpecTP), not from *that*-relatives (which, given the ban on null operators in A–positions, must involve operator movement to SpecCP). It certainly is conspicuous that all of Chung & McCloskey’s (1983) subject-relative examples are *who*-relatives, not *that*-relatives (cf. (6)). Jim McCloskey (p.c.) has confirmed to me that all the attested examples of extraction from highest-subject relatives that he has come across involve *wh*-relatives; but he adds that, for him, replacing *who* with *that* would not render the examples ungrammatical, which suggests that it is not an absolute requirement that the relative be introduced by a *wh*-operator. I do not have an account at this time of successful extraction from highest-subject *that*-relatives. In light of the text discussion, I am certainly committed to the idea that null operators cannot stay in A–positions. A general possibility that comes to mind is that *that*-relatives can mimic the behaviour of highest-subject *wh*-relatives by analogy. But a question that would inevitably arise for such an approach is how a derivation that is rejected in syntax can somehow make it to PF by way of its grammatical analogue.

(26) a. $[_{CP} \text{ NULL-OP}_i [C_{[-\text{PHON}; \text{SEM}]=\text{that}} [_{TP} t_i [T \dots]]]$
 b. $*[_{CP} \text{ NULL-OP}_i [C_{[-\text{PHON}; \text{SEM}]=\emptyset} [_{TP} t_i [T \dots]]]$
 c. $*[_{CP} [C_{[-\text{PHON}; \text{SEM}]=\emptyset} [_{TP} \text{ NULL-OP}_i [T \dots]]]$

(26a) obeys the VMH as well as the *Op*-in-A’ constraint, hence delivers a grammatical output. But (26c) violates the ban on null operators in A–positions, and (26b) the VMH — straightforwardly so on a *derivational* outlook on (5); a *representational* construal of (5) will only rule (26b) out if T never raises to C in relatives (though it might so raise in embedded questions, depending on one’s outlook on (22e)). These are the only possibilities that standard English has at its disposal. Hence standard English forces *that* to surface in (24).

4.2 The VMH and grammatical highest-subject contact relatives in different varieties of English

Besides the structures in (26a–c), there is one additional logical possibility, given in (26d). In (26d), the null operator raises from a position in the complement of T straight into SpecCP (cf. also Chomsky 2004:16). This derivation of course meets the *Op*-in-A’ constraint like (26b), with which (26d) shares the absence of phonological material in C. But unlike (26b), the derivation in (26d) satisfies the VMH in (5) thanks to the fact that movement of the subject from an extraction site in T’s complement to SpecCP crosses T, a semantically contentful head: tense is meaningful, hence (26d) is not a case of semantically vacuous movement.

(26) d. $[_{CP} \text{ NULL-OP}_i [C_{[-\text{PHON}; \text{SEM}]=\emptyset} [_{TP} \text{ ———} [T_{[-\text{PHON}; \text{SEM}]} [t_i \dots]]]]$

The derivation in (26d) results in a surface output for highest-subject relative clauses in which no overt operator or complementiser shows up in the left periphery — in other words, it delivers the version of (24) lacking *that*. In standard English, this is ill-formed: standard English cannot avail itself of the derivation in (26d) — *contra* what Chomsky (2004:16) assumes. But there are varieties of English in which highest-subject ‘contact relatives’ (to use Jespersen’s 1961:Vol. III, pp. 81, 132ff. term) are well-formed. I would like to argue that such varieties typically differ in another important way from standard English as well: they allow the subject to surface in a position below SpecTP. This is precisely what is expected from the point of view of the VMH-based account of (26b vs d): highest-subject ‘contact relatives’ can grammatically result only from a derivation of the type in (26d), whose grammaticality is in turn contingent on the subject being able to be A’-launched from, hence fully A–licensed in, a position lower than SpecTP. Let me lay out the details.

In many varieties of English (including Appalachian English, Belfast English, Scots English, and Ocracoke English³⁰), non-pronominal plural subjects often do not show agreement with the finite verb. Thus, speakers of such varieties accept (27b) alongside (27a), though they typically will not accept (28b).³¹

(27) a. the girls like pizza (28) a. they like pizza
 b. $^? \text{the girls likes pizza}$ b. $^? \text{they likes pizza}$

For Henry (1995), the distribution of agreement on the finite verb is a reflex of whether or not a Spec–Head relationship is established between the subject and the functional head checking agreement (AgrS for her, T in current Agr-less structures) — thus, when the subject raises into the specifier position of this functional head, agreement automatically ensues; when no raising transpires, no Spec–Head relationship is established, and no phi-feature agreement is morphologically manifested. This is summarised in (29a).

30 On Appalachian English, see Wolfram & Christian (1976:83); on Belfast English, see Milroy (1981), Henry (1995); on Scots English, see Adger & Smith (2005), Hazen (1996, 2000); on Ocracoke English, see Hazen (1996, 2000).

31 For Appalachian English, Wolfram & Christian’s (1976:83) statistics are very clear on this difference between pronominal and non-pronominal plural subjects: a full 59% of non-pronominal plural subjects occur in pattern (27b), whereas less than 1% of pronominal subjects partake in (28b). (See below, however, for a qualification based on the role of past tense.) The difference between pronominal and full-nominal subjects manifests itself also in the *wh*-agreement facts reported in Kimball & Aissen (1971): *the people who Clark/he thinks are in the garden* vs. (North-Eastern U.S.) *the people who Clark/ $^? \text{he}$ think are in the garden*.

- (29) a. [_{TP} <AGREEING SUBJECT> [_T T (...)] [(NON-AGREEING SUBJECT) ...]]
 b. [_{TP} <FULL-NOM/PRONOM SUBJECT> [_T T (...)] [(FULL-NOM/*PRONOM SUBJECT) ...]]

The contrast between (27b) and (28b) then follows on the further assumption that there is a positional difference between full-nominal and pronominal subjects (cf. (29b)); pronominal subjects must raise to the canonical structural subject position (the specifier of the functional head checking agreement) and then necessarily control agreement with the finite verb, whereas non-pronominal subjects may (but do not have to) stay in a position further down the tree, and when they do, they then fail to show morphological agreement with the finite verb.³²

Henry (1995) supports her claim that there is a positional difference between agreeing and non-agreeing subjects by observing, among other things, that the subject in Belfast English can be a negative polarity item (NPI) only if it does not agree with the verb, as the contrast between (30a) and (30b) shows (cf. Duffield 1993, Henry 1995:27–29).

- (30) a. any animals isn't coming (Belfast English)
 b. *any animals aren't coming
 c. [_{TP} <AGREEING SUBJECT: *NPI> [_T T [_{NegP} Neg (...)] [(NON-AGREEING SUBJECT: ✓NPI) ...]]]

These facts follow straightforwardly from the skeletal structure in (30c), with only the non-agreeing subject being c-commanded by Neg⁰.

Appalachian English allows non-pronominal plural subjects to show no concord with the finite verb as well. And there is evidence that in Appalachian English, too, lack of agreement with the finite verb is a reflex of the placement of the non-pronominal subject in a position below the canonical structural subject position, SpecTP. This is apparent in particular from the fact (illustrated in (31)); data taken from Montgomery & Hall 2004 and Wolfram & Christian 1976:113) that in negative concord constructions in Appalachian English, a negative subject may surface to the right of the finite auxiliary (hosting *-n't*), producing a subject-auxiliary inversion pattern (see Sells *et al.* 1996) which, unlike in standard English, even manifests itself in relative clauses, as shown in (31c). Since raising the finite auxiliary to a functional head position outside the inflectional domain is generally impossible in relative clauses (cf. **a room that under no circumstances would he stay in*), the word order in (31c), and — by extension — that in (31a,b) as well, arguably results from failure on the part of the non-pronominal subject to raise up into the canonical structural subject position, SpecTP. Interpreted this way, the facts in (31) (and esp. (31c)) support the conclusion, drawn for Belfast English by Henry (1995) as well, that non-pronominal subjects can stay in a position lower than SpecTP in Appalachian English.

- (31) a. didn't nobody get hurt or nothin' (Appalachian English)
 b. wasn't nothin' but acorns on the ground ... and wasn't nobody there
 c. it had this room that wouldn't nobody stay in
 (32) [_{CP} Op_i [_C that [_{TP} ——— [_T wouldn't_i [_{NegP} t_i [(...)] nobody_j t_j stay in t_k]]]]]]]

32 A placement difference between pronominal and non-pronominal subjects such that the former are high up the tree whereas the latter may be lower is familiar from the facts of standard English verb-particle constructions as well (*John looked (the information) up (the information/*it)*). It manifests itself in the domain of Object Shift in the Mainland Scandinavian as well. The fact that there is a difference between placement in SpecTP and a lower structural position when it comes to the morphological manifestation of the Agree relationship between T and the subject does not make the English varieties under discussion unique or unusual either. As is well known, Arabic forces only preverbal subjects to fully morphologically agree with the finite verb; similarly, in Italian dialects postverbal subjects may or must fail to morphologically agree with the verb. Though for the purposes of Case checking, T will effectively establish an Agree relationship with the subject throughout (i.e., regardless of whether the subject raises to SpecTP or stays in a lower position), only if the subject raises to SpecTP must this Agree relationship be fully morphologically manifest. This has led Guasti & Rizzi (2002) to make a distinction between pure Agree and Spec-Head agreement, the latter being a tighter relationship than the former. For interesting evidence from 'agreement attraction' phenomena in Romance (cf. English *the label on the bottles is/are dirty*) that shows that this distinction is psychologically real, see Franck, Lassi, Frauenfelder & Rizzi (2003).

For both Belfast English and Appalachian English, then, there appears to be reason to believe that non-pronominal subjects do not necessarily raise to the canonical structural subject position, SpecTP.³³ And interestingly, both varieties of English are also characterised by the fact that they allow what Doherty (1993) has dubbed 'subject contact relatives' (borrowing the term 'contact relative' from Jespersen 1961) — i.e., highest-subject relative clauses that are not introduced by an overt *wh*-operator or the complementiser *that*. Examples of subject contact relatives from Belfast English are given in (33); the subject contact relatives in (34) are from Appalachian English.³⁴

- (33) a. there's one woman in our street went to Spain last year (Belfast English)
 b. it's always me pays the gas bill
 c. I have one student can speak five languages
 d. he's the one stole the money
 (34) a. I got some kin people lived up there (Appalachian English)
 b. he's the funny lookin' character plays baseball
 c. my grandma's got this thing tells me about when to plant
 d. 'cause they was this vampire that killed people come in it
 e. at first, you wouldn't believe the characters come knocked on my door
 f. but he tied the company up some way to get a royalty off the timber was cut for the mines

For Belfast English, Henry (1995) argues that subject contact relatives of the type in (33) are not garden-variety highest-subject relatives with a *who* left out. An important argument to this effect is that they seem to only occur in presentational contexts (Henry 1995:126; see also Den Dikken 2003b) — contexts in which the 'head' of the relative clause is introduced on the scene. But though in Appalachian English as well, many subject contact relatives appear in presentational contexts, an investigation of the empirical material in Wolfram & Christian (1976:121) as well as preliminary fieldwork conducted by Judy Bernstein and Christina Tortora on the *Dante Oral History Project* (see fn. 34) reveals that they certainly do not seem to be strictly confined to such contexts: the subject contact relatives in (34d–f), unlike those in (33) and (34a–c), do not involve matrix clauses that introduce NPs as new players on the scene of the discourse. I will assume, in light of the Appalachian English facts, that subject contact relatives are not categorically restricted to presentational contexts in the English-speaking world, therefore.

Recapitulating so far, what we have seen is that (a) there are varieties of English that do not enforce agreement between the finite verb and a plural subject, (b) there is evidence that the non-agreeing subject in these varieties occupies a position lower in the tree than SpecTP, and (c) at least some of the varieties of English that allow non-agreeing plural subjects also feature so-called subject contact relatives of the type in (33) and (34). Appalachian English is particularly interesting in this connection because its subject contact relatives do not appear to be confined to presentational contexts, hence cannot be consistently reanalysed as non-relative topic-comment structures (*à la* Henry 1995) — the subject contact relatives in (34d–f) seem to be genuine relative clauses lacking an overt *wh*-operator or complementiser in their left periphery.

I would like to contend that properties (a)–(c) are interrelated. In particular, I would like to argue that it is precisely the fact that subjects can be licensed in a position lower in the tree than SpecTP (property (b)) that allows them to be relativised via null operator movement into the specifier of a zero-headed CP in highest-subject relatives (property (c)). Put differently, the grammaticality of (35a) is what facilitates a derivation of subject contact relatives via (35b), which satisfies the Vacuous Movement Hypothesis in (5) thanks to the fact that movement crosses over the semantically visible T–head.

33 I am not aware of any specific evidence to this effect in the extant literature on other varieties of English that allow lack of concord between a non-pronominal plural subject and the finite verb (including Scots English and Ocracoke English).

34 The Belfast English examples in (33) are from Henry (1995); the Appalachian English sentences in (34a–d) were taken from Hackenberg 1972, Wolfram & Christian 1976:121, whereas those in (34e–f) are from the *Dante Oral History Project* at Eastern Tennessee State University in Johnson City, TN.

- (35) a. $[_{TP} \text{ — } [_{T_{[\text{PHON};\text{SEM}]}]} \text{ [XP SUBJECT ...]}]]$
 b. $[_{CP} \text{ NULL-OP}_i \text{ [C}_{[\text{PHON};\text{SEM}]} = \emptyset \text{ } [_{TP} \text{ — } [_{T_{[\text{PHON};\text{SEM}]}]} [t_i \dots]]]]]$ (= (26d))

This approach to subject contact relatives makes the prediction that only *non-agreeing* subjects should be relativisable this way — in light of both the logic and the empirical evidence in support of Henry’s (1995) account, only non-agreeing subjects are allowed to stay in the low subject position indicated in (35a), hence it should only be non-agreeing subjects that could be launched up to SpecCP crossing over T. In fact, the analysis makes a more microscopic prediction, in view of the fact that the relative operator in (35b) is pronominal (arguably, PRO; recall fn. 28): it predicts that (35b) should succeed only in contexts in which (35a) is grammatical with *pronominal* subjects. For Belfast English, this tells us that (35b) should be altogether impossible: Henry (1995) shows that (35a) in Belfast English is categorically restricted to *non-pronominal* subjects (recall (27b) vs (28b)). Belfast English, therefore, is expected not to be able to produce genuine highest-subject relatives with null-operator movement to the specifier position of a zero-headed CP — which accords with Henry’s (1995) assessment of the empirical lie of the land: subject contact ‘relatives’ in Belfast English, according to Henry, are not genuine relatives but topic–comment structures instead (see also Den Dikken 2003b).

Appalachian English, on the other hand, does provide (35b) with a chance of survival. For in Appalachian English, pronouns are not categorically banned from the low subject position in (35a). Wolfram & Christian’s (1976:83) figures for the variety of Appalachian English they studied show that plural pronominal subjects co-occur with the singular *past-tense* form *was* in 79.8% of all cases — a spectacular finding in light of the fact that 0.7% of plural pronominal subjects are used with singular present-tense *is*.³⁵ This remarkable discrepancy in Appalachian English between *was* and *is* may suggest that the tense specification of T plays a role when it comes to the possibility on the part of pronominal subjects to stay in the lower subject position depicted in (35a): in particular, it may suggest that pronominal subjects are allowed to stay down low only in past-tense contexts. If this is the right interpretation of the facts (in other words, if *they was ...* is an anti-agreement effect resulting from low subject placement),³⁶ this bears directly on the distribution of (35b) in these varieties: what it suggests is that (35b) should succeed only in *past-tense* contexts, not in present-tense contexts, where pronominal subjects are never allowed to stay down low.

The few attested examples of subject contact relatives that are not amenable to a reanalysis in terms of topic–comment structures (i.e., those in (34d–f)) are compatible with this prediction: they are all past-tense examples.³⁷ Whether there is indeed a tense effect on the licensing of (35b) is something that Judy Bernstein, Christina Tortora, Raffaella Zanuttini and myself intend to investigate in detail in future research; for now, I have to leave a detailed verification of this prediction open. But in any case, what we can say at this time is that what little evidence there is in the extant literature certainly supports the line of thought outlined here.

35 See also Montgomery & Hall’s (2004:24) observation that ‘[i]n traditional Smokies speech [a variety of Appalachian English], *was* and *were* are used for both singular and plural [in all persons], but there is today and apparently has always been a strong preference for *was* in all persons and numbers’. This ‘levelling’ of subject-verb agreement is not specific to Appalachian English, as attested in the substantial literature on this phenomenon, which shows a high degree of complexity. In South-Western England, *were* is generalised for the affirmative and *wasn’t* or *weren’t* for the negative. In some US dialects, *was* is always used in affirmative clauses and *weren’t* in negative ones (see Hazen 1996 and 2000, Schilling-Estes & Wolfram 1994, Tagliamonte 1998, Tagliamonte & Smith 1999, Wolfram & Sellers 1999). Appalachian English differs conspicuously from other varieties of English that show this ‘levelling’, however, when it comes to the role of past tense. Thus, in Belfast English and similarly in Scots English (Adger & Smith 2005 on the Buckie variety), *they was* is equally non-existent as *they is*, unlike in Appalachian English.

36 There is an alternative interpretation — one that is syntactically much less interesting but which further research may show to be better supported by the data: it is possible that the past-tense form of *be* in (the relevant varieties of) Appalachian English is quite simply *was* throughout the paradigm (i.e., the past-tense paradigm of *be* does not include *were* at all).

37 The reader should bear in mind that *come* (seen in (34d)) is a past-tense form in Appalachian English.

5 Concluding remarks

Let me close the discussion by first of all presenting a brief recapitulation of the main findings, and then exploring the repercussions of the revised Vacuous Movement Hypothesis in (5) for the checking model for triggering movement.

Starting out from Chomsky’s (1986) weak formulation of the VMH in (2a) (‘vacuous movement is not obligatory at S-Structure’), I proceeded in this paper by exploring a stronger version of the VMH which blocks all vacuous movement. I showed that the strong VMH, in conjunction with a detailed analysis of the syntax of single and multiple *wh*-questions (based on seminal work by É. Kiss 1993), correctly derives the generalisation that extraction from English highest-subject *wh*-questions and relative clauses is possible precisely in contexts in which the subject remains *in situ* and does not raise to SpecCP — that is, in all highest-subject *wh*-relatives and also in *multiple wh*-questions in which the *wh*-subject is not interpreted as a focus (but instead receives a [+specific] interpretation). This result, confirmed by the contrast between (7) and (7’) and by the facts in (16), is not obtainable from any other account of *wh*-questions and relative clauses that I am familiar with.³⁸

It is worth emphasising that my analysis of highest-subject *wh*-questions and relative clauses has the *wh*-subject staying *in situ* only in a relatively small subset of cases: in highest-subject *wh*-relatives (where the *wh*-constituent it is not focused and where, due to C’s phonological and semantic emptiness, movement to SpecCP would inevitably violate the VMH unless the *wh*-subject is launched from a position in T’s complement), and in *multiple* questions in which the highest subject is interpreted as a wide-scope universal quantifier. In all contexts in which the subject is *focused* (including all *single* subject questions), the *wh*-subject is *ex situ*, as a consequence of the fact that focused subjects are barred from surfacing in SpecTP (recall (12)). In this respect, my analysis is significantly different from extant accounts (including Agbayani 2000, and the very interesting but unfortunately rather too categorical analysis presented in Surányi & Márkus 2005), which typically treat all highest-subject *wh*-constructions on a par.

The conclusion, drawn in the course of the investigation of the key facts in (7), that *wh*-foci are not allowed to stay *in situ* in SpecTP and must raise (to SpecCP in English embedded questions) instead, then led me to refine the VMH by giving both interpretive components (PF as well as LF) a say in it. Defining ‘vacuous movement’ as movement that does not cross any phonologically or *semantically* visible material, as in (5), in addition opened up a new perspective on the syntax of highest-subject relative clauses. In particular, in conjunction with the independently supported hypothesis that null operators are not allowed to stay *in situ*, ‘bipolar’ (5) was seen to make the prediction that highest-subject ‘contact relatives’ (i.e., relative clauses not introduced by an overt *wh*-operator or complementiser) can exist only in languages in which the null-operator subject can be launched from a position lower down the tree than SpecTP — a prediction diagnosing an empirically valuable distinction between standard English, which does not allow subjects to stay down low and which, concomitantly, does not have highest-subject contact relatives, and varieties of English that are more liberal when it comes to subjects being licensed in a position below SpecTP, where highest-subject contact relatives do indeed show up. Of course, a blanket assumption to the effect that *wh*-subjects *always* get launched from their vP–internal base position, as in Chomsky (2004:16), cannot derive this result.

38 I can think of one potentially interesting way of capturing the gist of the Chung & McCloskey (1983) data *without* an appeal to the VMH (or, equivalently, to notions such as those touched upon in fn. 7, above). Assume that CP is not inherently a phase (inherent phasehood being restricted to small clauses), but ‘inherits’ phasehood from a phase in its complement whenever the head of this phase makes its way up to C. In particular, assume that when *v* (which is the head of the vP phase) raises to T, this extends vP’s phasehood up to T; and that when *v+T* subsequently raises to C, this extends phasehood further up to CP (cf. Den Dikken 2006 on the notion of ‘phase extension’ as a result of movement of a phase head). Following Pesetsky & Torrego’s (2001) idea that (v+T)–to–C movement must obtain in all embedded CPs *except* those that feature movement of the highest subject to SpecCP (recall fn. 25), we then derive the result that CP is a phase *except* in the case of highest-subject *wh*-CPs. This will make extraction from highest-subject *wh*-CPs significantly easier than from other *wh*-CPs, which are predicted to be impenetrable. The result comes close to deriving the Chung & McCloskey (1983) facts — but *unlike* the VMH-based account developed in this paper, it has nothing to say about the contrast between (7) and (7’), i.e., the significance of the single/multiple-question distinction.

One of the results of the discussion of extraction from highest-subject *wh*-relatives and *wh*-questions in section 2 of this paper is that the VMH supports an analysis of A'-extraction as proceeding successively-cyclically through SpecCP. In this respect, it confirms the standard Chomskyan approach to long A'-movement. On the other hand, it raises significant questions about the motivation (or trigger) for movement. I have already had occasion to dwell on the implications of the ban on focused *wh*-subjects in SpecTP (12) in section 3.1, above. There, I expressed skepticism concerning the feasibility of a feature-driven approach to overt-syntactic movement operations that serve a discourse-functional purpose. But more generally, outside the realm of information-structurally 'driven' operations, the VMH in (5) leads one to conclude that a strictly probe-centred approach to all movement triggers is questionable. Consider, for instance, the case of a [+WH] C-head, as found in the syntactic structure of relative clauses and *wh*-questions. There is little doubt that there are languages in which this C-head receives *wh*-phrases in its specifier position in overt syntax in at least some syntactic contexts. Thus, in English, *wh*-phrases raise to the specifier of the [+WH] C-head of embedded clauses (embedded *wh*-questions and all relative clauses).³⁹ But if the discussion in this paper holds water, we cannot hold some property ('strength' or 'EPP') of the [+WH] C-head of embedded clauses responsible for this fact: what I have endeavoured to demonstrate is that, in keeping with the demands of the VMH, *some* highest-subject *wh*-questions and *all* highest-subject *wh*-relatives involve no *wh*-movement to SpecCP. If [+WH] C had some intrinsic property causing it to attract *wh*-phrases up to its specifier position in overt syntax, that would necessarily force overt-syntactic *wh*-fronting to SpecCP across the board, whenever [+WH] C is merged in overt syntax.

To the extent that there are probes that are both phonologically and semantically void, a theory incorporating the VMH should broadly countenance the possibility of movement driven by properties other than those of probes. Some researchers have already started exploring the possibilities of non-probe-driven movement (cf. e.g. Barbiers 1995, Van Riemsdijk 1995, Platzack 1996, Koenenman 2000, Zwart 2004, Franks in press, and Van Craenenbroeck 2006). This paper makes the modest theoretical contribution of bringing this issue into sharper focus.

Acknowledgements

This paper was written for oral presentation at the 29th GLOW Colloquium in Barcelona, on 8 April 2006. An earlier version was presented at the Linguistics Colloquium of the University of Kansas, Lawrence, on 27 March 2006. Many ideas reported in this paper were first developed in my fall 2005 Advanced Syntax seminar on the syntax of locality, at the CUNY Graduate Center. I thank all participants of the seminar for making this such a lively and productive seminar. I am also very grateful for the feedback provided by the Kansas audience. Finally, thanks are due to Judy Bernstein, Christina Tortora, and Raffaella Zanuttini for discussion of the material in section 4.2, which is part of a joint research proposal on the comparative morpho-syntax of Appalachian English. This paper is in many ways exploratory, and should be read in this light. Naturally, the responsibility for all of the paper's no doubt numerous imperfections is entirely my own.

References

- Adger, David & Jennifer Smith. 2005. Variation and the Minimalist Program. Ms., Queen Mary/University of London & University of York.
- Agbayani, Brian. 2000. *Wh*-subjects in English and the vacuous movement hypothesis. *Linguistic Inquiry* 31. 703–13.
- Barbiers, Sjeff. 1995. *The syntax of interpretation*. Diss., University of Leiden/HIL.

³⁹ Recall that in English *root* questions, *wh*-phrases never seem to raise to SpecCP in overt syntax (see Den Dikken & Giannakidou 2002; Den Dikken 2003a); see also Bošković (2002) for related discussion of Serbo-Croatian.

- Beghelli, Filippo & Tim Stowell. 1997. Distributivity and negation. In Anna Szabolcsi (ed.), *Ways of scope taking*. Dordrecht: Kluwer.
- Bošković, Željko. 1997. *The syntax of nonfinite complementation*. Cambridge, MA: MIT Press.
- Bošković, Željko. 2002. On multiple *wh*-fronting. *Linguistic Inquiry* 33. 351–383.
- Broekhuis, Hans & Joost Dekkers. 2000. The Minimalist Program and Optimality Theory: Derivations and Evaluation. In Joost Dekkers, Frank van der Leeuw and Jeroen van de Weijer (eds), *Optimality Theory: Phonology, syntax and acquisition*. Oxford: Oxford University Press. 386–422.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2004. On phases. Ms., MIT.
- Chung, Sandra & James McCloskey. 1983. On the interpretation of certain island facts in GPSG. *Linguistic Inquiry* 14.
- Craenenbroeck, Jeroen van. 2006. Transitivity failures in the left periphery and foot-driven movement operations. Paper presented at the TiN-dag, Utrecht University.
- Dikken, Marcel den. 2003a. On the morphosyntax of *wh*-movement. In Cedric Boeckx & Kleantes Grohmann (eds), *Multiple wh-fronting*. Amsterdam: John Benjamins.
- Dikken, Marcel den. 2003b. A comment on the topic of topic-comment. *Lingua* 115. 691–710.
- Dikken, Marcel den. 2005. When Hungarians agree to (dis)agree — The fine art of phi and art. Ms., CUNY Graduate Center.
- Dikken, Marcel den. 2006. *Relators and linkers. The syntax of predication, Predicate Inversion, and copulas*. Cambridge, MA: MIT Press.
- Dikken, Marcel den & Anastasia Giannakidou. 2002. From *hell* to polarity: 'Aggressively non-D-linked' *wh*-phrases as polarity items. *Linguistic Inquiry* 33. 31–61.
- Doherty, Cathal. 1993. The syntax of subject contact relatives. Ms., University of California at Santa Cruz.
- Duffield, Nigel. 1993. On case-checking and NPI licensing in Hiberno-English. *Rivista di Linguistica* 7. 215–44.
- É. Kiss, Katalin. 1993. Wh-movement and specificity. *Natural Language and Linguistic Theory* 11.
- É. Kiss, Katalin. 1998. Identificational focus and information focus. *Language* 74. 245–73.
- Franck, Julie, Glenda Lassi, Ulrich Frauenfelder & Luigi Rizzi. 2003. Agreement and movement: A syntactic analysis of attraction. Ms., Université de Genève.
- Franks, Steven. in press. Agnostic movement. In C. Davis, A. Deal & Y. Zabbal (eds), *Proceedings of NELS 36*. Amherst: GLSA.
- Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12. 155–84.
- Grimshaw, Jane. 1997. Projection, heads and optimality. *Linguistic Inquiry* 28. 373–422.
- Guasti, Maria Teresa & Luigi Rizzi. 2002. Agreement and tense as distinct syntactic positions. Evidence from acquisition. In Guglielmo Cinque (ed.), *The structure of DP and IP – The cartography of syntactic structures, Vol. 1*. Oxford/New York: Oxford University Press.
- Hackenberg, Robert. 1972. *A sociolinguistic description of Appalachian English*. Diss., Georgetown University.
- Hazen, Kirk. 1996. Dialect affinity and subject-verb concord: The Appalachian-Outer Banks connection. *SECOL Review* 20. 25–53.
- Hazen, Kirk. 2000. Subject-verb concord in a post-insular dialect: The gradual persistence of dialect patterning. *Journal of English Linguistics* 28. 127–44.
- Henry, Alison. 1995. *Belfast English and Standard English: Dialect variation and parameter setting*. Oxford/New York: Oxford University Press.
- Horvath, Julia. 2000. Interfaces vs. the computational system in the syntax of focus. In Hans Bennis, Martin Everaert & Eric Reuland (eds), *Interface strategies*. Amsterdam: Royal Netherlands Academy of Arts and Sciences. 183–207.
- Horvath, Julia. 2005. Separating 'focus movement' from focus. Ms., Tel-Aviv University.
- Jespersen, Otto. 1961. *A modern English grammar on historical principles*. London: F. Allen & Unwin Ltd.

- Kallulli, Dalina. 2000. Direct object clitic doubling in Albanian and Greek. In Frits Beukema & Marcel den Dikken (eds), *Clitic phenomena in European languages*. Amsterdam: John Benjamins. 209–48.
- Kayne, Richard. 1998. Overt versus covert movement. *Syntax* 1. 128–91.
- Kenesei, István. 1986. On the logic of Hungarian word order. In Werner Abraham & Sjaak de Mey (eds), *Topic, focus and configurationality*. Amsterdam: John Benjamins. 143–59.
- Kimball, John & Judith Aissen. 1971. I think, you think, he think. *Linguistic Inquiry* 2. 241–46.
- Koeneman, Olaf. 2000. *The flexible nature of verb movement*. Diss., Utrecht University/LOT.
- Koopman, Hilda & Anna Szabolcsi. 2000. *Verbal complexes*. Cambridge, MA: MIT Press.
- Koster, Jan. 1986. *Domains and dynasties: The radical autonomy of syntax*. Dordrecht: Foris.
- Lechner, Winfried. 2005. Semantic and syntactic effects of head movement. Paper presented at the 28th GLOW Colloquium, Université de Genève.
- Lipták, Anikó. 2001. *On the syntax of wh-items in Hungarian*. Diss., University of Leiden/LOT.
- Maling, Joan & Annie Zaenen. 1982. A phrase structure account of Scandinavian extraction phenomena. In Pauline Jacobson & Geoff Pullum (eds), *The nature of syntactic representation*. Dordrecht: Reidel.
- Milroy, James. 1981. *Regional accents of English: Belfast*. Belfast: Blackstaff Press.
- Montgomery, Michael & Joseph Hall. 2004. *A dictionary of Smoky Mountain English*. Knoxville: University of Tennessee Press.
- Pesetsky, David. 1987. *Wh-in-situ* and unselective binding. In Eric Reuland & Alice ter Meulen (eds), *The representation of (in)definiteness*. Cambridge, MA: MIT Press.
- Pesetsky, David. 1989. Language-particular processes and the Earliness Principle. Paper presented at the 12th GLOW Colloquium, Utrecht, 5 April 1989; abstract in *GLOW Newsletter*, ms., MIT.
- Pesetsky, David & Esther Torrego. 2001. T-to-C movement: Causes and consequences. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*. Cambridge, MA: MIT Press.
- Platzack, Christer. 1996. Germanic Verb Second languages – Attract vs. Repel: On optionality, A-bar movement and the symmetrical/asymmetrical Verb Second hypothesis. In E. Lang & G. Zifonun (eds), *Deutsch – Typologisch*. Berlin: Walter de Gruyter. 92–120.
- Rackowski, Andrea & Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36. 565–99.
- Reinhart, Tanya. 1995. Interface strategies. *OTS Working Papers*, Utrecht University.
- Richards, Norvin. 1998. The Principle of Minimal Compliance. *Linguistic Inquiry* 29. 599–629.
- Riemsdijk, Henk van. 1995. Push chains and drag chains: Complex predicate split in Dutch. Ms., Tilburg University.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*. Dordrecht: Kluwer. 281–337.
- Ross, John Robert. 1967[1986]. *Infinite syntax!* Norwood, NJ: Ablex Publishing Corporation [published version of 1967 MIT dissertation].
- Schilling-Estes, Natalie & Walt Wolfram. 1994. Convergent explanation and alternative regularization: *were/weren*’t leveling in a vernacular English variety. *Language Variation and Change* 6. 273–302.
- Selkirk, Lisa. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Selkirk, Lisa. 1995. Sentence prosody: Intonation, stress, and phrasing. In John Goldsmith (ed.), *The handbook of phonological theory*. Oxford: Blackwell. 550–69.
- Sells, Peter, John Rickford & Thomas Wasow. 1996. Negative inversion in African American Vernacular English. *Natural Language and Linguistic Theory* 14. 591–627.
- Surányi, Balázs & Andrea Márkus. 2005. Vacuous movement in subject questions. Ms., Hungarian Academy of Sciences & Eötvös Loránd Tudományegyetem, Budapest.
- Szendrői, Kriszta. 2003. A stress-based approach to the syntax of Hungarian focus. *The Linguistic Review* 20. 37–78.

- Tagliamonte, Sali. 1998. Was/were variation across the generations: View from the City of York. *Language Variation and Change* 10. 153–93.
- Tagliamonte, Sali & Jennifer Smith. 1999. Analogical leveling in Samaná English: The case of was and were. *Journal of English Linguistics* 27. 8–26.
- Trotta, Joe. 2000. *Wh-clauses in English: Aspects of theory and description*. Amsterdam: Rodopi.
- Wedgwood, Daniel. 2005. *Shifting the focus: From static structures to the dynamics of interpretation*. Amsterdam: Elsevier.
- Wolfram, Walt & Donna Christian. 1976. *Appalachian speech*. Arlington, VA: Center for Applied Linguistics.
- Wolfram, Walt & Jason Sellers. 1999. Ethnolinguistic marking of past *be* in Lumbee Vernacular English. *Journal of English Linguistics* 27. 94–114.
- Zubizarreta, Maria Luisa. 1998. *Prosody, focus and word order*. Cambridge, MA: MIT Press.
- Zwart, Jan-Wouter. 2004. Een dynamische structuur van de Nederlandse zin. *Tabu* 33.

revised draft — 28 March 2006

Marcel den Dikken
Linguistics Program
CUNY Graduate Center
365 Fifth Avenue
New York, NY 10016-4309
MDen-Dikken@gc.cuny.edu
<http://web.gc.cuny.edu/dept/lingu/dendikken/index.html>
